Rationalisation of steel grades and specifications using machine learning techniques

A Thesis Submitted for the Degree of Doctor of Philosophy

By

Sadegh Jalalian

Brunel Centre for Advanced Solidification Technology,
Brunel University London

2025

This thesis is dedicated to my family.

DECLARATION

I, Sadegh Jalalian, declare that the work presented in this thesis is my own. Where the work of others has been used, it has been fully acknowledged and referenced.

London, September 2025

Acknowledgements

First, I would like to express my deepest gratitude to my principal supervisor, Professor Hamid Assadi. Without his guidance, support, and supervision, this project would not have been possible. He has been far more than a supervisor to me; he has also been a mentor, a coach, and a friend, standing by me throughout the struggles and challenges of my PhD journey. His way of seeing the world and his wisdom have shaped my thinking and allowed me to view life and research from new perspectives. I am truly grateful for the knowledge and values I have gained under his supervision. In particular, the development of the KMRP methodology could not have been achieved without his invaluable guidance and ideas.

I would also like to thank my second supervisor, Professor Isaac Chang, as well as Professor Dmitry Eskin and Dr. Chamini Mendis. From the very beginning of this project, their patience, guidance, and constructive feedback were instrumental in helping me to develop the methodology. Their expertise in metallurgy and their generous knowledge sharing provided a strong foundation for this work and greatly supported its progress from the earliest stages.

I would also like to extend my sincere thanks to Dr. Yousef Jalali for generously taking the time to read my thesis. His valuable comments helped me to refine and strengthen the final version.

I gratefully acknowledge UKRI CircularMetal, under the leadership of Professor Zhongyun Fan, for providing the funding that made this research possible. I would also like to thank all members of BCAST, with whom it has been an honour to collaborate and work during this project.

I thank the examiners, Professor Hari Nadendla and Professor Zushu Li, for their careful reading of the thesis and for their constructive comments and recommendations, which have helped to improve the quality and clarity of this work.

I would also like to thank Dr. Roohollah Babaei-Mahani and his wife, Mrs. Akhtar, for their constant support and encouragement during the final phase of my PhD. Their kindness, help, and emotional support meant a great deal to me and will always be remembered.

I am deeply thankful to my friends, who added colour to this journey. Their encouragement, support, and the memories we shared not only lightened the challenges but also made this period of my life far richer and more meaningful.

And finally, to my family and closest friends, who have always stood by my side and made countless sacrifices to allow me to pursue my dreams, I owe my deepest gratitude. To my father, who taught me to work hard, to be a good person, and to always help others. To my mother, who showed me patience, love, and a way of seeing life with kindness. To my sister, who taught me to be brave and patient in everything I do. And to my brother, who reminded me that I am good enough to follow my dreams and has stood beside me in every moment of my life. From the bottom of my heart, I thank you. This achievement does not belong to me alone; it belongs to all of you, for my life is inseparably bound to yours.

Abstract

There are an excessive number of steel grades currently in use. However, many of them are used in the same application despite differences in chemical composition and processing conditions, and in some cases shows equivalent ranges of properties. These huge number of grades poses challenges for sustainable recycling and increases production complexity and cost. This study introduces a multi-phase, application-driven framework to simplify the steel grade system and reduce the number of grades by proposing a novel approach called K-Means Reduction Process (KMRP).

The framework was applied to 148 carbon and 288 stainless steel grades, including chemical composition, processing conditions, and mechanical properties (hardness, UTS, YS, and elongation). Machine learning models were first used to quantify the influence of alloying elements and processing conditions on mechanical performance. K-Means clustering was then applied to group grades based on performance to identify steels that shared equivalent property profiles, with four distinct clusters identified including ferritic/low-carbon steels, medium-carbon and martensitic steels, high-carbon steels, and austenitic steels. These clusters revealed significant redundancy, with multiple grades from existing steel classifications occupying the same mechanical property space.

In the reduction phase, KMRP identified the minimal set of grades required to preserve full mechanical property coverage within the generated clusters. Two sustainability-driven strategies were implemented: (1) tramp-element avoidance, favouring grades with low Cu and Sn, and (2) tramp-element tolerance, prioritising grades compatible with scrap-based recycling. While both approaches reduced reliance on critical raw materials (Mo, Ni, V, Ti), this study focused on the tramp-tolerance strategy as the most relevant for advancing circular economy objectives. Under this approach, the number of carbon steel grades were reduced by 38.4% (from 146 to 90) and stainless steel grades by 52.8% (from 288 to 136), while fully preserving the original mechanical property ranges, including UTS ranges of 295–2450 MPa for carbon steels and 120–1970 MPa for stainless steels, and elongation ranges of 7–41% and 2–55%, respectively.

These results demonstrated that KMRP can successfully simplify the steel grade system while supporting circularity, reducing dependency on critical elements, and improving the sustainability of future steel production. Moreover, the methodology is generalisable and can

be applied to other domains where reducing redundant options is essential, such as pharmaceutical applications.

Keywords: Circular economy, steel classification, clustering, reduction process, SHAP

Table of contents

Data-Driven Insights and Interpretation

Steel Behaviour and Rationalisation Outcomes

Decoding Results: Analytical Interpretation and Knowledge Extraction

List of Figures

List of Tables

List of Abbreviations

Metallurgy and Standards

AISI	American Iron and Steel Institute
ASTM	American Society for Testing and Materials
EN	European Norm (European Standard)
GBT	Guobiao Standards (Chinese National Standards)
UTS	Ultimate Tensile Strength
YS	Yield Strength
CRM	Critical Raw Materials

Data Science

ML	Machine Learning
AI	Artificial Intelligence
ANN	Artificial Neural Network
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
HC	Hierarchical Clustering
KMRP	K-Means Reduction Process
PCA	Principal Component Analysis
RF	Random Forest
SHAP	SHapley Additive exPlanations
XGBoost	Extreme Gradient Boosting
t-SNE	t-distributed Stochastic Neighbour Embedding
MAE	Mean Absolute Error
XAI	Explainable Artificial Intelligence

Chapter 1: Introduction

1.1 Project Overview

Our modern world is built on materials and they are everywhere. However, what if the various materials we rely on to build our world piece by piece, moving our lives along and powering our industries are silently undermining our ability to sustain the planet? Why do we use thousands of different materials for similar tasks? Is it because we truly need them, or have we accepted a lot of unnecessary complexity?

Steel is one of the examples of this challenge as it is widely used and economically important material in the world. It has played a pivotal role in industrial development since the 19th century and remains essential to modern infrastructure, transport, energy, and manufacturing (M., 1922). From a materials perspective, steel is an iron-based alloy whose properties are governed by the presence and concentration of additional alloying elements. Accordingly, steels are traditionally classified into broad types such as carbon, alloy, tool, and stainless steels based on characteristic ranges of alloying elements (e.g. C, Mn, Cr, Ni) (Bailey, 1994), as these compositions control properties such as hardness, durability, and formability (Reijnders, 2016). Building on this traditional classification framework, standard steel grades are specified within national and international standards, such as AISI, EN, and ASTM, where each grade is characterised by defined chemical composition limits, processing routes and associated property requirements.

According to the World Steel Association, steel production has been on the rise, with an estimated annual production of 1.96 billion metric tons in 2021 (World Steel Association, 2021). However, as demand continues to rise, concerns over the environmental and resource impacts of steel production are growing (Conejo et al., 2020). The steel industry worked under a linear model of production which is including the extraction, process, uses based on demand, and ultimately discard as a scrap (Hagedorn et al., 2022). The current model is putting a strain on resources, causing environmental damage, and generating a lot of waste. Also, there is the issue of disappearing steel, where the steel does not re-enter circulation, which some research has estimated to be at a rate of at least twelve percent of steel produced, probably higher (Bowyer et al., 2015).

Therefore, it is time to start thinking differently and shift towards a more circular approach. This shift involves not only improving recovery and recycling processes but also rethinking

the underlying structure of existing steel classification, questioning whether the vast number of steel grades currently in use is truly necessary. This complexity complicates scrap sorting, increases the risk of cross-contamination, and ultimately limits the reusability of recovered steel due to incompatible chemical compositions (Reck and Graedel, 2012; Björkman and Samuelsson, 2014). This research builds on that idea by investigating how the excessive diversity of steel grades can be reduced to avoid unnecessary complexity in production and supply chains while improving recycling efficiency and reducing environmental impact. In what follows, the motivation, aims and objectives, as well as the overall structure of this dissertation study will be discussed.

1.2 Motivation

Currently, there are more than 3,500 steel grades, each specifying detailed combinations of chemical composition, processing routes, and mechanical properties for particular applications (Javaid and Essadiqi, 2003). While the broad classification of steels into carbon, alloy, stainless, and tool steels has historically provided a useful framework, the existence of thousands of individual grades within these categories has led to a level of complexity that complicates grade selection, portfolio management, and end-of-life recycling.

Beyond the excessive number of grades, recent literature has increasingly pointed to alloy chemistry as a fundamental constraint on circularity (Daigo et al., 2017; Daigo and Goto, 2015; Dworak et al., 2022). In particular, the presence of critical raw materials (CRMs) and the sensitivity of many steel grades to tramp elements have emerged as key barriers to efficient scrap utilisation and large-scale recycling. As steel production shifts towards higher scrap fractions, these chemical constraints become increasingly difficult to manage within existing steel system.

At the same time, in the existing steel groups, many grades show overlapping property ranges despite having different compositions and processing routes, as documented for structural and automotive steels (García Gutiérrez et al., 2024). For instance, a large number of grades are used in common applications such as construction and automotive, despite offering similar ultimate tensile strength and elongation requirements at the product level. Excessively strict alloy specifications and unnecessary alloying components raise production costs, decrease resource productivity, cause environmental damage, and make recycling end-of-life products challenging due to chemical incompatibility between grades, which is not conforming with

circular economy principles where material recovery, reuse, and simplification are essential (Raabe, 2023a).

From an engineering perspective, the practical role of a steel grade is ultimately determined by its performance rather than its designation alone. Grades can therefore be compared based on the combinations of properties they possess, such as ultimate tensile strength, yield strength, hardness, and elongation. When multiple grades show similar combinations of these properties, they occupy the same region of mechanical behaviour and can fulfil equivalent application requirements despite differences in composition or standard designation. The ability of a selected subset of grades to collectively span the full range of mechanical properties required by applications is referred to in this study as mechanical coverage. This performance-based viewpoint provides a systematic basis for identifying redundancy within existing steel grade systems and motivates the need for a rational, data-driven reduction framework.

1.3 Significance

Despite recent progress in alloy design and materials informatics (Xiong et al., 2020; Guo et al., 2019; Lee et al., 2021; Yan et al., 2020; Xie et al., 2021), few studies have systematically assessed which grades are essential from a performance perspective. Most existing work focuses on predicting the properties, not reducing or rationalising the existing number of grades (Rocha and Marques, 2024). Moreover, steel classification has limited consideration through the lens of recyclability or environmental impact.

This reveals a clear research gap that there is no quantitative, data-driven framework that operates on current EN/ASTM/AISI/GB grades, identifies functionally equivalent steels based on mechanical performance and incorporates tramp element tolerance and CRMs use when deciding which grades are necessary and which are redundant. Filling this gap is significant because rationalising the number of steel grades, and favouring grades with higher tolerance to tramp elements and reduced CRM ranges, can make it easier to design recycling strategies and scrap-use policies that maintain mechanical performance while operating with a smaller, more impurity-robust set of grades.

Given the complexity and volume of available steel grade data and different dimensions of steel in terms of chemical, processing and properties, which will be discussed in chapter 2, traditional manual or expert-based methods are insufficient and challenging for identifying redundancies or patterns across grades. Therefore, the recent advances in AI and data science

present an opportunity to systematically identify functionally equivalent grades and rationalise the system toward fewer, more recyclable options, without compromising engineering applications.

1.4 Aim and Objectives

This study aims to propose an analytical basis for reclassification of steels into a significantly smaller number of grades than the currently existing 3,500 grades, to improve recyclability without compromising performance and production costs. It focuses only on mechanical properties due to limited access to standardised data on other characteristics such as corrosion resistance or weldability. Nevertheless, mechanical performance remains the primary driver in structural applications, making it a valid basis for rationalisation. To achieve the aim of this study the following objectives have been identified:

(1) Designing and creating a comprehensive steel grade database from multiple international standards.

(2) Quantification of the correlations between the alloy composition and processing conditions on one side, and the material performance on the other side using supervised learning approaches.

(3) Developing a novel, purpose-driven reclassification framework, the K-Means Reduction Process (KMRP), that systematically rationalises the existing steel system by integrating data analysis, machine learning, and optimisation techniques.

(4) Identification of recyclability-limiting alloying elements that can be replaced or eliminated by detection of identical grades that contain similar mechanical property coverage but differ in chemical composition and processing conditions using KMRP.

(5) Simplification of alloy systems based on their compositions, levels of performance, fields of application and reduction of reliance on CRMs and enhancement of tolerance to tramp elements by selecting grades with favourable recyclability profiles using KMRP.

1.5 Contribution and Impact

By focusing not only on property prediction and alloy design, which dominate much of the existing machine-learning literature on steels, but also on rationalisation, this work represents a novel contribution to both machine learning in materials science and the broader field of sustainable, data-driven decision support systems. It is proposing a multi-phase framework that

combines clustering, data reduction, and SHAP analysis to simplify alloy systems without sacrificing performance. Specifically, it introduces a novel hybrid algorithm, the KMRP, which extends conventional clustering by incorporating range-based property overlap checks and sustainability filtering criteria such as tramp element avoidance/tolerance and CRMs minimisation. Unlike existing unsupervised methods that only identify groups, KMRP adds a selection and reduction mechanism, enabling the practical reduction of steel grades within each cluster while ensuring full performance coverage. This structured approach enables users to retain only the most efficient, recyclable grades for a given application.

Furthermore, the development of a data-driven framework capable of reducing the number of steel grades offers significant benefits to industry and policy-makers. By identifying and promoting grades with enhanced recyclability, the framework supports more efficient scrap utilisation and contributes to the decarbonisation of the steel sector. This aligns with broader efforts to shift from the carbon-intensive Blast Furnace–Basic Oxygen Furnace (BF–BOF) route toward the lower-emission Electric Arc Furnace (EAF) process, which emits substantially less CO_2 per tonne of steel produced.

Beyond its environmental impact, the proposed framework can be embedded into decision-support tools or integrated into industrial software platforms. This would enable engineers and scientists to access a unified alloy dataset, analyse and detect similar alloys and their interchangeability, and streamline design choices, which can be informed by software tools like CALPHAD that can investigate and validate the designed alloy before laboratory experimentation.

The methodology is also transferable to other domains where reducing complex options is essential, such as drug discovery, by providing a scalable approach to simplify decision-making in multi-parameter systems.

1.6 Thesis Structure

In this chapter, the background, motivation, overarching aim, and specific objectives of the study have been introduced. The remainder of the thesis is organised as follows.

- **Chapter 2** provides a review of sustainability and recyclability in metals, with a focus on tramp elements and alloy-specific challenges. It then examines key aspects of steel, including its properties, existing grade systems, and classification standards. This

chapter establishes the industrial and environmental context of the study and highlights the inefficiencies inherent in the current steel system.

- **Chapter 3** reviews the role of machine learning in materials science, covering both supervised and unsupervised approaches, alongside explainable AI and data reduction techniques. This chapter identifies the methodological tools most relevant to the project and clarifies the research gap in their application to systematic steel grade rationalisation.
- **Chapter 4** presents the first methodological phase, describing the construction of a comprehensive steel dataset, exploratory analysis of its features, and the application of supervised learning coupled with explainable AI to uncover relationship between chemical composition, processing conditions, and mechanical properties.
- **Chapter 5** introduces the second methodological phase, KMRP. It details the clustering and reduction framework proposed to group steels with overlapping property ranges and to reduce redundancies while preserving performance and recyclability.
- **Chapter 6** presents the results of supervised learning applied to carbon and stainless steels, assessing both predictive performance and the metallurgical insights gained from model interpretation.
- **Chapter 7** presents the results of applying the KMRP framework, showing how steel grades can be reclassified into fewer representative groups without compromising property coverage.
- **Chapter 8** extends the analysis by interpreting KMRP outcomes under a tramp element tolerance strategy, demonstrating how rationalisation can also support sustainability objectives.
- **Chapter 9** concludes the thesis, summarising the key findings, discussing limitations, and identifying future directions for research and industrial application.

Chapter 2: Literature Review - Steel, Sustainability and Classification

As outlined in chapter 1, the existence of a large number of steel grades presents significant barriers to effective recycling, and resource efficiency because excessive grade diversity complicates scrap sorting, increases the risk of cross-contamination by tramp elements, and ultimately downgrades the quality of recovered steel during remelting. To address this, the current chapter provides a review of the relevant literature on sustainability and recyclability, with a particular focus on metal recycling strategies, the impact of tramp elements in recycling of metals, and the sustainability implications of specific alloying elements. The structure of the review progresses from high-level sustainability concepts to more technical issues that directly influence alloy recycling outcomes.

In addition, this chapter presents a detailed overview of the steel industry to contextualise the significance of existing steel categories in the current market, which shaped the data collection strategy of this study. The relationship between steel properties and their underlying chemical and processing parameters, followed by a critical review of current steel classification system is discussed. This foundation establishes the need for a more streamlined and data-driven approach to rationalising steel grades. To support this, the chapter concludes with an examination of classification principles from broader domains, where machine learning methods have been used to group structurally or functionally similar entities.

2.1 Sustainability and Recyclability

2.1.1 Overview of Circular Economy

Circular economy is a production and consumption model that obtains reusing, repairing and recycling the existing material or products as long as possible (Lüdeke-Freund et al., 2019). In general, the circular economy seeks to address major global issues like climate change, waste, and pollution by focusing on the design implementation of the three fundamental principles which obtains eliminating waste and pollution, circulating materials, and regenerating the natural world (Circular economy overview, 2021).

There has been an increasing number of sustainable products and methods being developed as a result of growing public concern over international environmental and social challenges. The concept of sustainable development was first introduced in the 1987 under the report of “Our Common Future” by the World Commission on Environment and Development (WCED)

(Commission, 1987). The main objective of sustainable development is "Improving economic and social welfare without compromising the capacity of future generations to do the same". Therefore, products must be made to be robust, simple to repair, and ultimately recyclable.

Recycling plays a critical role in advancing sustainable development, as it offers an alternative to traditional disposal methods and helps reduce environmental impact. Recycling is the process of turning waste material into new products and materials. This idea frequently takes the recovery of energy from waste materials into account. The capacity of a substance to regain the qualities it had in its initial state determines how recyclable it is (Villalba et al., 2002). This process assists conserve resources and reduce greenhouse gas emissions. In addition, it can limit the consumption of raw materials and stop the waste of potentially usable materials, which will cut down on energy use, air and water pollution.

2.1.2 Recycling of Metals and Alloys

According to the circular economy model, products must be made to be robust, simple to repair, and ultimately recyclable (Hapuwatte et al., 2025). Steel recycling is crucial to the circular economy because it saves precious resources and prevents usable materials from ending up in landfills. Iron and steel production accounts for approximately 7-9% of total global CO_2 emissions (Kim et al., 2022; De Ras et al., 2019).

Although the magnetic properties of steel make it easy to separate from mixed waste streams using simple magnetic sorting techniques (Veasey, 1993), this is not the primary reason steel is widely recycled. The main driver is the substantial reduction in energy use and CO_2 emissions when steel is produced from scrap instead of iron ore. Murray et al., (2024) reported that recycling could lead to savings of 1% of global CO_2 emissions. Manufacturing steel through scrap-based routes, particularly when processed in Electric Arc Furnaces (EAF), requires less energy than primary production and helps avoid the environmental and economic burdens associated with iron-ore extraction and ore-based steelmaking. When high-quality scrap is available, the reduced energy demand can also lower production costs, although the actual cost benefit depends on scrap prices, electricity tariffs, and product specifications.

Steel recycling is generally occurring in two ways including open-loop and closed-loop recycling (Nakamura et al., 2014). In open-loop recycling, recycled steel is used to make different products than it was originally intended for. This usually happens because the quality of the recycled steel may degrade or get contaminated during the recycling process. As a result,

new raw materials often need to be added to meet quality standards. On the other hand, closed-loop recycling, involves reusing recycled steel to make the same type of product, helping maintain the material's quality. This method is more environmentally friendly because it reduces waste and the need for new raw materials. Both strategies offer environmental benefits, and the steel industry has developed life-cycle assessment frameworks to show how recycling helps lower CO_2 emissions, reduces energy use, and cuts down on the need for iron ore compared to primary steelmaking routes (Broadbent, 2016).

Steel recycling is typically accomplished through a combination of recovery and remelting processes. At the collection stage, mechanical separation techniques such as magnetic sorting are applied to extract ferrous metals from mixed waste streams (Veasey, 1997). This is followed by sorting, cleaning, and preparation processes to ensure the scrap meets quality standards. The prepared scrap is then processed using the EAF route, which is the method for secondary (scrap-based) steelmaking. Compared with the conventional primary route (BF-BOF process), EAF steelmaking requires significantly less energy because it relies on recycled scrap rather than iron ore (Ohno et al., 2015a). In this process, scrap steel is placed in the furnace and melted using electric arcs. To clean the molten steel, materials like limestone or dolomite are added, which help remove impurities such as sulfur, phosphorus, and excess carbon (Teo et al., 2020). These impurities form a layer of waste material called EAF slag that floats on top of the molten steel. During melting, gases like carbon monoxide are released, causing the molten steel to boil and helping to remove unwanted elements like nitrogen and hydrogen. Once the steel is refined, it is poured out from the furnace, and the slag is collected separately. The slag cools and hardens into a rock-like material, which can be reused in construction or other industries. The EAF process is energy-efficient because it recycles existing steel instead of using raw iron ore. This reduces energy use and greenhouse gas emissions, making it an environmentally friendly option for steel production (Kishore et al., 2025).

Despite the well-established strategies and methods for steel recycling. However, some parameters make this process challenging. One of the most persistent barriers to high-quality recycling is the presence of tramp elements. These elements not only limit the potential applications of recycled steel but also constrain the possibility of achieving circular steel flows. In this study, understanding tramp elements is essential, as their presence directly affects decisions on evaluating steel grade suitability for retention or elimination during rationalisation.

2.1.3 Tramp Elements

Steel can be alloyed with various elements, such as carbon (which increases hardness and strength and improves hardenability), to enhance its characteristics. However, there are some elements called tramp elements that end up in steel by accident and are hard to remove (Worrell and Reuter, 2014). These elements remain in the steel cycle and keep reappearing each time the steel is recycled. Some tramp elements harm the properties of steel, while others are helpful for specific types of steel but not all. Iron reacts with purifying agents before tramp elements do, which makes it tough to remove them. Currently, there are no practical ways to remove these tramp elements from steel. In addition, some of these tramp elements, such as chromium, nickel, are intentionally added as alloying elements in specific steel grades.

A recent study on steel recycling in Europe highlighted the growing challenge posed by tramp elements in post-consumer steel scrap (Dworak et al., 2022). As steel recycling rates increase, particularly in regions with saturated steel markets like the EU, the accumulation of tramp elements such as copper, nickel, chromium, and tin are becoming a significant issue. The study projected that post-consumer scrap will rise from 80 million tonnes per year in 2020 to over 100 million tonnes by 2050. However, much of this scrap will be of lower purity due to the increasing presence of tramp elements, leading to a surplus of low-grade scrap that exceeds demand.

In another study, Daigo et al., (2017) focused on understanding how tramp elements accumulate in carbon steel during recycling in Japan. They analysed over 500 samples of carbon steel from various product forms like bars, plates, and sheets. They found that certain tramp elements, particularly nickel and molybdenum, made up more than 15% of Japan's annual consumption of these metals just through contamination in recycled steel. Copper was another major contaminant, accounting for about 4% of Japan's total electrolytic copper production, while significant amounts of tin were also found, mainly coming from tin-plated products. The study highlighted several challenges in managing tramp elements. For instance, copper and tin showed high levels of consistency across different recycled steel samples, indicating they are difficult to avoid during recycling. Although chromium theoretically should be removable during the smelting process, in practice, a considerable amount remains in the recycled steel, sometimes causing issues during manufacturing processes like deep drawing.

One of the key concerns is that steel made from end-of-life products tends to have higher concentrations of these tramp elements, which can lower the overall quality of the recycled

steel. However, as the accumulation of tramp elements in scrap is unavoidable, selecting or prioritising steel grades with inherently higher tolerance to these elements can enhance long-term recyclability. By favouring such grades at the design stage, the steel remains suitable for future recycling cycles without compromising material performance.

While tramp elements are a barrier to recycling, they represent only one aspect of the broader sustainability challenges associated with steel. Beyond their impact on recyclability, certain elements also pose environmental, economic, and geopolitical concerns due to their criticality, regional scarcity, or toxicity. Furthermore, some valuable alloying elements may be partially lost during recycling and leads reducing material efficiency (Charles et al., 2020).

These factors must be considered when evaluating steel grades for long-term circularity. The following section examines these elements and their effect on the steel to shape the rationalisation criteria of this study.

2.1.4 Element and Alloy-specific Sustainability Challenges

The presence or increasing concentration of specific elements can have both beneficial and harmful influence on the properties of steel. For instance, a higher chromium and nickel content can make recycling more challenging, while being beneficial for oxidation and corrosion resistance (Ohno et al., 2017). Also, copper and tin are the primary tramp elements (Daehn et al., 2017), which commonly degrade the quality of the recycled steel (Daigo and Goto, 2015; Ohno et al., 2015). In addition, manganese is a deoxidising agent that is added to steel, also to increase hardenability and strength. However, a high manganese content in steel can make it difficult to weld and machine, which may restrict how easily it can be recycled (Hagelstein, 2009). Vanadium is considered as a tramp element, while improving the wear resistance and being listed as a critical raw material (Blengini et al., 2020).

In addition to elements that affect recyclability, some also raise broader sustainability concerns. Certain alloying elements are classified as CRMs, which are essential due to their economic importance and supply risks. Tomala and Urbaniec, (2024) highlighted the vital role of CRMs in supporting clean technologies that drive the EU's transition toward a low-carbon and eco-innovative economy. In the context of steel production and recycling, alloying elements like chromium, nickel, and vanadium which many of them are classified as CRMs and are frequently used to enhance steel properties. However, their scarcity and criticality emphasise the need to keep these valuable elements within the recycling loop.

Despite the numerous sustainability concerns in steel recycling, such as the accumulation of tramp elements, the limitations of closed-loop recovery, and the growing dependence on CRMs, a pivotal question remains. Can we identify grades within existing standards that demonstrate comparable performance, and prioritise those that either tolerate tramp elements and minimise CRM content without compromising the mechanical properties required for their intended applications?

By identifying redundant or overlapping grades and simplifying the current system, it is possible to reduce classification complexity, prioritise more recyclable options, and lessen dependence on rare alloying elements. The smallest number of grades could also improve sorting efficiency, enable more consistent closed-loop recycling, reduce the need for corrective remelting adjustments, and minimise material waste and downcycling, which ultimately supports a more circular and resource-efficient steel economy.

However, to streamline the current steel system, it is essential to understand the current structure of the steel industry, including the types that are most commonly used and in which sectors. The following sections examine global steel production trends and classification systems to contextualise the rationale for simplification and identify the limitations of existing classification systems.

2.2 Steel

2.2.1 Big Picture of Steel Production

In this section, the existing data from World Steel Association, (2021) has been analysed to gain insights and information about the current state and trends of the global steel industry.

Crude steel is the initial form of steel before it is rolled into sheets or shaped into beams. It typically requires further processing and refining to produce finished steel products. Figure 2.1 shows the crude steel production from 1950 to 2021. The graph shows that crude steel production has increased dramatically over the past 70 years. In 1950, around 189 million tonnes of crude steel were produced. By 2021, that number had risen to over 1,951 million tonnes. As shown in the graph, the steel production raised dramatically after 2000 and can be attributed to several key factors such as rapid economic growth in developing countries, particularly China experienced a period of immense economic expansion after 2000 (Yu et al., 2020). This fueled a massive demand for steel in construction, manufacturing, and other

sectors. Also, many developing nations invested heavily in infrastructure projects like bridges, roads, and power grids after 2000. This again led to a significant increase in steel consumption.

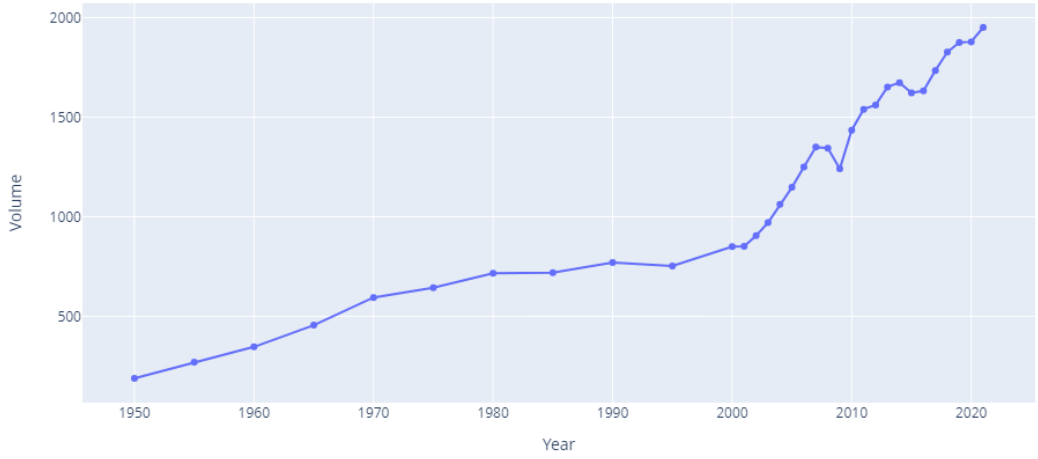


Figure 2.1: Line-plot world crude steel production from 1950 to 2021 (million tonnes). Data source from World Steel Association, (2021). Plot generated by the author.

Figure 2.2 demonstrates the major steel producing countries in 2020 and 2021. The map plot shows that China is the world's largest steel producer by a significant margin, accounting for more than half of global steel production. Besides China, the top steel-producing countries include India, Japan, US, Russia, South Korea, Turkey, Germany, Brazil, and Iran, all of which are significant players in the steel industry.

Figure 2.2: Major steel-producing countries 2020 and 2021 (in million metric tons). The size of the circle corresponds to the volume of steel produced by that country. Data source from World Steel Association, (2021). Plot generated by the author.

Any attempt to reduce the number of grades must be informed by market demand, functional diversity, and performance requirements to ensure that widely used and versatile grades are preserved while redundant or niche variants are reduced. Figure 2.3 shows the volume of various steel products exported globally from 2019 to 2021. The most popular exported steel

products were hot-rolled sheets and coils, with more than 70 million metric tons exported during this period. Hot-rolled sheets and plates provide a rougher surface than cold-rolled steel and are used in applications where corrosion resistance is necessary, such as in building and construction, automotive and transportation, and shipping containers (Imam et al., 2016). The second and third most popular products are ingots and semi-finished materials, which are useful for storage and transportation (Schulte et al., 2020), followed by galvanised sheets, which are primarily used in the manufacture of automobiles, refrigerated containers, construction sections, HVAC facilities, and furniture (Thierry et al., 2018). Importantly, these product forms can be manufactured from a diverse range of steel grades, depending on the desired mechanical properties and application requirements. However, when multiple grades are employed to produce similar product forms serving overlapping functions and opportunities arise for rationalisation.

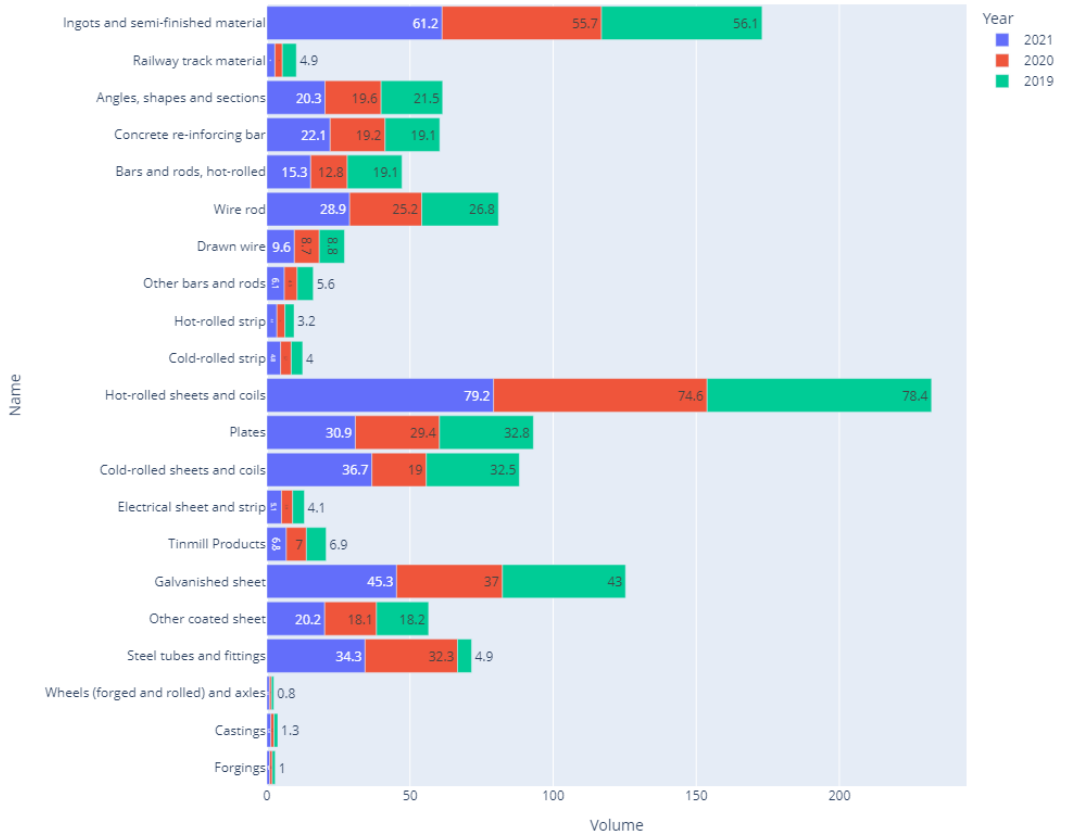


Figure 2.3: Bar plot of global steel exports by product category for 2019–2021. Bars are segmented by year (2019, 2020, 2021), with volumes shown in million metric tons. Data source from World Steel Association, (2021). Plot generated by the author.

In addition, Figure 2.4 shows the production of crude steel in the European Union from 2012 to 2020, categorised by quality. This bar plot indicates that carbon steel production (both non-alloy and other alloy combined) makes up a larger portion of the EU's crude steel production compared to stainless steel. Carbon steels are steels in which carbon is the main element influencing strength, with carbon content typically below 2% with minor quantities of other elements such as manganese and silicon (Davis, 1996). They are widely used in structural engineering, automotive components, and machinery due to their good balance of strength, ductility, and cost-effectiveness. In contrast, stainless steels contain at least 10.5% chromium, which provides corrosion resistance and makes them suitable for applications in chemical processing, food equipment, and medical instruments. A detailed overview of these steel categories is provided in Section 2.2.3.

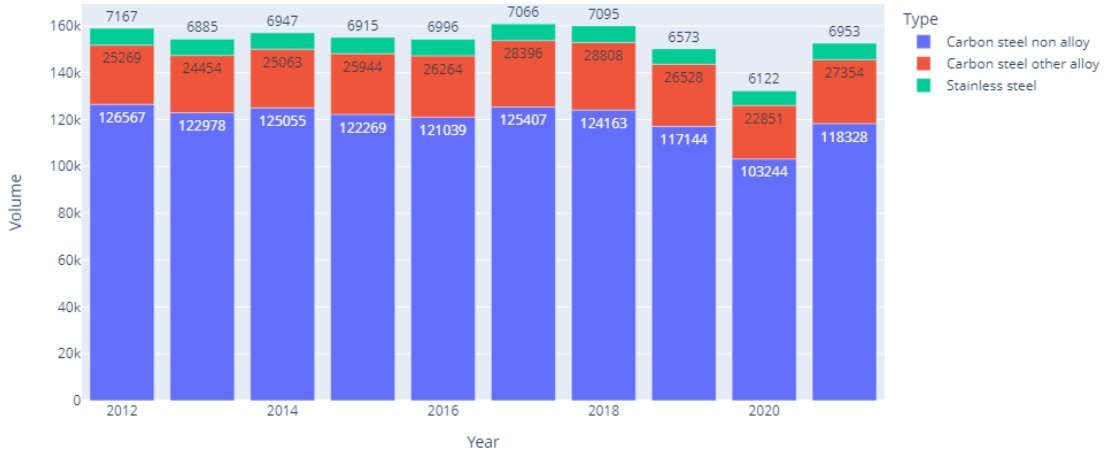


Figure 2.4: EU crude steel production by quality (2012–2021). Bars show annual production volumes (thousand tonnes) for stainless, non-alloy carbon, and other alloy carbon steels. Data source from World Steel Association, (2021). Plot generated by the author.

Figure 2.5 shows the EU steel consumption per sector for 2020 and 2021. The bar plot indicates that steel is mostly used in a wide range of construction applications, including rebar, beams, and sheets for roofing and wall panels. The second largest application of steel is in the automotive industry (Singh, 2016), where it is used in chassis, body panels, and engine components. The third most common application of steel is in mechanical engineering. This can be linked to the EU steel production where carbon steels are mostly used in constructions.

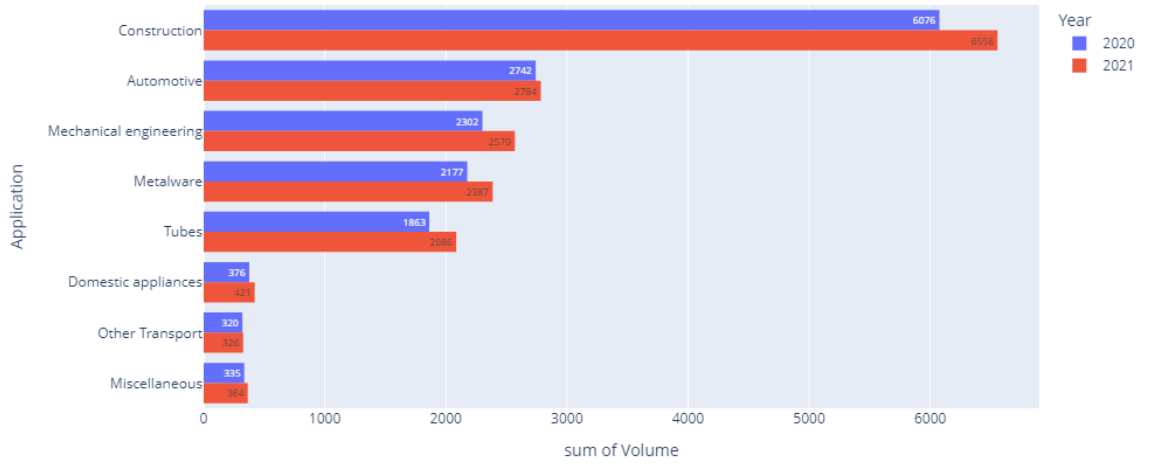


Figure 2.5: EU steel consumption by economic sector for 2020 and 2021. Bars represent total annual consumption per sector, separated by year. Data source from World Steel Association, (2021). Plot generated by the author.

Based on the global analysis of steel production, carbon and stainless steels account for the majority of international and EU output and are applied across a broad range of industries, from construction to automotive.

This contextual understanding is essential because it shaped the data collection of these two popular categories for this study. For rationalising the steel grade, it is essential to understand which steel properties most directly relate to application requirements. The following section presents these key steel properties, providing a foundation for evaluating which grades are necessary to retain and which could be reduced.

2.2.2 Steel Properties

In the context of reducing the number of steel grades without compromising performance, it is essential to understand which material properties are critical for end-use applications and how they are influenced. Steel properties arise from a complex interaction between chemical composition and processing conditions (Wiklund et al., 2017), which together determine a material's mechanical behaviour, corrosion resistance, and suitability for forming or welding. As discussed in section 2.1.2, these factors not only directly affect the performance but also impact production costs, recyclability, and compatibility across various applications. Therefore, identifying key properties and their governing inputs provides a foundation for this study.

The properties that are most important in steel from construction to manufacturing contain (Srikanth and Asmatulu, 2013):

Hardness**:** Hardness is the ability to withstand localised deformation and may indicate resistance to scratching and abrasion.

Ultimate tensile strength (UTS): Tensile strength is the maximum point for withstanding the material while being stretched or pulled. In other words, Tensile strength is the measurement of the required force to break the material.

Yield strength (YS): Yield strength is the maximum stress that can be applied before it begins to change shape permanently. In other words, it is the point that the material will no longer return to its original shape or dimension.

Elongation: elongation is the stretch percentage of the original length of the steel to the point of failure. In other words, elongation shows how ductile the steel is. Ductility determines the ability of a material to stretch out without becoming weaker in the process. The more ductility leads, the more formable product.

Module of elasticity: Elasticity is the ability of a material to deform under load and return to its original shape when the load is removed (Hooke, 1678). This behaviour is described by the relationship between stress (force per unit area) and strain (deformation relative to the original length). The modulus of elasticity or Young's modulus quantifies stiffness and is defined as the ratio of stress to strain within the linear elastic region of the stress–strain curve, where the material follows Hooke's law. For steels, a high modulus indicates strong resistance to elastic deformation under applied loads.

Machinability: It refers to how easily a metal material can be cut or machined. The weight of various criteria, such as cutting speed, surface finish, and tool life, is computed to determine machinability. Machinability is influenced by several factors, including hardness, microstructure, and the presence of free-cutting elements such as sulphur. Although higher carbon content can increase hardness, and may therefore reduce machinability, carbon is only one of several contributing factors. Different alloying additions are used to improve machinability, with sulphur being one of the most effective (Martin, 2006).

Among the above properties, the mechanical properties, particularly hardness, UTS, YS, and elongation, are most frequently specified in steel standards. These parameters define the strength, formability, and ductility of steel, serving as key criteria that engineers and industry professionals use when selecting suitable steel grades for specific applications.

However, Different criteria can affect the steel properties such as combination of chemical elements and also processing conditions.

2.2.3 Chemical Composition and Existing Steel Group

In manufacturing, different combination of chemical elements under different processing parameters influences the steel properties. For instance, carbon is the most important element in steel, and increasing the carbon content, increase the steel's strength and reduce its ductility (Zheng et al., 2010). Also, a small amount of chromium in combination with copper and nickel may increase steel corrosion resistance. Besides that, copper and molybdenum that are classified as tramp element and as CRMs, are mostly used in steel grades to increase the resistance to corrosion (Owen et al., 1976). However, as discussed in section 2.1, excessive alloying can increase production costs and create sustainability challenges. Their presence can limit recyclability and contribute to material downgrading during steel recycling. Thus, from a reclassification perspective, understanding which elements are essential for performance and which can be reduced is key to optimising both cost and circularity.

Steel can be classified in different ways, depending on the intended purpose, such as chemical composition, microstructure, processing method, product form, or standard designation system. One commonly used metallurgical approach groups steels into four categories based on their chemical composition, including carbon steels, alloy steels, stainless steels, and tool steels. This composition-based classification is widely recognised in the materials engineering literature and distinguishes these categories by their primary alloying elements and typical chemical ranges (Black, 2008; Kalpakjian et al., 2013).

Alloy steels: In this group, Different chemical elements besides carbon are used to improve steel's mechanical properties. Manganese, chromium, nickel, silicon, and molybdenum are common alloying elements that affect the steel's strength or corrosion resistance (Black and Kohser, 2017).

Tool steels: They are specially designed to offer high hardness, excellent wear resistance, and the ability to maintain their shape at elevated temperatures. They typically contain alloying elements such as tungsten, molybdenum, vanadium, or cobalt. These steels are widely used in the manufacturing of hand tools, cutting instruments, and machine dies (Trautman et al., 2023).

Carbon steels: In this group, Carbon is the main chemical element besides iron and any other element. In addition, Carbon steels are very popular in production due to their enough strength

and cheap manufacture (Papavinasam, 2014). Carbon steels are generally classified based on their carbon content into three different categories including low where carbon content less than 0.25%, medium where carbon content between 0.25% and 0.5% and high carbon steels where carbon content in this class is between 0.5% and 1.25%. The maximum carbon content in carbon steels is 2%, and any carbon content higher than 2% is classified as cast iron (Sotoodeh, 2022).

Stainless steels: Besides nickel, silicon, manganese, and carbon, stainless steels also contain 10-20 percent chromium as an alloying element. Stainless steels have a high resistance to corrosion (Britannica, 2021) and are categorised into four different types including martensitic, ferritic, duplex, and austenitic types which are widely used in chemical, oil and gas industries (Kangas and Chai, 2013).

Austenitic stainless steels are the most common type of stainless steel, accounting for around two-thirds of global stainless steel production (Michler, 2016). They have a unique austenitic microstructure, which is achieved by adding enough nickel, and sometimes manganese and nitrogen, to keep this structure stable across a wide temperature range.

Ferritic stainless steels have a ferrite microstructure similar to the carbon steel (Dutta, 2018). It typically contains between 10.5% and 27% chromium but has little to no nickel. Unlike austenitic stainless steel, ferritic stainless steel cannot be hardened through heat treatment and does not respond as well to strengthening by cold working. However, like carbon steel, it is magnetic, which sets it apart from the non-magnetic austenitic grades.

Duplex stainless steels have a mixed microstructure made up of both austenite and ferrite, which is why they are sometimes called austenitic-ferritic steels. Their composition usually includes 20-26% chromium, 1-8% nickel, 0.05-5% molybdenum, and 0.05-0.3% nitrogen. This combination gives duplex steels higher strength than standard austenitic stainless steels, while duplex stainless steels offer good ductility, their increased strength limits their formability compared to austenitic grades.

Martensitic stainless steels share a similar body-centered cubic (BCC) structure with ferritic stainless steel and carbon steels, but their higher carbon content, typically between 0.1% and 1.2%, allows them to be hardened through heat treatment. These steels are typically used in a hardened and tempered state, which gives them high strength along with moderate corrosion resistance. Martensitic stainless steels are ideal for applications like cutlery, surgical instruments, industrial knives, wear plates, and turbine blades. However, compared to ferritic,

austenitic, and duplex stainless steels, martensitic grades are generally less ductile (Guide, 2013).

Besides the chemical composition combination, different process parameters can influence the steel properties by modifying the shape of the iron atom's structures (Liang et al., 2022). The following section describes all the different processes and their effect on steel properties.

2.2.4 Process Parameters

In engineering metallurgy, steel processing is a major factor influencing mechanical and physical properties. It generally includes heat treatment, thermomechanical deformation, and surface hardening processes, which modify the microstructure and mechanical behaviour without changing the chemical composition (Krauss, 1993).

Rolling: It is a thermomechanical deformation process used extensively in steel manufacturing to convert cast products such as billets and slabs into plates, sheets, and structural shapes (Black et al., 1996). It is performed either above the recrystallisation temperature (hot rolling) or below it (cold rolling). Hot rolling reduces deformation resistance, allowing large shape changes without cracking while refining the grain structure and improving ductility. Cold rolling is carried out at lower temperatures, which increases strength and hardness due to strain hardening and provides improved surface finish and dimensional accuracy.

Annealing**:** It is a classical heat-treatment process in which steel is heated to a temperature above the critical range and then cooled very slowly within the furnace (Dieter and Bacon, 1976). The controlled furnace cooling promotes recrystallisation and the formation of coarse pearlite, leading to reduced hardness, improved ductility, and relief of internal stresses.

Normalising: It involves heating steel to an elevated temperature above its critical point, then removing it from the heat source and allowing it to cool in room-temperature air (Lyman, 1948). Unlike annealing, where the steel is cooled slowly inside the furnace (resulting in even slower cooling), normalising employs faster air cooling. This difference in cooling rate means normalised steel has a finer grain structure and typically higher strength and hardness, but slightly less ductility, than annealed steel.

Quenching**:** It is a heat-treatment process of steel in which rapid cooling is applied after heating at a high temperature (Lyman, 1948). Different quenching types are classified based on cooling conditions, including water quenching, oil quenching, and vacuum quenching. In general, Quenching applies to adjust hardness and to add toughness.

Tempering**:** Following quenching, the steel exists in a hard but brittle martensitic condition containing high residual stresses. Tempering involves reheating the quenched steel to a temperature below the eutectoid temperature (A_1) and holding it for a controlled period. This allows partial decomposition of martensite, relieves internal stresses, and produces a tempered martensitic structure with improved toughness and ductility, although with some reduction in hardness (Harris, 1994). Higher tempering temperatures generally result in lower hardness and greater toughness.

Carburising: It is the heat treatment process to harden the surface of the materials in which the metal part or low carbon surface is heated in a carbon-rich gas atmosphere (Sharma, 2003). This technique enhances surface hardness while maintaining the original toughness.

As described, steel varies widely in terms of its chemical composition, processing conditions, and resulting properties. To meet different performance requirements, thousands of grades have been developed and traditionally grouped based on their composition which has been described. To bring consistency to this complexity, countries and regions have established their own steel grade systems, each with unique naming based on either chemical elements, mechanical properties, or both. The next section outlines the major standardisation organisations and explains how these systems have contributed to the emergence of over 3,500 recognised steel grades (García Gutiérrez et al., 2024).

2.2.5 Steel Grades System and Standards

The chemical elements, properties, fabrication methods, heat treatments, and steel shapes are all communicated through the steel grade (ASTM International, 2025). Grade designation systems provide a standardised vocabulary that allows engineers, manufacturers, and customers to describe and specify steels accurately. It is important to distinguish between steel classification and steel grade designation systems, as these terms serve different purposes in metallurgy. Classification refers to the grouping of steels based on their characteristics, such as chemical composition, microstructure, or manufacturing route and was introduced earlier in Section 2.2.3, where steels were grouped compositionally into carbon, alloy, stainless, and tool steels. In contrast, steel grade systems and standards provide a naming convention used by international organisations to identify specific steels within these broader classes (Lyman, 1948).

A number of international organisations publish such standards, each using its own conventions.

American Iron and Steel Institute (AISI)

In AISI, carbon steels are determined by four-digit numbering, and stainless steels determine by a three-digit numbering system with a "type" prefix for identification. In addition, some steel grades have suffixes that indicate compositional changes, such as 303Se, which indicates the addition of selenium to the composition. The AISI compositions and designations serve as primary standards in various industries (Klar and Samal, 2007).

Society of Automotive Engineers International (SAE)

Similar to AISI, the first digit of the four-digit number issued to alloy and carbon steels identifies the primary alloying contents (Jeffus, 2016). The second number represents the top-grade element, while the final two digits represent the steel's carbon content. SAE also assigns five-digit numbers for certain special-purpose and high-strength steels, but stainless steels use the same three-digit designations as the AISI system. There is a chance that AISI grades will be referred to as SAE grades, and AISI/SAE standards frequently use the same steel identification number.

Table 2.1: AISI/SAE designation system.

AISI/SAE Designation	Type
1xxx	Carbon Steel
2xxx	Nickel Steel
3xxx	Nickel-chromium steels
4xxx	Molybdenum steels
5xxx	Chromium steels

The American National Standards Institute (ANSI) and American Society of Mechanical Engineers (ASME) also use their own numbering system. In addition, one of the famous and established standards is American Society for Testing and Materials (ASTM). ASTM uses a specific designation system for metals that starts with a letter, for example "A" for ferrous materials, followed by a sequential number. These designations are often specific to certain products. For instance, A548 refers to cold-heading quality carbon steel wire used for making

tapping or sheet metal screws. When referring to metric ASTM standards, the designation includes the suffix "M" to indicate the use of the metric system (Bringas, 2004).

According to the above steel grade systems, and so many grades across different standards it is worth to ask whether all of them are necessary. In many cases, steels from different or same designation systems may share the same properties and performance but are named differently (Bringas, 2008). This makes it difficult to identify overlaps and adds complexity. To address this problem, first find a systematic way to group similar steels based on shared characteristics is crucial step. This leads to the broader concept of classification but not just in the context of metallurgy, instead as a general data-driven approach. The following section explores how classification is applied across different fields, and how these principles can support a more streamlined and meaningful organisation of steel grades.

2.3 An Insight into Data Classification

Many objects in life are categorised or divided into categories or groups depending on their features. Classification is related to the process of recognising, comprehending, and categorising objects into different groups, also known as sub-populations (Nisbet et al., 2018). It plays a key role in organising complex information by facilitating the identification and interpretation of diverse elements within a system.

The concept of classification was first formalised by Carl Linnaeus, who is recognised as the father of taxonomy (Tsafou and Jensen, 2016). He introduced the Linnaean system, a hierarchical structure that still used in biology to name and organise living organisms based on genus and species. Over time, the classification became important, especially with the emergence of data.

The problem of classification is in fact one of the major applications of machine learning not only in materials / steel research but also in a variety of other disciplines. Drug discovery is a most relevant example of such applications, with some similarities to the problem of steel classification. Drug discovery is the process of making drugs for the targeted disease (analogous to engineering application of steel), with minimal side effects on the body of patients (analogous to negative environmental impact). Lead compound identification (analogous to alloy composition) is one of the essential steps for finding the relevant drug for a particular disease in the drug discovery process (Lu et al., 2016). A further critical factor in steel reclassification is consideration of different process parameters, which have an impact on

the final properties of the product, such as strength and durability. In view of the relevance of the drug classification to the problem of rationalisation of steel grades, a review of the methodology in the former is given next.

Many studies utilised machine learning tools to classify drugs based on their similarities and differences. For instance, Malhat et al., (2014) applied various unsupervised methods, including K-Means, bisecting K-Means, and the Ward algorithm, to group drugs and identify lead compounds. The K-Means performed better in terms of clustering consistency as measured by standard deviation. In addition, Lu et al., (2016) applied the K-Means algorithm to classify drugs targeting two types of lung cancer and successfully identified a distinct cluster with potential anti-cancer activity.

In similar study, Madugula et al., (2021) combined PCA and K-Means clustering to classify approved drugs based on structural similarities for repurposing, using molecular and biological activity data. In addition, Korkmaz et al., (2014) used supervised learning with feature selection and SVM to distinguish between drug-like and non-drug compounds. Hasan et al., (2019) used the Hierarchical Clustering method to investigate the toxicity doses of drugs in the drug development process and successfully classified similar toxic doses of drugs into a single group.

From these studies, it is clear that machine learning, particularly clustering methods such as K-Means clustering, is used to group drugs based on their chemical similarities. In all cases, Euclidean distance was used to measure the similarity between data points. One thing that stands out is that the classification was mostly done using molecular descriptors or chemical features of the drugs, not based on their therapeutic effects or application areas. Furthermore, none of the studies explicitly focused on reducing the number of drugs by identifying and reducing redundant candidates within clusters. This presents a methodological gap that directly parallels the objective of steel grade rationalisation, where the goal is not only to classify similar grades but also to minimise the number of distinct grades by identifying the overlap and interchangeability of steel grades.

The use of classification methods is not limited to drug discovery; it has also been applied in various other fields, including agriculture and astronomy. The soil contains various attributes, including chemical elements, as well as electrical conductivity and pH values, which are relevant to this work. For instance, Hayatu Hassan et al., (2020) applied K-Means to classify

soil samples based on fertility for plant development and could successfully classify the soil into four distinct groups, with only one group showing a higher fertility rate than the others.

In similar applications, Data-driven methods have been applied to astronomical applications where unsupervised techniques are used to group stars based on their behavioural or spatial characteristics. For instance, Valenzuela and Pichara, (2018) classified variable stars by analysing similarities in their light curves using a query-based method, which allowed astronomers to select one star and identify others with similar behaviour based on Euclidean distance. In addition, Darma and Wulan, (2021) applied DBSCAN to classify stars into main-sequence and non-main-sequence groups, using spatial distances between data points to identify patterns in stellar behaviour. In general, K-means clustering analysis is frequently applied in astronomical research for categorising celestial objects within transient search projects (Wozniak et al., 2001; Zhang and Zhao, 2004).

After reviewing the above studies, it is evident that when classification involves multiple parameters, machine learning plays a significant role in uncovering underlying structures within the data and grouping items according to their similarities across these parameters.

2.4 Summary

This chapter reviewed the existing approaches to steel classification and recycling. On the sustainability side, prior work has shown that recycling significantly reduces energy consumption and CO_2 emissions. However, the persistence of tramp elements such as copper, and tin, together with dependence on critical raw materials like chromium and molybdenum, limits the quality and circularity of secondary steel. While strategies such as closed-loop recycling and life-cycle assessment frameworks have been explored, little attention has been given to whether reducing the number of grades, or prioritising grades with higher tramp element tolerance and lower CRM content, could enhance recyclability and long-term sustainability.

From an industrial perspective, the review of steel production, consumption, and grade systems highlighted the dominance of carbon and stainless steels, which account for the majority of global demand. Their widespread use across construction, automotive, and engineering sectors has contributed to the existence of more than 3,500 grades. While classification systems such as AISI, SAE, and ASTM ensure consistency and quality control, they also introduce duplication, with many grades offering similar mechanical performance under different names

or standards. Although prior work documents this complexity, there is limited discussion on systematic approaches to reduce redundancy.

In addition, insights from broader domains such as drug discovery, agriculture, and astronomy showed how classification methods can successfully group entities with shared features. These diverse applications highlight how machine learning, particularly clustering techniques used for classification tasks, has become an essential tool across scientific domains. This growing use of data-driven grouping methods sets the stage for the next chapter, which explores the emergence of machine learning in real-world applications and its role in materials science.

Chapter 3: Literature Review - Machine Learning (ML)

As discussed in chapter 1, this study utilises machine learning as a tool for rationalising steel grades. This chapter begins with the historical emergence and evolution of machine learning in scientific and industrial domains, followed by its adoption in materials science, where data-driven methods increasingly complement experimental approaches. The chapter then focuses on supervised and unsupervised methods relevant to this study, particularly for property prediction, clustering, and highlights the use of explainable AI to interpret model outputs. Finally, it reviews reduction and elimination algorithms from different domains, which form the methodological foundation for this research.

3.1 Emergence of ML in Real World

Machine learning is a field of computer science and artificial intelligence that allows computers to learn from data without being explicitly programmed. By analysing historical or real-time data, it can generate predictions or automated responses, making it particularly useful for tasks where traditional rule-based algorithms may be inefficient or impractical (Paluszek and Thomas, 2017).

The origins of machine learning go back to 1950, when Alan Turing introduced the Turing Test to assess whether a computer could show intelligent behaviour (Turing, 2009). Two years later, Arthur Samuel developed the first self-learning computer program for playing checkers, and later, in 1959, he formally defined machine learning as the field of study that gives computers the ability to learn without explicit programming (Samuel, 1959). This marked an early turning point in the development of artificial intelligence.

Further progress came in the early 1960s, when Frank Rosenblatt and his team at Cornell University built a system inspired by the human nervous system to recognise alphabetic characters (Kussul et al., 2001). Although this work initially received limited attention, it laid the foundation for pattern recognition. A major breakthrough followed in 1986, when David Rumelhart and James McClelland introduced the backpropagation algorithm, enabling multi-layer neural networks to adjust weights through error correction and improving predictive capabilities (Rumelhart et al., 1986). From this point, machine learning gradually emerged as a distinct field and separated from traditional artificial intelligence and pattern recognition.

Currently, machine learning is applied across a wide range of sectors, including healthcare, transport, finance, and manufacturing, where it supports applications such as image recognition (X. Wang et al., 2024), fraud detection (Ali et al., 2022), recommendation systems (Fanca et al., 2020), and predictive maintenance (Carvalho et al., 2019).

3.2 Emergence of ML in Material Science

Machine learning was first applied in materials science by Fischer et al., (2006), who demonstrated that it could predict crystal structures more efficiently than quantum mechanical methods, which were computationally expensive and time-consuming. Since then, it has been applied to a range of materials, including solar materials (Yu and Zunger, 2012), water photo splitting materials (Castelli et al., 2012), carbon capture and gas storage materials (Lin et al., 2012). These studies highlighted the potential of machine learning to accelerate materials discovery, but they were limited to functional materials where properties are mainly determined by chemical composition.

In contrast, structural materials such as steels present a more complex challenge, as their mechanical properties depend not only on composition but also on processing history. Traditional regression methods, which assume linear relationships between parameters, struggle to capture these nonlinear dependencies (Nielsen and Martins, 2021). The recent development of machine learning addresses this limitation by assigning weights to input variables and adjusting them during training to minimise prediction error (Almeida, 2002). In doing so, it can capture the nonlinear relationships between composition, processing conditions, and mechanical properties in structural materials. This capability has accelerated materials discovery, reduced reliance on trial-and-error experiments, and improved the accuracy of property prediction.

To address the challenges of modelling nonlinear behaviour, it is essential to select ML approaches that align with the problem and the type of feedback available during training. When labelled examples are provided, supervised learning methods can be used to learn the relationship between inputs and outputs (Baştanlar and Özuysal, 2014). In contrast, when only unlabelled data are available, unsupervised learning techniques are applied to uncover underlying structure or similarity patterns within the dataset. Other approaches, such as semi-supervised learning, which combines a small amount of labelled data with larger unlabelled datasets (Pise and Kulkarni, 2008), and reinforcement learning, where an agent learns through evaluative feedback (Sutton and Barto, 1998), also exist, but they are not relevant to the aims

of this study. The following section introduces these approaches in more detail and explains their relevance to this research.

3.3 Supervised Machine Learning

Supervised machine learning algorithms learn by example. In other words, in the supervised learning approach, a teacher guides the learning process. In this method, a given dataset acts as a teacher; its role is to train the model. Once the model gets trained, it can start making a prediction or decision when new data is given (Pink, 2016). According to Figure 3.1, supervised machine learning is generally divided into two types: regression and classification. The regression method is mainly used to find the relationship between the dependent and independent variables. Also, it can be used to predict the numerical value based on previously observed data. In contrast, the classification method is mainly used to predict the category that the data belongs to (Torgo and Gama, 1996).

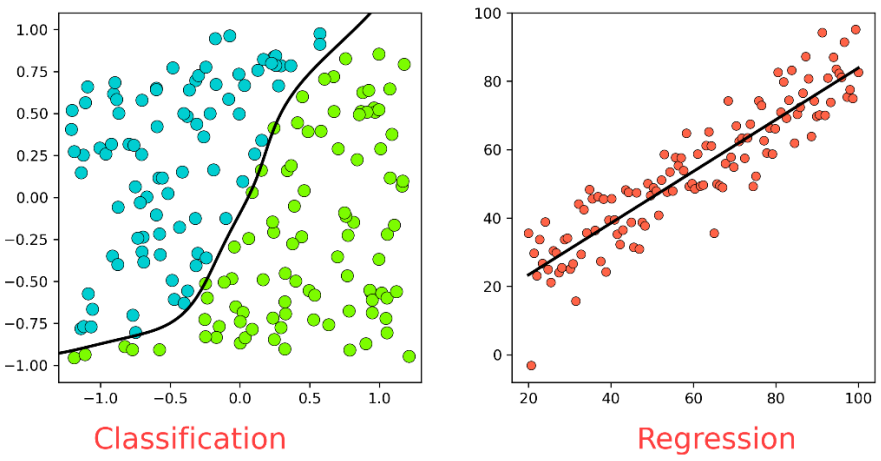


Figure 3.1: Example of supervised machine learning approaches (classification and regression), generated by the author using simulated data.

Supervised machine learning approach is widely used in different areas of material science. For instance, Ward et al., (2016) used supervised machine learning approaches for predicting the properties of inorganic materials; where they developed a decision tree ensemble to predict the three physically distinct properties including band gap energy, specific volume, and formation energy with high accuracy (91% using reduced-error pruning). The model identified the most influential attributes, demonstrating supervised learning ability to uncover structure–property relationships in materials.

As discussed in previous section, in steel-related applications, the mechanical and physical properties of steel mainly depend on different parameters such as chemical compositions and

process parameters (Fig. 3.2). Understanding how these variables interact is essential for various tasks, including alloy design, quality control, and the characteristics of existing steel grades classification. Supervised learning methods, particularly regression-based models, are well-suited to capture these interactions, enabling the identification of key features that influence performance.

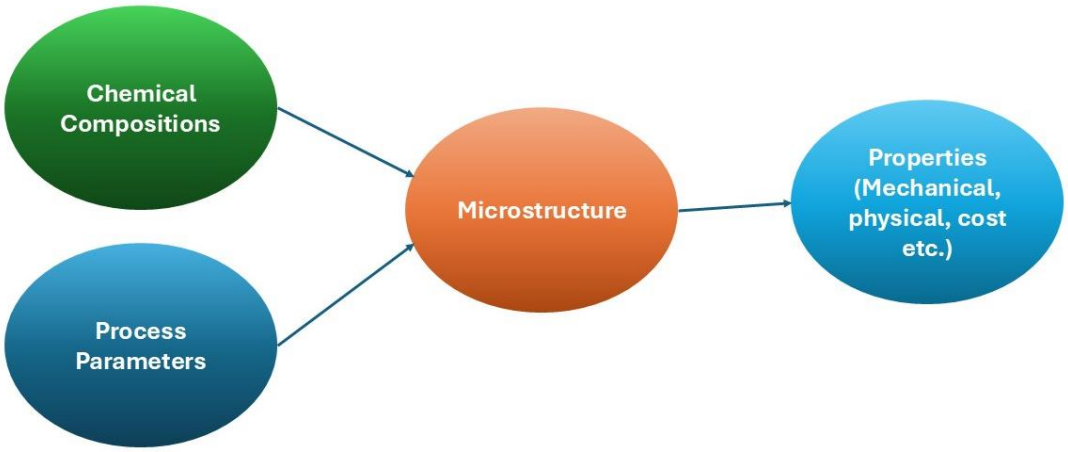


Figure 3.2: Correlation between different steel parameters.

Among the many algorithms available for regression problems, two families have been most widely applied in steel-related research including Artificial Neural Networks and tree-based ensemble models such as Random Forest and Extreme Gradient Boosting. The following introduce these algorithms and explain their relevance before turning to specific applications.

Artificial Neural Networks (ANNs)

ANNs are a class of supervised learning algorithms inspired by the structure of the human nervous system (Schmidhuber, 2015). Their origins date back to the 1970s, and they have since become widely applied to both regression and classification problems. ANN composed of a large number of highly interconnected processing elements known as the neuron to solve problems (Montesinos López et al., 2022). As shown in Figure 3.3, the architecture of ANN consists of three layers: Input layer, hidden layer and output layer. For the learning process, the data feds into the network via the input layer, and then the model learns by different combinations of input units, and consequently, it gives output. The neural network works by a feedback process called backpropagation (Bebis and Georgiopoulos, 1994). It means the expected outcome compares with the generated output and dedicates the weight by computing

the differences between them. This process performs iteratively until the margin of error is minimal.

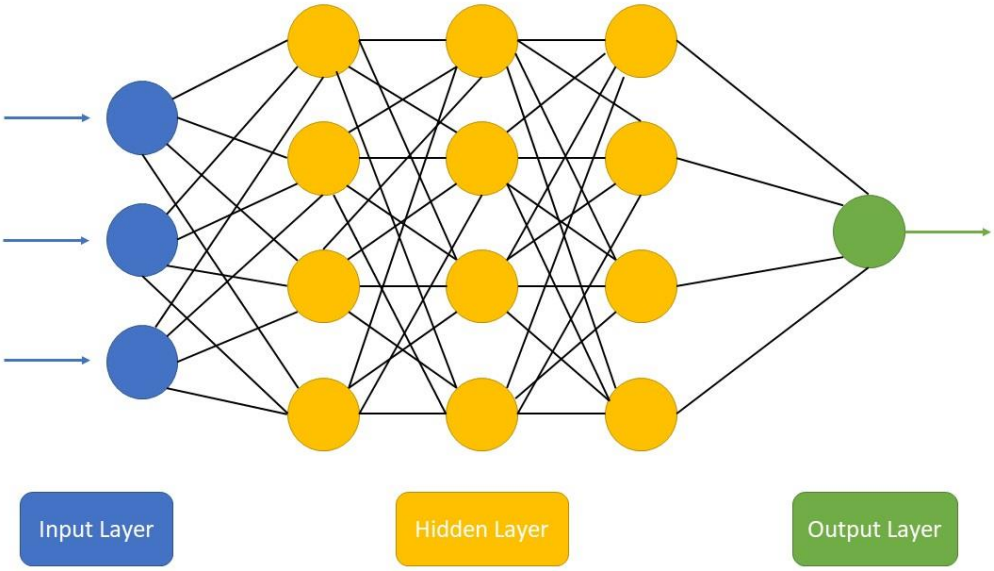


Figure 3.3: Example Architecture of a Feedforward ANN. The network is composed of three fundamental layers: an input layer (blue nodes) where data is fed into the model; one or more hidden layers (yellow nodes) where non-linear computations are performed; and an output layer (green node) which provides the final prediction. The flow of information is unidirectional, and the lines represent weighted connections between neurons.

Figure 3.4 illustrates the structure of a single artificial neuron. Each neuron receives a set of inputs $x_1, x_2, \ldots, x_n$, each associated with a learnable weight w_i. These weighted inputs are aggregated and combined with a bias. The neuron then applies an activation function $f(\cdot)$ to this aggregated value to generate the output (Walczak and Cerpa, 2003).

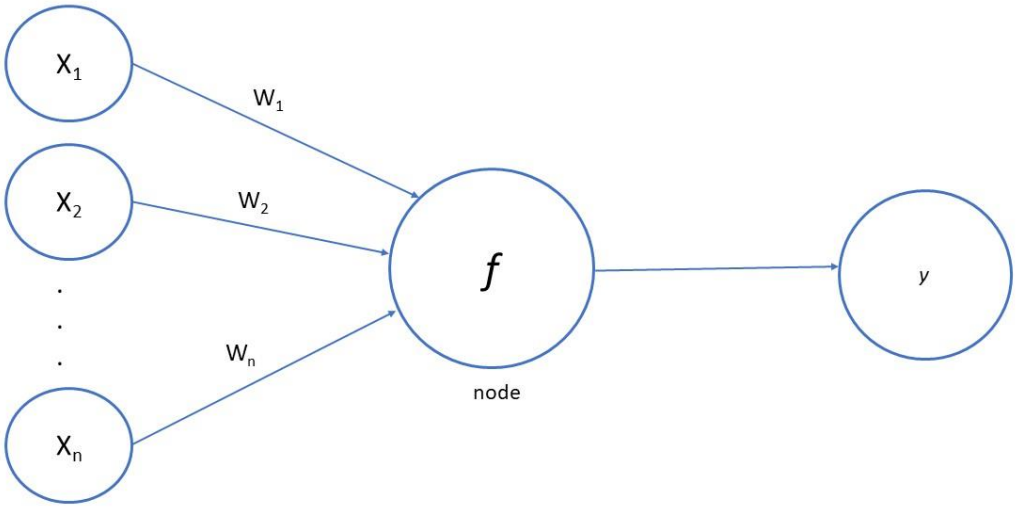


Figure 3.4: Schematic representation of a single artificial neuron demonstrating weighted inputs, activation function, and output.

This behaviour represented mathematically (Zou et al., 2009) as shown in equation (3.1):

$$\hat{y} = f\left(\sum_{i=0}^{n} w_i x_i + b\right) \tag{3.1}$$

In equation 3.1, $\hat{y}$ is the output of the node, $f(\cdot)$ is the activation function, w_i is the weight assigned to each input x_i, and b is the bias term.

Activation functions are essential because they allow the neuron to model complex, non-linear relationships that cannot be captured by linear transformations alone (McCulloch and Pitts, 1990). They are mathematical operations applied to the output of each neuron that determine whether and to what extent the neuron is activated. Several activation functions are commonly used in ANN-based models to transform the summed weighted inputs and pass the resulting signal to the next layer, enabling the network to construct hierarchical representations suitable for tasks such as regression, classification, and pattern recognition (Szandała, 2021). The most widely used activation functions in regression tasks are described below.

Sigmoid function: It maps input values to the (0, 1) range and is mathematically defined as:

$$f(x) = \frac{1}{1 + e^{-x}} \quad (3.2)$$

The sigmoid function has historically been used in binary classification models because its bounded output can be interpreted as a probability (Cybenko, 1989). However, empirical studies have shown that its gradient becomes extremely small for large positive or negative inputs, which leads to slow parameter updates during training and makes the function less suitable for deeper neural network architectures (Dubey et al., 2022).

Hyperbolic tangent (tanh) function: It maps input values to the range -1 to 1 and is mathematically defined as:

$$f(x) = \tanh(x) \quad (3.3)$$

The tanh function is zero-centred and often provides better optimisation behaviour compared with the sigmoid function. Nevertheless, like the sigmoid, it suffers from gradient saturation when inputs fall into extreme positive or negative ranges, which can hinder training efficiency in deeper networks (Lecun et al., 2000).

Rectified Linear Unit (ReLU): It was became widely adopted after Glorot et al., (2010) demonstrated its advantages in training deep networks. It is defined as:

$$f(x) = \max(0, x) \quad (3.4)$$

ReLU avoids saturation for positive inputs, enabling faster and more stable optimisation. Although neurons may become inactive when inputs remain negative, the function remains highly effective for regression tasks, including the prediction of steel mechanical properties.

While ANNs perform well at capturing such complex relationship, they typically require large datasets and are often criticised for limited interpretability (Dou et al., 2023). These limitations have led researchers to use alternative ensemble methods, such as RF and XGBoost.

Random Forest (RF)

RF is an ensemble learning method introduced by Breiman, (2001). It is based on constructing a large number of decision trees, each trained on a random subset of the data and features, and then aggregating their outputs (Rigatti, 2017). For regression tasks, it constructs an ensemble of unpruned decision trees using bootstrap samples of the training data, with random feature selection at each split to ensure tree diversity. Each tree produces a regression estimate based

on its training subset, and the predictions from all trees are aggregated, typically by averaging, to yield the model's final continuous output.

Mathematically the overall prediction is:

$$\hat{y}(x) = \frac{1}{k}\sum_{k=1}^{k} T_k(x) \qquad (3.5)$$

Where, x is the input sample, K is the total number of trees in the ensemble, $T_k(x)$ is the output of the k-th decision tree, and $\hat{y}(x)$ is the aggregated RF prediction. Averaging over multiple trees reduces variance and increases robustness by mitigating the overfitting typically observed in individual decision trees (Biau, 2010).

For instance, if three trees predict YS values of 500 MPa, 520 MPa, and 510 MPa for the same steel composition, the RF prediction would be their average, 510 MPa. This illustrates how aggregation stabilises the model and improves generalisation performance on unseen data.

Averaging improves robustness but does not explicitly correct systematic errors; boosting addresses this by refining mistakes sequentially.

Extreme Gradient Boosting (XGBoost)

XGBoost is an ensemble method that extends the principle of gradient boosting, proposed by Chen and Guestrin, (2016). Unlike RF, where trees are built independently and their predictions are averaged, XGBoost builds trees sequentially, with each new tree correcting the residual errors of the previous ensemble. This sequential construction is important because it enables the model to progressively reduce both bias and variance. Rather than relying on the averaging of many independent weak learners, XGBoost improves accuracy by systematically focusing on the mistakes made by earlier trees.

This iterative correction can be expressed mathematically, where the overall prediction is updated round by round as new trees are added:

$$\hat{y}_i^{(t)} = \sum_{k=1}^{t} f_k(x_i), \qquad f_k \in \mathcal{F} \qquad (3.6)$$

In equation 3.6, x_i is the input sample, $\hat{y}_i^{(t)}$ is the predicted value after t trees, f_k represents the k-th regression tree, and $\mathcal{F}$ denotes the functional space of all possible trees. During training, XGBoost optimises a regularised objective function:

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i{}^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) \tag{3.7}$$

Where y_i is the true target value, l is the loss function (e.g., mean squared error), $\hat{y}_i^{(t-1)}$ is the model prediction from the previous round, f_t is the new tree added at iteration t, and $\Omega(f_t)$ penalises model complexity to reduce overfitting. This combination of loss minimisation and regularisation allows XGBoost to achieve both high predictive accuracy and strong generalisation (Wiens et al., 2025), making it particularly valuable for modelling nonlinear relationships in steel datasets where both composition and processing history must be considered.

In summary, ANNs, RF, and XGBoost represent three supervised learning methods that are particularly relevant in steel-related studies. The discussion here has focused on their underlying mechanisms, highlighting how they approach regression problems and why they are suitable for modelling composition–processing–property relationships. With these foundations, the next sections turn to the wider range of supervised learning applications in materials science, where these and other algorithms have been employed for property prediction, microstructure classification, quality assessment, and defect diagnosis.

3.3.1 Predicting Properties using Supervised Learning

Several studies have demonstrated the utility of supervised machine learning in predicting core mechanical properties of steel such as YS, UTS, elongation, and hardness. In materials science, the size and scope of datasets strongly influence the choice and performance of algorithms. Many studies rely on relatively small datasets due to the cost and effort of data collection, whereas access to industrial databases has enabled larger-scale investigations. Accordingly, different supervised learning models have been tested under these conditions to evaluate their suitability for property prediction.

For smaller datasets, several studies have highlighted the benefits of ensemble methods. Xiong et al., (2020) applied four ML models (RF, Linear Least Squares (LLS), k-Nearest Neighbours (KNN), ANN) to 360 low alloy and carbon steel data. LLS is a simple regression method that

fits a linear model by minimising the sum of squared errors, while KNN is a non-parametric model that predicts a target value based on the average of the k most similar data points in the feature space. Among these, RF achieved the best performance based on correlation coefficient (R) and relative root mean square error (RRMSE). Feature selection revealed that tempering temperature and alloying elements such as carbon, chromium, and molybdenum were the most influential factors in predicting tensile strength and hardness of carbon steels. In similar studies, Santos et al., (2009) used ANN and KNN to predict UTS in cast iron based on foundry data. ANN achieved an accuracy of approximately 84% and outperformed KNN in terms of RMSE and mean square error (MAE). Wang et al., (2025) used 106 samples of automotive steels to compare several supervised learning models, including ridge regression, Support Vector Regression (SVR), Gradient Boosted Decision Trees (GBDT), RF, and Adaboost. Ridge regression is a linear model that incorporates L2 regularisation, which penalises large model coefficients and therefore discourages overly complex fits (Hoerl and Kennard, 1970). By shrinking coefficient magnitudes towards zero, L2 regularisation reduces the model's sensitivity to noise in small datasets and helps prevent overfitting. In contrast, SVR is a kernel-based method that learns a regression function within a specified error tolerance, enabling it to model nonlinear relationships while controlling model complexity through the margin parameters (Vapnik, 2000). GBDT and Adaboost are boosting algorithms that construct ensembles of weak learners, where each successive model focuses on correcting the errors made by earlier ones (Freund and Schapire, 1997; Friedman, 2000). After benchmarking the algorithms individually, Wang et al. combined them into stacked ensembles, which consistently outperformed single models. The RF–Adaboost combination performed best for YS, GBDT–Adaboost for UTS, and a three-model ensemble (Adaboost–GBDT–SVR) achieved the highest accuracy for fracture elongation. These findings demonstrate that ensemble methods can exploit complementary strengths across algorithms and deliver superior performance, particularly when working with limited or noisy data. In addition, Zhao and Yang, (2025) analysed 245 high-strength steel samples and showed that XGBoost outperformed other models, achieving the lowest RMSE and the highest R^2, while also revealing that niobium additions enhanced strength more effectively than vanadium or titanium. Dabiri et al., (2022b, 2022a) investigated tensile strength and ultimate strain in reinforcement bars using around 400 experimental samples under both spliced and non-spliced conditions. Across models including nonlinear regression, ridge regression, ANN, decision trees, and RF, they found that nonlinear regression performed better for the purpose of tensile

strength prediction, while tree-based models achieved R^2 values above 85% for ultimate strain, effectively capturing complex stress–strain behaviour without destructive testing.

These studies indicate that tree-based models such as RF and XGBoost are particularly effective for small, noisy datasets because of their robustness to noise and reduced risk of overfitting (Kiangala and Wang, 2021). However, tree-based models can struggle to capture highly complex or nonlinear interactions between alloying elements and processing conditions. This is because decision trees rely on feature-aligned splits, which makes it difficult for them to represent smooth or strongly coupled relationships across multiple variables without requiring a very large number of partitions. In such cases, ANNs offer an advantage, as they are designed to model nonlinear relationships more flexibly. Although ANNs often require larger datasets and careful regularisation, they have also been successfully applied in small-sample studies.

Chou et al., (2016)combined Particle Swarm Optimisation with ANN to optimise the chemical composition of steel bars, successfully predicting tensile strength and yield point, with C, Si, and Mn identified as key contributors. In addition, Ozerdem and Kolukisa, (2009) applied ANN to predict properties of Cu–Sn–Pb–Zn–Ni cast alloys using composition alone. Singh et al., (1998) estimated tensile and yield strength of ferrite–pearlite steels based on composition and rolling parameters, using a Bayesian framework to reduce overfitting and identify variable importance. Other studies, such as Jia et al., (2011), Sterjovski et al., (2005), Hwang et al., (2010), and Lalam et al., (2019), also developed ANN-based models for properties like toughness, hardness, and strip performance, though these were often restricted to specific alloy types or production settings. In a more recent study, Park et al., (2024) applied ANN to predict the yield strength of austenitic stainless steel using 160 samples. Their input features included the chemical composition and welding heat input. Despite the relatively small dataset, the ANN model achieved better performance in terms of R^2 and mean percentage absolute error (MAPE), a metric that expresses prediction error as a percentage of the true value, making it useful for comparing performance across different scales.

In contrast to small experimental datasets, studies based on large industrial databases have demonstrated how greater data availability enables more comprehensive modelling and even alloy design. Guo et al., (2019) used around 63,000 samples to predict YS, UTS, and elongation, where regression trees and RF achieved superior MAE and RMSE. Importantly, the large dataset allowed revealed property boundaries and identified optimum alloying

compositions. In addition, Lee et al., (2021) utilised 55,473 records of TMCP steels provided by Hyundai Steel Co. and applied sixteen machine learning algorithms. Non-linear ensemble methods, particularly RF, achieved the strongest predictive accuracy. To extend these predictions into alloy design, the authors coupled the RF model with the Non-Dominated Sorting Genetic Algorithm II (NSGA-II), a multi-objective optimisation method that searches for Pareto-optimal solutions by balancing competing objectives rather than selecting a single best outcome. In this framework, RF acted as a fast surrogate model that predicted mechanical properties for many hypothetical compositions, while NSGA-II used these predictions to identify optimal combinations of C, Mn, Nb, and Si that satisfied multiple performance targets simultaneously. This integration demonstrated how supervised learning can be combined with evolutionary optimisation to support data-driven alloy design in industrial settings. Large dataset applications have demonstrated the full potential of ANN-based methods as well. Xie et al., (2021) employed a Deep Neural Network (DNN) to predict YS, UTS, elongation, and impact energy (Akv) using 11,101 datapoints from hot-rolled steels. Unlike earlier studies, this dataset included both chemical composition and detailed process parameters, allowing the model to capture complex, nonlinear relationships between composition, processing history, and mechanical behaviour. The DNN achieved higher predictive accuracy than traditional ML models, highlighting the promise of deep learning when applied to large-scale and multi-dimensional datasets.

A review of the above studies shows that three algorithms including RF, XGBoost, and ANNs, consistently appear among the strongest performers for predicting the mechanical properties of steels. Tree-based ensemble models (RF and XGBoost) are particularly effective for small or noisy datasets because bootstrap sampling and ensemble averaging reduce variance and limit overfitting. ANNs, although typically associated with larger datasets, have also shown competitive performance in small-sample studies when combined with careful feature selection or regularisation, and they remain especially advantageous when modelling highly nonlinear interactions between composition, processing history, and mechanical behaviour, relationships that tree-based models may struggle to represent efficiently. Overall, the relative performance of these algorithms depends on dataset size, the level of nonlinearity, feature complexity, and the robustness of preprocessing strategies. Despite their strong predictive capabilities, a key limitation across existing research is the lack of attention to explicit treatment conditions such as whether the steel was normalised, annealed, or quenched and how these influence the mechanical properties. In addition, no studies have systematically considered the standardised

steel grades as broader case study to explore the relationship between composition, processing conditions, and properties like hardness, YS, UTS, and elongation across different steel categories.

While supervised learning has mainly been used to predict the mechanical properties of steel, it has also been increasingly applied to microstructure classification. This is important because the microstructure has a significant influence on its overall performance. The following section examines how these methods, particularly image-based approaches, are helping to automate and improve the classification of steel's microstructure.

3.3.2 Microstructure Classification using Supervised Learning

Supervised learning techniques have also been applied to classify steel microstructures. The microstructure is steel's inner structure and represents steel's mechanical properties (Bodyakova and Belyakov, 2023). Hence, Steels have different microstructure appearances, which are influenced by different parameters such as alloying elements, heat treatment, cooling rate, etc.

Nowadays, many studies have focused on microstructural characteristics because it is the main element for understanding the material process and property. In general, the performance of the steel under different applications mostly depends on the distribution, shape, and size of phases in the microstructure (Durmaz et al., 2023). Thus, the correct classification of steel microstructures is a complex task. Traditionally, microstructural characterisation was performed manually by expert metallurgists who compared microscopy images with standard reference micrographs using visual assessment and professional judgement (Vander Voort, 2004). This approach is inherently subjective and lacks reproducibility, as the outcome depends heavily on the experience and interpretation of the examiner. However, these days this task has become more understandable, accurate, and accessible with the help of AI, especially the supervised learning methods. For instance, Gola et al., (2018) applied support vector machines (SVM), a margin-based classifier that separates classes using an optimised decision boundary, with feature selection to classify steel phases, achieving 87.15% accuracy, while Azimi et al., (2018) developed a Fully Convolutional Neural Network (FCNN), a deep-learning architecture designed for pixel-wise image segmentation, which outperformed traditional approaches and achieved 93.94% accuracy in low-carbon steels.

Although these studies highlight the potential of supervised learning for microstructural recognition, the present work did not rely on image-based datasets, which were limited in availability for many steel grades. Instead, microstructural information is represented explicitly in tabular form, using categorical descriptors such as ferritic, martensitic, and austenitic. In this way, microstructure is treated as an input feature rather than a prediction target.

Related work has also examined supervised learning in the context of specific microstructural phases like martensitic steels, which are widely used in industry due to their high strength and wear resistance (Nogara and Zarrouk, 2018). A key focus in this area is predicting the martensite start temperature (Ms), an important factor in designing steel composition and heat treatments. Early studies Bhadeshia, (1981) and Capdevila et al., (2002) used neural networks and linear regression to estimate Ms, showing good results within narrow compositional ranges. More recently, Zhang and Xu, (2021) applied Gaussian Process Regression (GPR), a probabilistic, kernel-based method that models uncertainty explicitly, on 1,119 carbon steel samples to predict Ms across a wider alloying range (153–938 K), achieving high accuracy based on RMSE and MAE metrics. While effective, these models are dependent to specific phase transitions and are not directly aligned with property prediction. However, it shows that supervised learning can find the relationship between alloying elements and martensitic steels.

While various studies have applied supervised learning to classify microstructures in steel, its use extends beyond structural recognition. Supervised learning has also been applied in quality assessment tasks, particularly in monitoring and predicting outcomes in continuous casting and related manufacturing processes.

3.3.3 Quality Assessment using Supervised Learning

Continuous casting is a critical stage in steel production where physical, mechanical, and chemical factors interact. Traditional quality control methods, such as ultrasonic testing (Nsengiyumva et al., 2021), are difficult to apply in real-time production settings and may miss defects caused by variations in process conditions.

To address quality challenges in continuous casting, supervised learning has been used to link process parameters with product outcomes. Jiang et al., (2011) applied backpropagation neural networks to correlate casting speed, superheat, and water ratio with cooling rate, while José et al., (2015) expanded the model by including chemical composition and mold oscillation, identifying sulfur and phosphorus as key predictors of segregation. More recently, Li et al.,

(2018) showed that ensemble methods combining SVM and KNN improved prediction of temperature and pressure, demonstrating the value of supervised learning for real-time quality monitoring.

Overall, these studies demonstrate that supervised learning not only enhances prediction accuracy but also facilitates a more systematic understanding of complex patterns, particularly in how various input features influence the targeted features in steel production.

Building on the role of supervised learning in assessing steel quality during production, another critical application lies in the diagnosis and classification of defects. The area where supervised machine learning has shown great promise in improving early detection and reducing manual inspection dependency.

3.3.4 Diagnose Defects using Supervised Learning

Supervised learning has also been applied to defect diagnosis in steel production, where early detection is critical for maintaining quality. Patel and Jokhakar, (2016) developed a defect cause analysis system using RF to diagnose cooling temperature deviations in mild steel coil production. Feature selection was guided by domain experts. Among various algorithms tested including neural networks, SVM, decision trees, and the RF achieved the highest accuracy in fault diagnosis.

Similarly, Chen and Kaufmann, (2022) applied supervised learning models to predict surface defects in casting processes, using data from a steel and cast iron foundry. Six regression-based models were evaluated, with the Extremely Randomised Trees algorithm performing best in terms of MAE, MSE, and RMSE. The SHAP framework as explainable AI technique was used to interpret feature importance, identifying factors such as excessive organic binders, high fines content, and low mold gas permeability as key contributors to defects.

Further advancing defect detection, Madhavan et al., (2024) proposed a two-stage system combining deep learning and classical machine learning. Convolutional Neural Networks (CNNs) were first used to extract features from surface images of hot-rolled steel, leveraging six pre-trained models. The optimal combination (Functional Tree classifier using DenseNet201 features) achieved an impressive 99.72% accuracy in defect classification.

These examples illustrate how supervised learning, particularly ANN, RF, and XGBoost, has been successfully applied across different aspects of steel research, from property prediction to defect detection, demonstrating their ability to capture relationships between composition,

processing, and performance. However, supervised learning is only suitable when labelled data are available, that is when the target property for each sample is already known. In many materials-science problems, the objective is not to predict a known property but to discover structure within the data itself, such as grouping steels with similar behaviour or revealing hidden relationships among composition, processing, and properties. For such tasks, supervised learning is not appropriate. Instead, unsupervised learning provides the tools needed to uncover patterns, cluster similar materials, and reduce dimensional complexity. These capabilities are especially relevant to steel grade rationalisation and data-driven reclassification. The following section reviews how unsupervised techniques have been applied in the context of steel research.

3.4 Unsupervised Machine Learning

Unsupervised machine learning has emerged as a powerful approach for extracting insights from unlabelled data, with applications in various domains including networking, computer vision, and natural language processing (Happonen et al., 2022). Unlike supervised learning, which relies on labelled examples to guide the model, unsupervised learning analyses data without any target values and seeks to reveal underlying structure or relationships that are not explicitly provided (Valkenborg et al., 2023). In general, there are different types of unsupervised machine learning which contain:

Anomaly detection: In this type of unsupervised learning, the model will find and discover unusual behaviours and patterns in the systems. It is mainly used in fraud detection in bank transactions (Hilal et al., 2022).

Association mining: In this type, the model will find the group of items that frequently occur together, and it is mainly used in basket analysis and customer purchase behaviour applications (Pedrycz et al., 2007).

Clustering: In this type, the model aims to find meaningful patterns and structure in a data set. Clustering techniques are commonly applied to identify patterns of similarity and dissimilarity among objects, by quantifying how closely related or distinct the observations are (Madhulatha, 2012). This approach is particularly relevant to the current study, which aims to reduce the number of steel grades. To achieve this, it is essential first to identify grades with similar characteristics, which then informs the broader objectives of this research. However, a critical element of any clustering method is the choice of distance metric, which defines how similarity or dissimilarity between data points is quantified. Because clustering algorithms group

observations based on how "close" they are to one another in feature space, the distance metric effectively shapes the geometry of the data and determines which samples are considered similar. Different metrics emphasise different aspects of the data: some prioritise absolute differences in magnitude, others focus on directional similarity, and some are more robust to noise or scaling differences. Common metrics include:

Euclidean distance: Measures the straight-line distance in multi-dimensional space that is well-suited for continuous numerical data (Vandeginste et al., 1998).

$$d_E(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \tag{3.8}$$

Where, x and y denote two data points represented by n features.

Manhattan distance: Calculates the sum of absolute differences across dimensions and be more robust for datasets where noise or measurement uncertainty is significant (Nixon and Aguado, 2020).

$$d_M(x, y) = \sum_{i=1}^{n} |x_i - y_i| \tag{3.9}$$

Cosine similarity: Evaluates the angle between vectors, typically used for high-dimensional or sparse data. It is commonly applied in text analysis or vector-based data where the magnitude of the vectors is less important than their direction (Han et al., 2012).

$$\cos(\theta) = \frac{x.y}{\|x\|\|y\|} \tag{3.10}$$

In materials science, Euclidean distance has been the dominant choice because mechanical and chemical descriptors are continuous and naturally represented in geometric space (Kamel et al., 2025). Given the central role of distance metrics in shaping clustering behaviour, the following section introduces the main clustering algorithms used to group data based on these similarity measures.

3.4.1 K-Means Algorithm

K-means algorithm is one of the common and popular methods for clustering problems due to its simplicity and scalability. In K-Means (Peña et al., 1999), the model randomly dedicates a data point as the centre point and finds the distance between each data point to the centre point using Euclidean distance (Fig. 3.5). Then performs, repetitive calculations to optimise the position of centroids. This process will be stopped until no movement occurs in the centroid's positions.

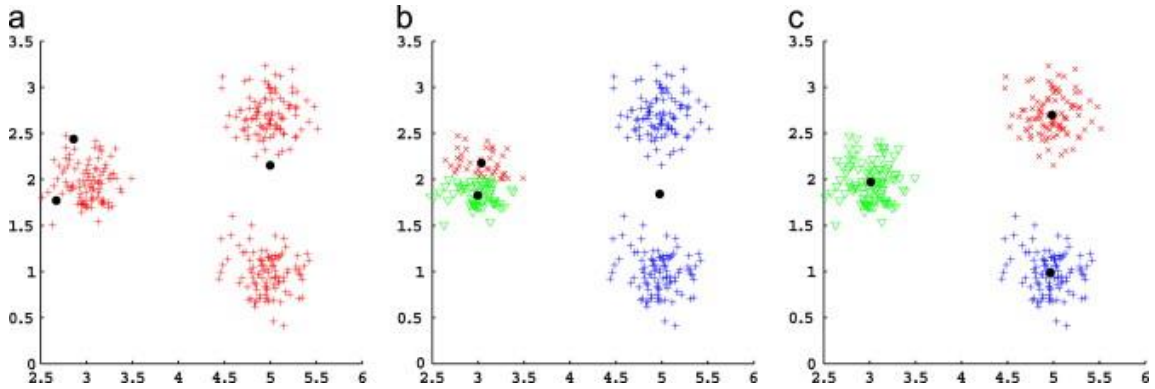


Figure 3.5: The mechanism of K-Means algorithm (Tzortzis and Likas, 2014). (a) Initial random placement of centroids. (b) Assignment of data points to the nearest centroid based on Euclidean distance. (c) Updated centroid positions after recalculating the mean of assigned points, followed by re-assignment; the algorithm repeats these steps until convergence.

Suppose the dataset is given by $X = \{x_1, \dots, x_N\}$, where each point $x_n \in \mathbb{R}^d$ represents a d-dimensional feature vector. The aim of K-Means is to partition the data into K disjoint clusters $C_1, \dots, C_K$, where each cluster C_k is represented by a centroid μ_k. For a point x_j, cluster membership is determined by assigning it to the nearest centroid.

The algorithm seeks to minimise the within-cluster sum of squared Euclidean distances:

$$v = \sum_{k=1}^{k} \sum_{x_j \in C_k} \left(x_j - \mu_k\right)^2 \qquad (3.11)$$

This objective function measures how tightly the data points are grouped around their respective centroids. A smaller value of v indicates that points within each cluster are close to their centroid, reflecting greater compactness and improved separation between clusters. During optimisation, K-Means iteratively alternates between (i) assigning each data point to the nearest centroid and (ii) updating each centroid to the mean of its assigned points, progressively reducing v until convergence.

A key factor influencing the performance of K-Means is the choice of the number of clusters (K), which directly affects how the data are grouped (Yuan and Yang, 2019). There are three main approaches for finding optimal number of clusters including Elbow method, Gap statistic and Silhouette coefficient. The Elbow method evaluates the reduction in within-cluster variance (SSE) across different K values and identifies the point where further gains become marginal (Kansal et al., 2018). The Gap statistic introduced by Tibshirani et al., (2001) compares clustering results to those of a random reference dataset, with the optimal K corresponding to the maximum gap between observed and expected performance. The Silhouette coefficient balances cohesion within clusters and separation between clusters, with higher average scores indicating more appropriate and stable partitions (Rousseeuw and Kaufman, 2009).

Because of this balance between efficiency, scalability, and the availability of diagnostic tools, K-Means has been widely adopted in commercial domains such as user segmentation (Bandyopadhyay et al., 2021) and recommendation systems (Yassine et al., 2021), as well as in scientific fields like medical diagnostics, astronomy, and soil classification (section 2.3). In steel research, however, its application has been more limited and often focused on surface analysis and production scheduling. For instance, Kitahara and Holm, (2018) applied unsupervised learning and computer vision techniques to classify surface defect images from scanning electron microscopy (SEM). CNNs were used to extract 4,096 image descriptors, which were then reduced using Principal Component Analysis. These compressed representations were visualised using t-SNE and finally clustered using K-Means. Their approach successfully distinguished between human-identifiable surface defects and more subtle fracture features, demonstrating clear cluster boundaries and separation. In another application, Švec et al., (2016) used K-Means and fuzzy clustering for scheduling steel production in continuous casting. By grouping 151 steel grades based on similarities in chemical composition, liquidus temperature, and inter-smelting pot temperature across 27 variables, the study helped optimise casting sequences. The elbow method was used to identify five optimal clusters for production planning.

While these examples show that K-Means has been successfully used in steel research, the gap remains in applying it specifically for the purpose of steel grade rationalisation. Existing studies have mainly focused on grouping steels based on chemical composition, but limited study has concentrated only on clustering steel grades by their mechanical properties. Shifting the similarity assessment from compositional attributes to mechanical properties would provide a

more meaningful basis for identifying functionally equivalent grades and redundant grades (objective 4). Alongside K-Means, other clustering techniques such as Hierarchical Clustering offer alternative ways of assessing similarity, particularly when the dataset may contain nested or hierarchical structures. The following section introduces this approach.

3.4.2 Hierarchical Clustering

Hierarchical clustering is another widely used unsupervised method for grouping data based on similarity (Halkidi, 2009). Unlike K-Means, which partitions the dataset into a predefined number of clusters, Hierarchical clustering builds a nested structure of clusters by iteratively merging (agglomerative) or splitting (divisive) groups of data points. In its agglomerative form as the most common approach (Müllner, 2011), each observation starts as its own cluster, and the algorithm repeatedly merges the two closest clusters until all data points are combined into a single hierarchy. The results are typically represented using a dendrogram, which visually illustrates the sequence of merges and allows the optimal number of clusters to be chosen.

The choice of distance metric plays a critical role, with Euclidean distance being the most common. However, other measures may also be applied, depending on the nature of the dataset. Beyond the distance metric, the linkage criterion determines how distances between clusters are defined. One of the most widely approach is Ward's method (Murtagh and Legendre, 2011), which aims to minimise the total within-cluster variance at each merge.

The cost function for merging clusters A and B is given by:

$$\begin{aligned}\Delta(A,B) = &\sum_{i\in A\cup B} \|\vec{x_i} - \vec{m}_{A\cup B}\|^2 \\ &- \sum_{i\in A} \|\vec{x_i} - \vec{m}_A\|^2 \\ &- \sum_{i\in B} \|\vec{x_i} - \vec{m}_B\|^2 \\ &= \frac{n_A n_B}{n_A + n_B} \|\vec{m}_A - \vec{m}_B\|^2\end{aligned} \tag{3.12}$$

In equation 3.12, A and B determine the distances of the clusters, $\overline{m_j}$ determines the centre of the j cluster. In addition, the number of data points in j cluster determines by n_j. Finally, Δ determines the cost of merging two clusters (A and B). In general, the sum of squared in the Hierarchical clustering approach starts with zero because each data point is considered as a cluster. However, this number will be increased by measuring the distance between each

cluster. The purpose of the Ward method is to make this number as small as possible (Murtagh and Contreras, 2012).

Hierarchical clustering has also been used in steel research. For instance, Zhitenev et al., (2021) applied and compared hierarchical clustering with K-Means to group non-metallic inclusions based on chemical composition. Both methods produced similar results, but hierarchical clustering offered a clearer view of the data structure. In another study, Eroglu and Guleryuz, (2025) used both K-Means and hierarchical clustering along with feature selection and parameter tuning to detect and classify faults in steel plates. Their approach improved prediction accuracy by 18.25 percent compared to standard models.

Hierarchical clustering mostly applied for tasks such as fault detection or compositional grouping. Its computational cost makes it less scalable for large datasets of steel grades, which may explain its absence from rationalisation studies (Gagolewski et al., 2016). Nevertheless, its ability to reveal nested structures offers complementary insights that simple partitioning methods like K-Means may overlook. Alongside Hierarchical clustering, other approaches such as DBSCAN can capture more complex relationships, as introduced next.

3.4.3 Density-based Spatial Clustering of Applications with Noise (DBSCAN)

DBSCAN is a density-based clustering algorithm introduced by Ester et al., (1996) that has become widely applied in domains where clusters may not follow simple geometric structures. Unlike partitioning algorithms such as K-Means, which impose a predefined number of clusters and assume convex, isotropic cluster geometry, DBSCAN derives the cluster structure directly from the spatial distribution of the data. As a result, the number of clusters emerges naturally from the density landscape rather than being imposed as an external parameter. This property is particularly advantageous when analysing datasets characterised by non-spherical, irregular, or heterogeneous cluster shapes, as frequently encountered in real-world scientific data(Perafan-Lopez et al., 2022). DBSCAN core strength lies in distinguishing between dense regions of data and sparse regions, which are classified as noise or outliers.

To operationalise this density-based perspective, DBSCAN evaluates the local neighbourhood of each data point through two key parameters: Eps (ε) which is the radius that defines the neighbourhood boundary, and MinPts which denotes the minimum number of points required to form a dense region. If a point has at least MinPts within its Eps radius, it is considered a core point and becomes the starting point for forming a cluster (Fig. 3.6). Points that are within

the Eps radius of a core point but do not meet the MinPts requirement themselves are called border points. Points that do not belong to any cluster are classified as noise (Fahim, 2023). This structure allows DBSCAN to categorise data into core, border, and noise points, making it robust to outliers and well-suited for real-world applications where data boundaries are not well defined.

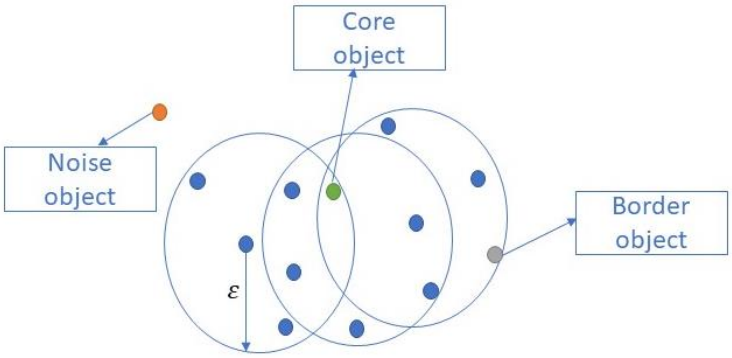


Figure 3.6: Core, border and noise in DBSCAN algorithm.

Despite these advantages, DBSCAN has several limitations. Its performance is highly sensitive to the selection of described parameters (Bushra et al., 2024). Choosing inappropriate values can lead to poor clustering results. Additionally, the algorithm can be computationally intensive, particularly when searching for neighbours during cluster expansion, which becomes costly for large datasets (Lv et al., 2016). DBSCAN also struggles to separate clusters that lie close together but differ in density, which can reduce its effectiveness in heterogeneous datasets.

Nevertheless, DBSCAN has been widely applied in domains where density rather than distance is the key factor for grouping. For instance, it has been used to identify spatial clusters of business activity or crime hotspots (John et al., 2023; Tu et al., 2022). Its use has also extended to steel-related research. Jin et al., (2023) used DBSCAN as the core of the proposed VHCA-DBSCAN anomaly detection method, enabling density-based spatial clustering to identify suspected outliers within multi-dimensional energy data in the steel industry. In similar research, Bordekar et al., (2025) applied DBSCAN to three-dimensional point cloud data derived from pixel-level classification of pores and inclusions using an SVM model. DBSCAN successfully grouped defect points into distinct clusters and enabled their individual characterisation in terms of size, shape, and geometry. This application demonstrates how DBSCAN can move beyond simple detection to support detailed structural analysis.

DBSCAN has shown promise for detecting anomalies and classifying defect structures in steel, especially in imaging and energy data. However, its density-based nature has not yet been leveraged for steel grade rationalisation, where many grades overlap in property ranges. This gap highlights the need to explore whether density-based clustering can reveal functionally similar grades and reduce redundancy in the existing steel classifications.

While different clustering methods provide useful ways of grouping similar items, their effectiveness can be reduced when dealing with high-dimensional steel datasets. To mitigate this, dimensionality reduction methods such as PCA are often introduced.

3.4.4 Principle Component Analysis (PCA)

PCA is a type of unsupervised learning approach which is commonly used for reducing the dataset dimension (Ding and He, 2004). It transforms the raw data variables into a smaller one that contains most of the information of raw data in it.

In other words, PCA transforms the high dimension of data into a smaller dimension to facilitate the model's training and visualisation phases. This process of extracting necessary information is computing by the average degree to which each point differs from the mean (variance). Mathematically, this transformation works by identifying new directions called principal components that capture how the data varies. These directions are found by analysing how different variables in the dataset relate to one another, either by looking at their covariance or by using a method called singular value decomposition (SVD) (Wall et al., 2002). Each principal component points in the direction where the data spreads out the most. By keeping only the first few of these components, the size of the dataset can reduce while still keeping most of the important information. This helps make the data easier to work with and interpret, especially when dealing with lots of variables.

In steel research, PCA is increasingly adopted to reveal hidden patterns in alloy data and assist in feature reduction for downstream machine learning models. For instance, Kim et al., (2020) applied PCA as part of an unsupervised pipeline for microstructural image segmentation. By reducing the dimensionality of morphological descriptors extracted from steel phases (e.g. pearlite, grain boundary ferrite), PCA helped accelerate clustering via the Bayesian Gaussian Mixture Model (BGMM). The resulting unsupervised method achieved accurate classification without requiring labelled image data or expert-driven segmentation rules. In addition, Zhang and Saniie, (2021) proposed a PCA-based method to compress the 3D ultrasonic data using

unsupervised machine learning techniques. In order to compress the data, this study looked for a latent pattern in the data and then eliminated the redundant information in 3D ultrasonic data. This study used different unsupervised machine learning which was included PCA, Incremental-PCA, Independent Component Analysis (ICA), and Exploratory Factor Analysis (EFA) in order to find the latent pattern for compressing the 3D ultrasonic data. PCA has been used in order to reduce the number of 3D ultrasonic features in data. Also, to minimise the training time and involve minimum memory for the training phase, Increment-PCA has been used.

In addition, Cheng and Chen, (2021) used PCA to reduce the number of variables when predicting the tensile strength of hot-rolled strip steel. They started with 28 input features, including both chemical composition and process parameters, and used PCA to reduce these to 8 key components. This helped simplify the data and remove repeated or less useful information. These components were then used as inputs for a Gradient Boosted Decision Tree (GBDT) model. The results showed that using PCA improved the model's accuracy and helped better explain how different inputs affect tensile strength, especially when the relationships were complex and nonlinear.

Based on the review of unsupervised learning algorithms, there is a lack of development in applying clustering methods specifically for steel classification. In particular, none of the studies considered mechanical properties as the basis for identifying similarities among standardised steel grades. This leaves unexplored the possibility of identifying the grades that may deliver equivalent performance despite having distinct chemical compositions. In this study, clustering techniques and dimensionality reduction are applied to support a new classification framework that promotes steel recyclability and sustainability. However, machine learning models are typically considered black boxes, making it difficult to interpret how predictions are made. To address this challenge, the next section explores Explainable AI (XAI) techniques, which aim to interpret model predictions and provide transparency in data-driven steel classification as well as uncover the relationship between chemical composition, processing conditions, and material performance.

3.5 Explainable Artificial Intelligence (XAI)

The increasing adoption of machine learning in materials science has highlighted a challenge that most high-performing models behave as "black boxes", offering little transparency into how predictions are generated. This lack of interpretability creates barriers for industrial

adoption, where scientific likelihood and engineering justification are as important as accuracy. XAI addresses this problem by providing tools to interpret model outputs and clarify how features influence predictions (Mohseni et al., 2021).

XAI was first used to explain the results of an expert system via the applied rules (Swartout, 1981). Since AI research has risen, the popularity of explainable AI methods has increased. At the moment, the goal of explainable AI methods is to help human experts understand why an AI made a particular decision (Pearl, 2019). Figure 3.7 illustrates the evolution of popularity for major XAI libraries on GitHub. The plot shows a clear and sustained rise in adoption across all approaches, with SHAP (SHapley Additive exPlanations) emerging as the most widely used method. This rapid increase reflects SHAP's strong theoretical foundation in cooperative game theory, its ability to provide consistent and locally accurate feature attributions, and its broad compatibility with tree-based models and neural networks. By contrast, earlier techniques such as LIME and rule-based systems exhibit slower growth, partly because they can suffer from instability or limited model coverage. Taken together, these trends demonstrate a shift in both academia and industry toward XAI approaches that offer mathematically rigorous, model-agnostic interpretations, an aspect that is particularly relevant to materials science applications, where model transparency is essential for validating structure–property relationships and supporting engineering decisions.

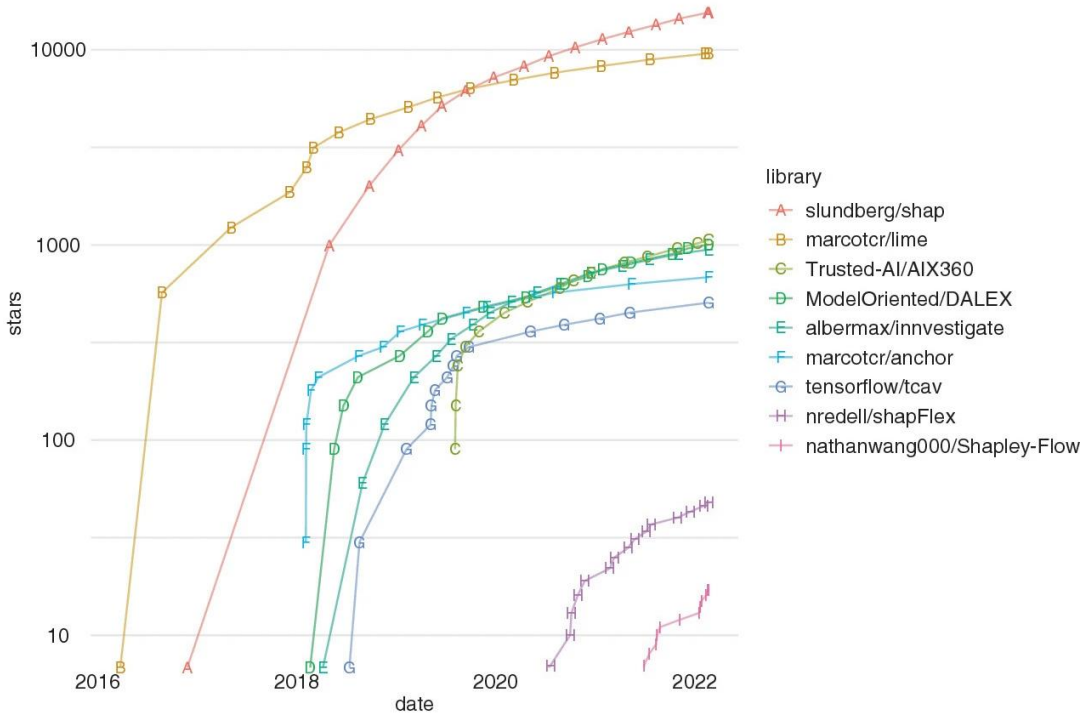


Figure 3.7: The number of stars on GitHub for the most popular repositories of explainable AI approaches since their development (Kolajo and Daramola, 2023).

3.5.1 SHapley Additive exPlanations (SHAP)

The concept of the Shapley value originates from game theory (Shapley, 1953), and it is designed to distribute the contributions of players when they achieve an outcome. In machine learning, Shapley values can quantify the contribution of each feature in a model to the overall prediction (Štrumbelj and Kononenko, 2014). The Shapley value corresponding to feature X_j in a model is calculated in equation 3.13:

$$Shapley(X_j) = \sum_{S \subseteq N \setminus \{j\}} \frac{|S|!\,(p - |S| - 1)!}{p!} [f(S \cup \{j\}) - f(S)] \tag{3.13}$$

Where p is the total number of features, $N \setminus \{j\}$ is the set of all possible combinations of features excluding X_j, S is a subset of features from $N \setminus \{j\}$, $f(S)$ is the model prediction using the features in S, and $f(S \cup \{j\})$ is the model prediction using the features in S plus X_j.

In general, the shapely value method finds out how important each feature is (Lundberg and Lee, 2017). It measures how much each input, like a feature, helps a complex machine learning model work well. Computationally, for a given model and input sample, SHAP evaluates how the predicted output changes when a feature is included versus excluded from different feature subsets, weighting these differences according to the Shapley value formulation. This process yields an additive decomposition of the prediction, where the sum of individual feature contributions equals the difference between the sample's prediction and a chosen baseline (e.g., the mean model output). By averaging local Shapley values across many samples, SHAP produces global importance rankings that are both locally accurate and consistent, meaning that if a model depends more strongly on a particular feature, the assigned importance for that feature cannot decrease.The SHAP method has found applications in various areas of the steel industry. For instance, Jeon et al., (2021) utilised it by training four machine learning algorithms to manage the hardness of low-alloy steels across different tempering conditions. The results showed that the tempering temperature feature had the most significant effect on the variation in hardness. Also, Jeon et al., (2022) applied SHAP to RF to predict austenite-grain growth with greater accuracy and result showed that the reheating temperature was the most important factor affecting austenite-grain size. In addition, Costa et al., (2022) applied two machine learning methods to measure the YS and TS. The SHAP method, based on the RF

model, showed that Fe, Ti, and C are the most important factors for measuring YS, while Fe, P, and Cr were the most important factors for measuring the TS. Moreover, Frie et al., (2024) applied SHAP to interpret a RF model for predicting fatigue strength in steel components. The analysis revealed that elements like chromium and nickel, along with loading type and specimen size, were key factors influencing fatigue performance and failure mechanisms. Furthermore, Jeon et al., (2023) applied SHAP to uncover the key factors influencing a machine learning model designed to predict the martensite start temperature (Ms) of alloy steels. The study evaluated multiple machine learning models, with ANN achieving the highest predictive accuracy. SHAP was used to determine the most significant variables affecting Ms, revealing that carbon had the strongest influence in Ms, followed by manganese, prior austenite grain size, nickel, molybdenum, chromium, copper, and silicon. In addition, (Millner et al., 2023)applied various machine learning techniques alongside SHAP to investigate the impact of different factors on the mechanical properties of steel sheets, including tensile strength, yield stress, and elongation. The study evaluated multiple ML models, where XGB and Ridge models selected for SHAP analysis due to their computational efficiency and strong performance. Through SHAP analysis, the study identified chemical composition (including Mn, Nb, and Cr) as a primary factor influencing mechanical properties, followed by annealing parameters and rolling conditions. Additionally, annealing temperature and nitrogen content were found to play a crucial role in determining yield stress, tensile strength, and elongation.

Collectively, these studies demonstrate SHAP's strength in linking alloying elements and processing parameters to mechanical outcomes, thereby bridging predictive accuracy with metallurgical interpretability.

3.5.2 Local interpretable model-agnostic explanations (LIME)

LIME is another popular XAI approach designed to explain the predictions of any classifier or regressor in a reliable and interpretable way (Ribeiro et al., 2016). It approximates the complex model with a simple, interpretable surrogate model around the instance of interest. This allows researchers to identify the specific factors influencing a single prediction, rather than the overall model (Nguyen et al., 2021).

LIME works by generating synthetic samples near the instance being explained and obtaining predictions for these samples from the original model. These samples are then weighted based on their proximity to the instance being analysed. Using these weighted samples and their corresponding output probabilities, LIME constructs a linear model to approximate the original

model's behaviour locally. The weights of this model are used to determine the importance of input features for the specific prediction.

LIME has been effectively used in the medical domain to interpret model predictions. For instance, Alabi et al., (2023) applied LIME to explain how tumour stage, metastasis, and patient demographics influenced survival predictions. In similar study, Wu et al., (2023) used LIME with deep learning models to identify key predictors of heart disease and diabetes, such as cholesterol levels, blood pressure, glucose, and BMI. LIME has also been used in materials science to interpret machine learning models. Park et al., (2022) applied it to explain how chemical composition and processing steps affect mechanical properties in 7xxx aluminium alloys, finding zirconium had the most positive impact on strength and ductility. Tiwari et al., (2023) used LIME to interpret clustering results in 6xxx alloys, identifying magnesium, silicon, and copper as key drivers of yield strength. In addition, Swateelagna et al., (2024) applied LIME to reveal that enthalpy of mixing, atomic size difference, and valence electron concentration were key factors influencing phase formation in high-entropy alloys.

The key distinction between LIME and SHAP lies in both their theoretical foundations and the scope of explanations they provide. LIME constructs a simple surrogate model (usually linear) in a small neighbourhood around the instance of interest, yielding local, model-agnostic explanations that approximate the black-box behaviour only in that region. In contrast, SHAP computes feature attributions based on Shapley values from cooperative game theory, which satisfy properties such as local accuracy and consistency and can be aggregated across many samples to obtain stable global importance rankings and dependence relationships. As a result, LIME is well suited for analysing individual predictions with low computational cost, whereas SHAP provides both local and global interpretability with greater theoretical guarantees and across-dataset consistency in feature attributions (Alabi et al., 2023). Due to SHAP's advantage over LIME in providing both local and global interpretability, this study applied SHAP to analyse the relationship between chemical composition, processing conditions, and mechanical properties of steel grades. In addition to interpreting the results of supervised learning models, SHAP analysis was also used to support the understanding of steel grade clustering based on mechanical properties. This enables to explain the formation of newly proposed steel grade groups and assess their relevance for improving recyclability and sustainability.

While explainability enhances the interpretability of machine learning outputs, a further challenge in this study is how to reduce the number of steel grades once groups of functionally

similar grades have been identified. The next section therefore reviews data reduction strategies, with particular emphasis on methods for detecting redundant or equivalent records that can inform the rationalisation of steel grades.

3.6 Reduction Process

Reduction and overlapping the number of options in the dataset are major problems nowadays. For instance, in manufacturing, different alloy grades can perform similarly in the production phase but differ in their attributes (Tiwari et al., 2023). Also, in healthcare, some medications might work the same under certain conditions but have different ingredients and production methods (Groza et al., 2021). Data reduction techniques aim to address this by eliminating redundancy while preserving the essential variability of the dataset (Meng et al., 2016). In data science, such approaches range from dimensionality reduction to compression methods, and deduplication strategies that identify functionally equivalent records.

This section reviews relevant data reduction techniques, including dimensionality reduction and clustering deduplication, which lay the foundation for developing a simplified and more sustainable steel classification system.

3.6.1 Dimension Reduction Techniques (DR)

In a dataset, certain features or attributes may be not important for machine learning. Some attributes are irrelevant, while others have no significant impact on the outcomes of machine learning algorithms. The purpose of dimension reduction techniques is to compress the dataset by transforming it from a higher-dimensional matrix to a lower-dimensional one while maintaining the original data meaning (Sadegh-Zadeh et al., 2024). This not only reduces computational cost but also enhances clustering performance and enables clearer visualisation of similarity patterns.

Broadly, DR techniques can be categorised into global and local approaches. Global methods aim to preserve relationships across the entire dataset, capturing large-scale variance or independence structures, whereas local methods focus on maintaining the neighbourhood relationships of each point with its nearest neighbours.

Global techniques such as PCA and Independent Component Analysis (ICA) retain structure at the dataset level (Wang et al., 2021). PCA, already introduced in Section 3.4.4, identifies orthogonal directions that maximise variance, while ICA extends this idea by finding

components that are statistically independent, which has proved useful in tasks such as hyperspectral image analysis (Jayaprakash et al., 2018), financial forecasting (Grigoryan, 2016), and bioinformatics for denoising high-dimensional cancer datasets (Jimoh et al., 2018). Both PCA and ICA are particularly effective when the underlying data relationships are largely linear.

However, some data like steel data often contain non-linear interactions between variables. In such cases, local DR methods such as t-distributed Stochastic Neighbour Embedding (t-SNE) (Van Der Maaten and Hinton, 2008) and Uniform Manifold Approximation and Projection (UMAP) are more suitable. t-SNE, converts pairwise distances into conditional probabilities to preserve local neighbourhood structures. It has been widely applied in domains such as image classification (Ren et al., 2020), materials science (Liu et al., 2024), and biological data exploration (Liu et al., 2021). Despite its strength in capturing local fidelity, t-SNE struggles with preserving global structure and is computationally demanding. To overcome these limitations, UMAP offers a graph-based alternative that preserves both local and global structure more effectively than t-SNE. It assumes the data lies on a Riemannian manifold and constructs a weighted graph to represent it. UMAP has been applied in domains ranging from COVID-19 mutation visualisation (Hozumi et al., 2021) to time-series clustering (Péalat et al., 2022), showing its versatility across complex datasets. This characteristic reflects a better balance between preserving global and local data structures, making the resulting visualisations easier to interpret.

Nevertheless, despite the advantages of newer nonlinear methods such as UMAP, PCA remains widely used in materials informatics due to its strong theoretical grounding, interpretability, and extensive validation across composition–process–property datasets (Section 3.4.4). Studies have shown that transforming correlated variables into orthogonal principal components helps to concentrate the dominant variance into fewer dimensions and reduce noise arising from multicollinearity. This is particularly beneficial when clustering materials, as PCA prevents highly correlated variables from being over-weighted in distance-based algorithms and generally improves cluster separability and stability. For these reasons, PCA continues to serve as a standard preprocessing step in steel and alloy data analysis. While dimensionality reduction focuses on simplifying the feature space by removing irrelevant or redundant variables, another stream of data reduction targets the size of the dataset itself. This is achieved through data compression techniques, which are outlined in the following section.

3.6.2 Data Compression

Data compression is another data reduction technique which reducing the size of the data to optimise storage and transmission time (Dantas et al., 2024). These techniques can uncover and utilise patterns in the original data and compact it to the smallest size. This makes compression a size-oriented rather than structure-oriented approach to reduction.

The origin of data compression dates back to 1983 and is linked to Morse code, where letters were compressed for telegraph communication (Hamming, 1986). The study of data compression became popular in the late 1940s when a systematic approach to allocating code words based on block probabilities was developed by Claude Shannon and Robert Fano. This method was later optimised and improved by David Huffman in 1951 (Huffman, 1952).

Compression techniques are broadly divided into lossless and lossy methods (Fitriya et al., 2017). Lossless compression is the process of converting the original data into compressed format without reducing the loss of information of the original data. On the other hand, lossy compression converts the original data into a compressed format with some differences, but these differences are designed to ensure that essential information from the original data is not lost.

These techniques have used in varied applications. For instance, text compression requires lossless methods to avoid semantic distortion (Sayood, 2018), whereas image and video compression often tolerate lossy techniques by exploiting recurring pixel or frame patterns (Al-Shaykh and Mersereau, 1998; Sayood, 2012). In medical imaging and archiving, accuracy is paramount, making lossless techniques more suitable (Xin and Fan, 2021).

The limitation of compression methods is that they primarily address data size rather than structural redundancy. While effective for unstructured formats such as text, images, and audio (Pu, 2006), they are less relevant for structured datasets like those used in this study, where the size of the data is not a concern. For this reason, compression is briefly reviewed here but is not applied in the steel grade rationalisation framework.

3.6.3 Data Deduplication

In the context of this study, data deduplication is especially relevant, as it reflects the core objective of identifying and reducing redundant steel grades with overlapping performance characteristics. Unlike traditional compression methods that focus on reducing file size,

deduplication methods operate at the record level and seek to detect and remove entries that are duplicate or near-duplicate in terms of their attributes, thereby preserving functional diversity while eliminating unnecessary variation.

Data deduplication, also known as record linkage or entity resolution, involves detecting and removing duplicate records that refer to the same real-world entity (Steorts et al., 2014). It is widely used in various domains to reduce storage demands, improve data quality, and enhance system performance. In cloud computing environments, deduplication reduces both storage and bandwidth requirements by replacing duplicate copies with references to a single instance (Mandagere et al., 2008; Hovhannisyan et al., 2018). Many major services such as Amazon S3, Dropbox, and Microsoft Azure implement deduplication to optimise performance and enhance storage utilisation (Kaur et al., 2018).

In large-scale storage systems, deduplication is commonly implemented in three stages: chunking, duplicate identification, and then storing (Meister et al., 2012). In this process, backup files are divided into chunks, and a specific fingerprint is given to each chunk. A system identifies a chunk's uniqueness or redundancy through a search for its presence in a store environment. Chunks that are not redundant are stored, and redundant chunks are eliminated. Unlike traditional data compression techniques, deduplication eliminates redundant information not only between individual files but between groups of files together.

Many studies have developed an approach to identify and remove duplicate data. For instance, in early studies, Jagadish et al., (1999) proposed an algorithm called fascicles for the semantic compression which clusters records sharing similar values under defined constraints. In addition, Khan et al., (2012) proposed a numeric data conversion technique combined with K-Means clustering to reduce record comparisons. A divide-and-conquer strategy was then applied to detect duplicates within each cluster, demonstrating improved accuracy and computational efficiency. Experimental results on a restaurant dataset have proven that the algorithm is working as it allows for duplicate removal as well as the right balance between clustering efficiency and accuracy.

Moreover, Devi and Thigarasu, (2015) presented an advanced method that combines Multiple Hidden Markov Models (MHMM) with fuzzy clustering to detect duplicates in multi-column datasets. Their approach grouped similar records using fuzzy logic, improving flexibility in identifying partial matches. However, low true positive rates were initially a limitation. To address this, they included semantic similarity through Fuzzy Ontology, an ontology-based

framework that encodes domain concepts and their relationships using fuzzy logic, enabling the system to recognise similarities based on meaning rather than exact matching. This enhancement significantly improved detection precision and recall. Omar et al., (2022) applied a random forest approach to detect duplicates in the RLdata500 dataset. They used binning to convert continuous numeric values into categorical ranges and applied phonetic encodings like Soundex and Onca to handle textual variations. The method generated a proximity matrix from tree-based co-occurrence, with similarity thresholds used to flag duplicates. By removing the need for labelled training data, their approach offered an efficient and scalable solution suitable for real-world, diverse datasets.

Padmanaban et al., (2013) presented a three-phase approach combining Q-gram-based string similarity, feature selection using Particle Swarm Optimisation (PSO), and classification via Naïve Bayes. Initially, similarity scores such as the Dice coefficient and Tversky index were calculated using Q-grams. PSO then identified the most informative features, reducing the dimensionality of the problem. Finally, a Naïve Bayes classifier was trained to distinguish between duplicate and non-duplicate records. This approach, tested on Restaurant and Cora datasets, achieved 89% accuracy, showing that optimising similarity features can effectively reduce redundant records. Extending this study, Hamad and Salih, (2016) applied a hybrid method that enhanced Q-gram similarity scoring with an artificial neural network (ANN). Each record was first assigned a unique key, and Q-gram-based comparisons identified potential duplicates. The ANN then refined this selection to distinguish true matches from false positives. This method, evaluated across multiple data warehouses, achieved a 96.94% accuracy rate, outperforming the earlier approach by improving detection precision.

In the bioinformatics domain, Chen et al., (2015) explored the BARDD approach, which used association rule mining on metadata and sequence similarity to identify duplicates. Although effective in previous experiments, its performance significantly dropped when applied to a larger GenBank dataset, identifying just 0.2% of duplicates. In addition, Koumarelas et al., (2020) focused on improving the duplicate detection accuracy by systematically applying data preparation techniques prior to the detection process. Their methodology included standardising formatting, removing special characters, and normalising entries like addresses. Moreover, Shahri and Shahri, (2006) developed a six-step deduplication framework including clustering, attribute selection, similarity measurement, and fuzzy rule application. Their method also included defining linguistic terms and a user-guided merging step to determine representative records. The system's flexibility allowed for customised configurations, making

it suitable for large-scale data integration where varying criteria and user-defined preferences must be accommodated.

With the rise of Generative AI, duplicate detection in text data has seen significant advancements compared to traditional entity matching methods. A recent study by Ormesher, (2024) proposed an advanced duplicate detection method using Large Language Model (LLM) and sentence embeddings. In this approach, the LLM encodes each record into a dense numerical vector, known as a sentence embedding that captures semantic meaning rather than exact lexical similarity. This allows records that "mean the same thing" to be positioned close together in vector space, even if they differ in wording. By transforming structured records into unified text strings and embedding them into vector space, the approach applies clustering (DBSCAN) to identify semantically similar entries. Tested on real-world datasets, this method outperformed traditional entity matching techniques, demonstrating the effectiveness of embedding-based deduplication.

While most existing deduplication studies focus on identifying redundant textual records or resolving syntactic similarity, the concept of deduplication has not been explored in materials engineering contexts. In this study, the objective is not to remove exact duplicates but to rationalise the steel system by identifying grades whose mechanical property ranges substantially overlap. Unlike previous approaches that apply clustering primarily to minimise computational load, this study uses clustering to group steel grades with similar mechanical property profiles. This enables comparisons within clusters to identify whether grades with lower recyclability can be replaced by those that are more recyclable but functionally equivalent. This addresses a key gap in the literature which is the absence of deduplication frameworks that include sustainability constraints such as tramp element content and critical raw materials alongside performance overlap. By extending deduplication beyond textual domains and integrating domain-specific priorities, this study offers a novel contribution to sustainable, data-driven classification.

3.7 Summary

Machine learning has been widely applied in steel research to address diverse challenges, including property prediction, microstructure classification, process monitoring, and defect detection. Supervised learning methods, such as ANN, RF, and XGBoost, have consistently demonstrated strong performance, particularly in modelling mechanical properties and supporting quality assessment tasks such as continuous casting and defect diagnosis. These

approaches were well-suited to quantifying correlations between composition, processing, and performance, but they remain limited to prediction tasks and do not address hidden similarities or redundancies between grades. This limitation motivates the complementary use of unsupervised learning.

Unsupervised learning, including K-Means, HC, and DBSCAN, has been used to group steels based on microstructure, composition, or production parameters for revealing structural patterns in the data. However, existing applications are limited to clustering and do not extend to systematic grade rationalisation. In particular, limited studies have coupled clustering with automated reduction strategies capable of eliminating redundant grades while considering multiple criteria such as property coverage, tramp element tolerance, and CRM dependency.

XAI techniques, particularly SHAP, have contributed to improving transparency in steel modelling, enabling the identification of chemical and processing factors most relevant to predictive outcomes. At the same time, data reduction techniques such as dimensionality reduction and deduplication have been applied in other domains to simplify complex datasets. However, deduplication approaches in prior work have generally relied on a single criterion and have not been adapted to materials science, where functional equivalence must be balanced with sustainability considerations.

To summarise, prior research has demonstrated progress in prediction, classification, and exploratory grouping, but a critical gap remains. A limited framework currently integrates clustering, explainability, and multi-objective reduction for the purpose of rationalising steel grades. This study addresses this gap by developing an application-driven classification approach designed to simplify the steel system while preserving mechanical performance and enhancing recyclability. The next chapter introduces the proposed methodology developed to meet this objective.

Chapter 4: Method I - Data collection, Exploratory Analysis, and Machine Learning-based Relationship

This chapter presents the first methodological stage, focused on collecting, pre-processing, and exploring steel grades data. The goal is to ensure the dataset is standardised, interpretable, and analytically suitable for the advanced clustering and grade rationalisation framework introduced in next chapter.

The chapter begins by outlining the data acquisition process using publicly available databases and international standards. It then describes pre-processing techniques, including standard scaling, feature encoding, and clustering suitability assessment using the Hopkins statistic, applied to examined steel datasets to harmonise variables and assess whether meaningful groupings exist among the existing steel grades. This is followed by a machine learning-based exploratory modelling approach to assess how chemical composition and processing variables influence mechanical performance across traditional steel classes. These steps provide a foundation for the development and application of the KMRP clustering methodology introduced in chapter 5.

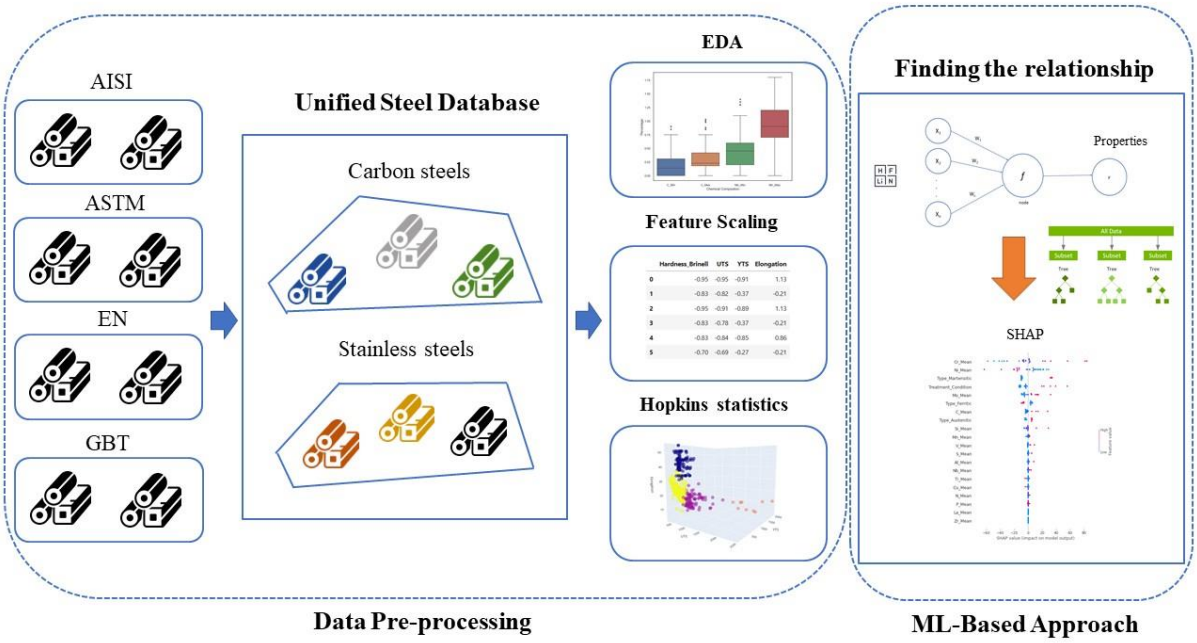


Figure 4.1: Workflow of the first methodological stage, covering data acquisition, pre-processing, and exploratory machine learning analysis.

4.1 Data Collection

This study focuses on carbon and stainless steel grades due to their widespread industrial use and strategic importance, as discussed in chapter 2. Data were collected from publicly available international standards and databases, with primary sources including MatWeb (MatWeb, 2023), MakeItFrom (MakeItFrom.com, 2023), and Steel-Grades (S&G, 2024). These platforms were selected based on their comprehensive coverage of commercially available steel grades and their ability to provide detailed technical specifications, including chemical composition, mechanical properties, and processing history.

Matweb: It was established in 2011 and is well-known as one of the most widely used material properties databases available online, valued for its simplicity and the trust it has built among users.

MakeItFrom: It is a material database that gathers its information from literature, academic sources, and supplier documentation. It provides a broad collection of material properties across various categories. Unlike MatWeb that has a search engine, MakeItFrom organises materials into categories, allowing users to explore and navigate through different material types manually.

Steel-Grades: S&G was launched in 2010 and it has become a leading resource for steel grade information, offering data on more than 2000 steel grades and over ten million material properties from over 30 countries. Its strength lies in its wide coverage of international steel standards. However, unlike MatWeb and MakeItFrom, it offers only general information on processing conditions and property ranges, lacking detailed variations. Nonetheless, it remains a key resource for standardised steel grade data.

Because no single database provided a complete profile for all steel grades within the traditional classification systems, the three sources were used in a complementary manner. In many cases, a steel grade listed on one platform lacked specific details such as property values or processing history. In addition, not all grades were available across all platforms and required cross-verification to ensure complete coverage of standard carbon and stainless steel grades. This cross-referencing strategy enabled the construction of a unified and comprehensive dataset, suitable for the analytical tasks of this study.

To ensure consistency across these heterogeneous reporting formats, a set of harmonisation procedures was applied to the integrated dataset. Mechanical properties were standardised to

MPa for UTS and YS, and to Brinell Hardness (HB) for hardness. When hardness values were reported in Vickers (HV) or Rockwell (HR) scales, they were converted to HB using established metallurgical conversion relationships (Lyman, 1948). Composition values reported as ranges (min–max) were retained in their original form, while differences in processing descriptions (e.g. “annealed”, “normalised”) were mapped into a common binary schema (e.g. annealed = 1, not annealed = 0).

In addition, because multiple international standards describe equivalent steels under different naming conventions, several grades appeared more than once across data sources. To ensure transparency and avoid introducing bias into the clustering analysis, all potential cross-standard duplicates were systematically reviewed. Duplicate detection was performed by comparing chemical composition, processing conditions, and mechanical properties; when all three fields matched, the entries were treated as representing the same underlying material. In such cases, only one record was retained, and all corresponding designations were merged into the one grade name. Table 4.1 summarises the consolidated duplicates used in this study.

Table 4.1: Summary of consolidated cross-standard duplicate steel grades.

AISI designation	***EN designation***	***UNS designation***	***Final Name Used***
AISI 304	*EN 1.4301*	*UNS S30400*	*AISI 304*
AISI 316	*EN 1.4401*	*UNS S31600*	*EN 1.4401*
AISI 410	*EN 1.4006*	*UNS S41000*	*EN 1.4006*
AISI 1010	*EN C10E*	—	*AISI 1010*
AISI 1045	*EN C45E*	—	*AISI 1045*
AISI 440C	*EN 1.4125*	*UNS S44004*	*EN 1.4125*

As discussed in chapter 2, global estimates indicate that more than 3,500 steel grades exist across international standards (Javaid and Essadiqi, 2003). However, a substantial proportion of these grades either lack publicly available mechanical property data, provide only partial chemical composition, or present incomplete information on processing routes. In many cases, the same grade appeared across different databases with inconsistent or missing fields. Such incomplete records cannot support reliable clustering, property comparison, or rationalisation and therefore were excluded from this study.

For this reason, the dataset for this research was intentionally restricted to grades with complete and comparable information. Each record was required to contain:

(1) full chemical composition,

(2) a defined processing route, and

(3) complete mechanical property information (YS, UTS, elongation, or hardness).

Applying these strict inclusion criteria resulted in a final dataset containing 368 unique alloy grades, represented by 541 individual records and 43 attributes. To ensure transparency and enable full verification of the analysis, the complete dataset used in this study has been archived in a publicly accessible Zenodo repository (Jalalian, 2025) under DOI: 10.5281/zenodo.17704410.

The dataset comprises 146 carbon and 288 stainless steels, spanning all major industrial categories, including low, medium, and high carbon steels, and austenitic, ferritic, and martensitic stainless steels. Although this dataset does not encompass all 3,500+ global steel grades, the majority of omitted grades lacked full, publicly accessible specifications necessary for analytical reliability.

The dataset also contains an uneven representation of steel types, with stainless steels more than carbon steels. This imbalance reflects data availability rather than intentional sampling bias. To mitigate potential analytical bias, all machine-learning models were trained on standardised features, and clustering was performed in a normalised mechanical property space. Moreover, grade reduction and interpretation were conducted within steel types rather than through direct comparison between carbon and stainless steels, ensuring that decisions were driven by performance similarity within comparable metallurgical families rather than by class frequency. Results are therefore interpreted both at the cluster level and within traditional steel categories to avoid over-generalisation.

Thus, the selected dataset provides a representative, diverse, and methodologically robust foundation for developing and validating the steel grade rationalisation framework. A summary of the attributes, corresponding examples, and data types is presented in Table 4.2. The attributes are divided into two categories of independent (I) and dependent (D).

Table 4.2: Carbon and stainless steel data structure used for the present study.

Attribute	***Example***	***Data Type***
Chemical Composition (I)	C, Mn, P, S	Float
Mechanical Properties (D)	Hardness, Yield, UTS, Elongation	Int
Processing Condition (I)	Annealing, Quenching	Boolean
Standard	ASTM, AISI, EN, GB	String
Type	Carbon, Stainless	String
Grade Name	AISI 1010, ASTM A240	String

Chemical composition data included 15 elements such as C, Mn, Cr, Ni, and Cu, with minimum and maximum values recorded to capture their variations across grades. These elements form the foundation for modelling mechanical performance and were also critical for evaluating recyclability and sustainability particularly in relation to tramp element tolerance and the presence of CRMs. Summary statistics such as mean, standard deviation, and range are presented in appendix A to provide a detailed overview of their distribution within the dataset.

In addition to chemical composition, processing conditions play a key role in shaping the mechanical properties of steel grades and are therefore treated as independent variables in this study. Because processing information was reported inconsistently across ASTM, UNS, AISI, EN and GB standards, typically without temperatures, times or cooling rates, only the descriptive treatment labels provided by the sources were used. Fourteen processing treatments were identified during data collection, and as described in the harmonisation step, all were standardised into a binary format (treatment applied = 1, not applied = 0) to ensure consistency across sources. As summarised in appendix A, processing treatments are well distributed across the dataset, with most being applied to a significant portion of the records. For instance, annealing are presented in 53% of entries, while cold rolling and quenching are observed in over half of the samples. This distribution indicated that no single treatment dominates the dataset, ensuring a balanced representation of common industrial processing routes. Such balance is important for allowing fair comparative analysis and for developing reliable machine learning models that can learn the true effects of processing treatments on mechanical properties, without being biased toward a particular treatment type.

The dependent attributes in this study are limited to four key mechanical properties that are crucial for meeting application requirements including hardness, YS, UTS, and elongation. As discussed in chapter 2, these mechanical properties are critical for determining the strength, durability, and application suitability of different steel grades. Crucially, they were also the only properties consistently reported in complete, numerical form across all three data sources used in this research. Other important performance properties, such as toughness, weldability, formability, and corrosion resistance, were reported inconsistently across standards, often qualitatively (e.g., “good weldability”, “excellent corrosion resistance”) or not reported at all. Including such incomplete or non-standardised properties would have introduced bias and compromised the reliability of the clustering and rationalisation framework.

For these reasons, the present analysis focuses on the four mechanical properties while acknowledging that extending the framework to incorporate additional property domains remains an important direction for future work, once reliable data become available. Table 4.3 provides a statistical overview of these properties, including their minimum, maximum, mean, and standard deviation within the dataset.

Table 4.3: Summary of mechanical properties in the dataset.

Mechanical property	***count***	***Mean***	***Std Dev***	***Min***	***25%***	***50%***	***75%***	***Max***	***Type***
Hardness (Brinell)	538	190.0	46.0	49	143	180	210	710	Int
UTS (MPa)	538	606.1	241.0	123	470	580	680	2450	Int
YS (MPa)	538	379.3	230.6	131	240	310	430	2050	Int
Elongation (%)	538	27.2	11.6	2.0	18.0	25.0	37.0	55.5	Float

4.2 Exploratory Data Analysis

In this step, exploratory techniques and data visualisation methods have been utilised to examine the dataset, to facilitate a better comprehension of the variables and their interrelationships. These include the correlation analyses, and visualisation techniques, aiming at identifying patterns, outliers, and potential limitations in the dataset. Since this study focusing on two different steel groups including carbon and stainless steels, the analysis of steel grades in each group is conducted separately to gain better insights from the data.

4.3 Data Pre-processing

Data pre-processing techniques was performed in this phase to prepare the data for machine learning algorithms. This process included examining null values, assessing the clustering tendency of the data using the Hopkins statistic, and scaling the data to ensure it is properly formatted for machine learning models. The details of these pre-processing steps and their mechanisms are explained below.

Data cleaning is a crucial pre-processing step to ensure dataset quality, particularly when dealing with missing values, inconsistent entries, or formatting issues (Prakash, 2024). In this study, special attention was given to handling missing values, as discarding entire rows with incomplete data would have led to a significant loss of information due to the limited number of available records. To address this, two imputation strategies were applied based on the distribution characteristics of each variable. For features showing approximately symmetric

distributions, the mean was used to fill in missing values, as this approach preserves the central tendency without introducing bias. However, for features with skewed distributions (a common scenario in mechanical property data such as elongation or UTS) the median was used instead. This strategy is recommended in the literature for handling non-normal data, as it is more robust to outliers and skewness (Satyam Kumar, 2020).

Hopkins statistics is used to examine the clustering tendency of a dataset by calculating the likelihood of a uniform / randomised data distribution (Banerjee and Dave, 2004). The Hopkins statistic ranges from 0 to 1, where values close to 0 suggest that the dataset is not well-suited for clustering, while values near 1 indicate that the dataset has a strong clustering tendency and is suitable for clustering analysis. Hopkins statistic is applied to assess whether steel grades in the dataset demonstrate the natural groupings based on their mechanical properties. The mechanism of the Hopkins statistics is as follows:

1. Randomly select n steel grades from the dataset D, represented as $\{x_1, x_2, \ldots, x_n\}$.
2. Generate n synthetic points within the same feature space, denoted as $\{y_1, y_2, \ldots, y_n\}$.
3. Measure the distance (d_i) from each selected steel grade (x_i) to its nearest neighbour in the dataset D, for i = 1, 2, …, n.
4. Measure the distance (e_i) from each synthetic point (y_i) to its nearest real steel grade in the dataset D, for i = 1, 2, …, n.
5. Calculate the total sums of W and U, where W is the sum of all e_i values (distances from synthetic points) and U is sum of all d_i values (distances from real steel grades).
6. Calculate the Hopkins Statistics (H):

$$H = \frac{W}{W + U} \tag{4.1}$$

Feature scaling is used to rescale the values of variables, so they share a standard scale, and to make ML methods converge more quickly. Each attribute in the scaled dataset has been transformed to have a mean of zero and a standard deviation of one (Milligan and Cooper, 1988). In terms of mechanical properties, elongation is expressed as a percentage, while UTS and YS are measured in MPa. To ensure consistent measurement scales, mechanical properties need to be scaled properly. This is essential for accurately calculating distances between different steel grades and for the ANN algorithm to effectively predict mechanical properties.

4.4 Investigating the Influence of Chemical Composition and Processing on Mechanical Properties

While earlier sections of this chapter focused on data preparation, this section transitions into the analytical phase, where key relationships between input variables (composition and treatment processes) and target mechanical properties are investigated. The objective here is to reveal both data-driven patterns and metallurgical consistencies, leading to the identification of which features most strongly differentiate steel grades in terms of performance. These insights provide the necessary foundation for the clustering and rationalisation framework developed in chapter 5.

4.4.1 Neural Network Architecture

To model the relationship between chemical composition, processing conditions, and mechanical properties, a Multilayer Perceptron (MLP) neural network was implemented (Bebis and Georgiopoulos, 1994) based on insights from the literature review in chapter 3. The input layer received 30 features, including chemical compositions and binary indicators of processing treatments, while the output layer predicted four target properties including hardness, UTS, YS, and elongation.

The architecture of ANN consisted of 30 features including chemical composition and processing condition representing input layer, while the output layer contained the four mechanical properties (Table 4.2). The network included three hidden layers with 256, 128, and 64 neurons, respectively, selected through trial and error to balance prediction accuracy and model complexity. The hidden layers employed the ReLU activation function, chosen over sigmoid and tanh because the steel-property dataset contains wide numerical ranges and highly non-linear interactions between alloying elements and processing routes. ReLU prevents gradient saturation, preserves stable gradients for positive values, and enables faster and more reliable optimisation in deep networks (Nair and Hinton, 2010). Additionally, a 20% dropout has been applied to prevent overfitting by randomly deactivating neurons during training, ensuring the model generalises better (Srivastava et al., 2014). Model training was performed using backpropagation with the Adam optimiser (Kingma and Ba, 2014). Adam was selected because it adaptively adjusts learning rates for individual parameters, offering robustness when feature scales differ substantially, such as ppm-level alloying elements and MPa-level mechanical properties (Xue et al., 2022). During model development, Adam consistently

produced faster and more stable convergence than standard stochastic gradient descent, which further justified its use for this multi-output regression task.

This architecture was designed after iterative experimentation to balance model complexity and predictive accuracy. The model was not intended for classification or decision-making tasks, but rather for exploratory analysis aimed at quantifying how chemical composition and processing conditions influence mechanical properties. By evaluating prediction performance and later interpreting feature contributions, the neural network provides insights into which input variables most strongly affect the variation in properties within existing steel classifications. The same architecture will later be reused in chapter 9 to evaluate the predictive validity of synthetically generated alloys, proposed as part of the future direction of this research.

4.4.2 XGBoost Architecture

To benchmark ANN model performance and identify the most suitable approach for evaluating mechanical properties from chemical composition and processing data, the XGBoost algorithm was implemented (Chen and Guestrin, 2016b). XGBoost was selected as a comparison point due to its high predictive accuracy, efficiency, and robustness in handling non-linear feature interactions between chemical composition, processing conditions, and mechanical properties, making it a valuable tool for uncovering patterns in steel grades.

The following steps were implemented in this study, following the framework described by (Sandeep et al., 2023):

1. **Setting an initial estimate:** The model begins with a simple initial prediction, usually the average value of the mechanical properties (e.g., hardness, UTS, YS, and elongation). This serves as the model's first estimate before any learning takes place.
2. **Identifying prediction errors:** The model then calculates the difference (residuals) between the actual mechanical property values and the initial predictions. These residuals indicate how much the predictions need to be adjusted to be more accurate.
3. **Training the first decision tree:** Instead of predicting the target values directly, the first tree is trained to learn the baseline errors. It looks for the best split points in chemical composition and processing conditions that help minimise these errors.

4. **Improving with additional trees:** Rather than starting over, XGBoost adds new trees sequentially, each learning from the mistakes of the previous trees. This step-by-step correction process continuously refines the predictions.
5. **Optimising the learning process:** During training, XGBoost minimises a loss function, such as Mean Squared Error (MSE) (Battaglia, 1996), to ensure the model is effectively learning from the data and making better predictions.
6. **Avoiding overfitting:** To keep the model from memorising noise, XGBoost applies L1 and L2 regularisation (Demir-Kavuk et al., 2011), preventing trees from becoming overly complex and ensuring they generalise well to unseen data.
7. **Determining when to stop:** The model continues adding trees until it reaches a stopping condition. In this study, early stopping was applied, meaning the training process stopping when further improvements in validation accuracy were minimal.

To optimise the performance of the XGBoost model, several hyperparameters fine-tuned that control the complexity, learning process, and regularisation of the model. A relatively large number of boosting rounds (n_estimators = 1400) was selected to allow the model to learn subtle patterns in the data, while early stopping after 600 rounds avoided unnecessary training and overfitting. The tree depth was limited to four (max_depth = 4) to control model complexity and prevent the algorithm from overfitting to noise in a moderately sized dataset. A small learning rate (0.010) was used to allow gradual refinement of predictions, helping the model converge more reliably (Gonsalves and Upadhyay, 2021). To further regularise the learning process, randomness introduced through subsampling (subsample = 0.85) and feature sampling (colsample_bytree = 0.7), which prevents the model from becoming overly dependent on specific training instances or features. L1 (reg_alpha = 3.5) and L2 (reg_lambda = 8.0) regularisation were applied to control overfitting and promote sparsity, particularly important when modelling with many compositional and processing variables. Finally, a minimum child weight of six (min_child_weight = 6) ensured that each decision node was supported by a sufficient number of training samples, reducing the likelihood of splits that capture random variation (Tune XGBoost “min_child_weight” Parameter, 2024).

These hyperparameters were selected through trial and error to balance accuracy, efficiency, and generalisation. The model was trained separately for each mechanical property (hardness, UTS, YS, and elongation).

4.4.3 Model Evaluation and Performance

This section evaluates the predictive performance of the developed models in estimating the mechanical properties of steel grades. Two primary error metrics were applied including the MAE and Standardised MAE.

MAE calculates the average absolute difference between the predicted values and the actual target values. It is defined as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \qquad (4.2)$$

In equation 4.2, y_i represents the actual mechanical property value and $\hat{y}_i$ is the predicted value and n is the total number of steel grades in the dataset.

Due to the different units used for each mechanical property, we need to standardise the MAE results to better understand the model's error. This is done as follows:

$$Standardised\ MAE = \frac{MAE}{Mean\ of\ each\ property} \qquad (4.3)$$

In equation 4.3, It ensures that elongation errors are measured on the same scale as UTS and YS.

In addition, since the examined dataset is small, traditional train-test splits could lead to issues like overfitting or underfitting. To ensure a more robust evaluation, the k-Fold Cross-Validation technique (Refaeilzadeh et al., 2009) was used. The data was partitioned into k subsets where each model was trained on k−1 folds and tested on the remaining one, with the process repeated k times. Final metrics were averaged across all folds to produce stable evaluations.

In addition, to evaluate the performance of the models, R^2 applied, which assesses how well the model explains the variance in mechanical properties (Goyal et al., 2022). It is calculated as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(property_i^{actual} - property_i^{predicted}\right)^2}{\sum_{i=1}^{n}\left(property_i^{actual} - property_i^{mean}\right)^2} \qquad (4.4)$$

The R^2 values range between 0 and 1, where a value close to 1 indicates that the model effectively predicts the selected mechanical properties, capturing strong relationships between chemical composition, processing conditions, and mechanical properties. Conversely, a value close to 0 suggests that the model fails to capture any meaningful patterns in the data.

These evaluation metrics were used not only to assess model accuracy but also to benchmark the relative performance of the ANN and XGBoost architectures across the four target properties. This benchmarking supports the reliable model for capturing structure–property relationships within traditional steel classifications.

4.4.4 SHAP Analysis

To gain deeper insights into the relationships captured by the machine learning models, the SHAP framework (Nguyen et al., 2021) was applied to the best-performing models, as identified in the evaluation phase. As discussed in chapter 3, SHAP provides a systematic and model-agnostic method for interpreting predictions by quantifying the marginal contribution of each input feature, including chemical composition and processing conditions, to the predicted mechanical properties.

The SHAP analysis is performed through the following steps:

1. Compute SHAP values: The impact of each input feature (e.g., chemical composition, processing conditions) on the predicted mechanical properties is calculated.
2. Feature importance ranking: SHAP ranks the independent attributes based on their overall contribution to the model's predictions.
3. Visualisation techniques: Summary plot is used to explain the results.

SHAP was incorporated to verify that the model relied on physically meaningful relationships and to ensure transparency in the prediction process. This interpretability step supports model validation and informs the clustering and grade-rationalisation stages developed in the next chapter.

4.5 Summary

This chapter presented the first stage of the methodological framework, which defines how a unified dataset of carbon and stainless steel grades is constructed, pre-processed, and prepared for analysis. It outlined the approach for compiling data from multiple international sources, applying standardisation techniques such as feature scaling and encoding, and assessing

clustering suitability using the Hopkins statistic to assess the presence of natural groupings in the existing steel classification. It also described the proposed use of machine learning models to explore the influence of chemical composition and processing conditions on mechanical properties through data science lens. Furthermore, SHAP analysis was introduced as a method for interpreting model outputs and identifying influential features. These preparatory steps form the foundation for the proposed clustering and rationalisation strategy detailed in the next chapter, aimed at simplifying the steel grade system.

Chapter 5: Method II - K-Means Reduction Process (KMRP)

Following upon the data preparation and exploratory analysis presented in the previous chapter, this chapter introduces a multi-stage methodological framework for rationalising the existing steel grade system. The objective is to reduce the number of grades while maintaining mechanical performance and promoting recyclability. The proposed framework, called KMRP, consists of three sequential phases: (1) identifying the optimal number of clusters based on mechanical similarity within the existing steel grade dataset; (2) grouping grades with similar mechanical properties, thereby detecting those that share overlapping performance ranges despite differing in chemical composition; and (3) systematically reducing redundancy within each cluster through a filtering strategy grounded in metallurgical logic and sustainability criteria.Unlike standard K-Means or conventional data-reduction approaches, the KMRP framework combines clustering and systematic reduction into a single, performance-preserving workflow. K-Means is used in Phase II of the framework to group steel grades based on their mechanical similarity; however, the algorithm itself does not determine which grades should be retained and reduction phase. KMRP extends this by introducing a reduction phase that evaluates, for every grade within each cluster, whether its removal would compromise the original min–max mechanical property ranges or introduce gaps in the property distribution. These property-coverage safeguards are not part of K-Means or any existing cluster-based elimination methods. Only when a grade is fully covered by others in the same cluster does KMRP apply sustainability-oriented prioritisation criteria (tramp-element tolerance and CRM dependence) to decide which grade to retain. This integration of performance-space preservation with metallurgically motivated filtering distinguishes KMRP as a rationalisation framework rather than a standard clustering algorithm.To initiate the process, multiple clustering algorithms are evaluated to determine the most effective method for capturing structural similarities in the performance space of existing grades. Once the optimal clustering method is selected and appropriate clusters are formed, the reduction phase targets grades that share similar mechanical behaviour but differ in alloying or processing characteristics. Within each cluster, a hierarchical filtering procedure is applied to retain the minimal set of representative grades that offer (i) broader mechanical property coverage, (ii) higher tolerance to tramp elements, and (iii) reduced reliance on CRMs. The resulting reclassification system simplifies the current steel grade system, aligns with application-based requirements, and supports circularity goals through improved recyclability and material efficiency with the smallest number of grades.

5.1 Clustering Method Selection

As the first step in the proposed KMRP framework, this section evaluates multiple clustering algorithms to identify the most suitable method for grouping steel grades based on mechanical similarity. Since the aim is to detect performance-based groups that reflect overlapping property ranges across existing grades, careful selection of the clustering method is essential. Three widely used clustering algorithms (K-Means, HC, and DBSCAN) are considered due to their differing assumptions regarding cluster shape, density, and separability, as discussed in chapter 3. Their suitability is assessed using both quantitative clustering metrics and visual inspection of cluster structure to ensure alignment with the objectives of grade rationalisation.

5.1.1 Overview of Clustering Algorithms

K-Means clustering: As discussed in chapter 3, K-Means was applied as a partition-based method (Jin and Han, 2010), with the number of clusters determined using the Elbow method and validated with Silhouette scores. Its computational efficiency and scalability made it a strong candidate for identifying broad patterns across thousands of steel grades. However, its assumption of spherical cluster structures and the need to predefine k required careful validation to avoid artificial groupings.

Hierarchical clustering (HC): HC was applied using Ward's linkage, which prioritises minimising within-cluster variance during merging (Halkidi, 2009). Unlike K-Means, HC does not require a predefined k, and the dendrogram structure allows exploration of nested similarities (Tullis and Albert, 2013). This was particularly relevant for the steel dataset, where overlapping property ranges could create natural sub-hierarchies. Nevertheless, HC is computationally intensive in high-dimensional spaces, and its non-iterative assignment mechanism limits adaptability once merges are fixed.

DBSCAN: DBSCAN was also applied to test a density-based approach capable of detecting irregularly shaped clusters and isolating outliers (Ester et al., 1996). The parameters (neighbourhood radius) and MinPts (minimum points) were tuned using k-distance graphs to balance cluster compactness against excessive noise classification (Ma et al., 2023). While DBSCAN is effective for handling heterogeneity, it is sensitive to parameter settings, and the skewed distribution of mechanical properties posed a challenge in maintaining consistent cluster formation.

Following the explanation of the mechanisms of popular clustering algorithms, the next step involves assessing their effectiveness in classifying steel grades. This was achieved using quantitative clustering metrics including Silhouette score and Davies–Bouldin Index, which evaluate cluster cohesion and separation, alongside PCA-based visual inspection to assess cluster structure and identify any overlap.

5.1.2 Evaluating Clustering Performance

To determine which method performs best and is most effective in grouping steel grades, in this study three different approaches was evaluated using the following metrics:

Silhouette score: It measures how well a data point belongs to its assigned cluster by evaluating two factors: cohesion (how closely it is grouped with similar points) and separation (how far it is from points in other clusters) (Rousseeuw, 1987). A higher score (closer to 1) means the clusters are well-defined and distinct, while a lower score (closer to 0 or negative) suggests that the clusters overlap or are not clearly separated.

Davies-Bouldin Index: It evaluates the quality of clustering by measuring how similar each cluster is to the one most similar to it (Davies and Bouldin, 1979). It does this by comparing the distance between clusters (inter-cluster distance) to the spread of points within each cluster (intra-cluster distance). A higher DBI value suggests poor clustering, meaning the clusters are either too close to each other or not well-defined. In contrast, a lower DBI value indicates that the clusters are well-separated and compact, which generally reflects a more effective clustering solution.

In addition to these numerical metrics, cluster distribution analysis was conducted through dimensionality reduction and visualisation. This step allowed manual inspection of cluster boundaries, particularly useful for validating whether steel grades within each cluster showed meaningful differentiation in mechanical properties.

5.1.3 Justification for Selecting K-Means

Based on the combined performance evaluation described above, and the detailed results presented in section 7.1, K-Means was selected as the most effective clustering method for the following rationalisation process. Among the three algorithms assessed, it produced the highest Silhouette score (0.58) and the lowest DBI (0.65), indicating superior cohesion within clusters and clear separation between them.

This performance reflects the suitability of K-Means for the four-dimensional standardised mechanical property space (hardness, UTS, YS and elongation). Its iterative centroid-refinement mechanism allowed it to capture the dominant patterns in the data more effectively. In contrast, HC was limited by its lack of a refinement mechanism, since once data points are merged into clusters, assignments cannot be revised, reducing its adaptability to overlapping property ranges (Yu et al., 2023). DBSCAN, although advantageous for discovering clusters of arbitrary shapes, struggled to handle the skewed distribution of mechanical properties. Its sensitivity to the parameter and inability to differentiate clusters without treating many points as noise turned it unsuitable for this application.

While K-Means formally assumes spherical clusters with roughly comparable variance, these assumptions were evaluated against the structure of the data. After scaling, the mechanical property space showed strong linear relationships (particularly between UTS and YS) that produced coherent, approximately convex groupings, as confirmed by PCA visualisations and internal validation metrics. Under these conditions, the Euclidean distance criterion of K-Means aligned well with the underlying structure of the dataset, yielding stable, interpretable clusters that support the requirements of the reduction strategy. For these reasons, K-Means was adopted as the foundational clustering method within the proposed KMRP framework.

5.2 Mechanism of the K-Means Reduction Process (KMRP)

This section introduces a multi-step method for identifying and reducing the number of steel grades based on mechanical property similarity, minimal use of critical raw materials, and enhanced recyclability. The KMRP is designed to group grades that show overlapping performance and to select representative subsets that preserve the full range of mechanical properties while minimising critical raw material dependency and maximising recyclability. As shown in figure 7.1, the process contains three phases: (1) determining the optimal number of clusters, (2) applying K-Means clustering to cluster similar grades, and (3) reducing the number of grades within each cluster while ensuring that key property ranges are maintained. Unlike standard K-Means, which only produces clusters, KMRP adds a property-coverage check and sustainability-based elimination stage, making it a full rationalisation framework rather than a clustering method alone. This approach supports steel rationalisation by identifying and removing grades with redundant performance characteristics.

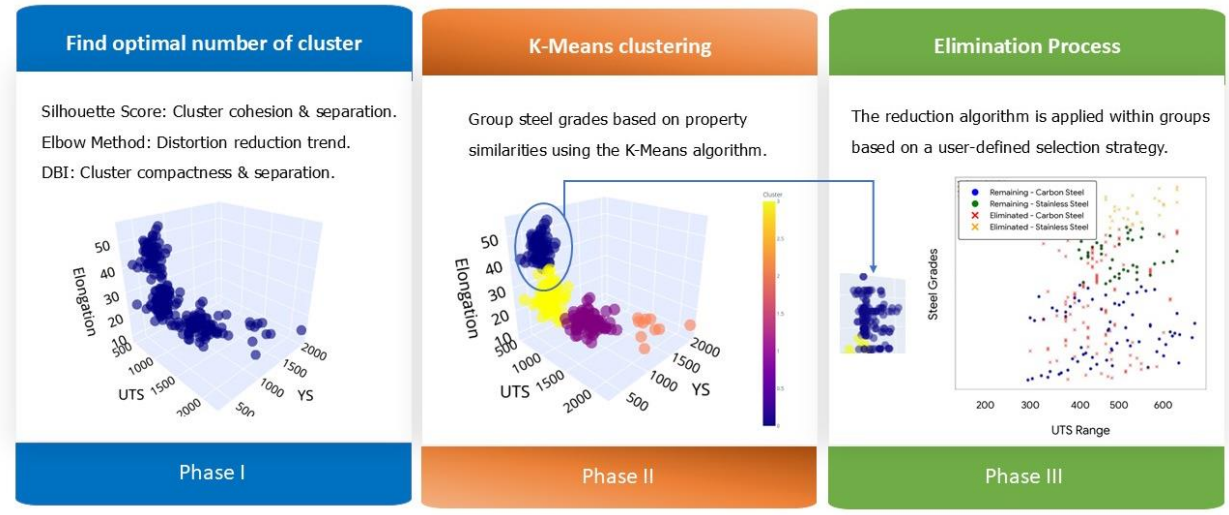


Figure 5.1: K-Means Reduction Process (KMRP) workflow.

To complement the workflow shown in Figure 5.1, algorithm 5.1 expresses the KMRP as an explicit sequence of operations. Presenting the method in pseudocode clarifies how the framework is executed in practice and highlights the decisions involved in retaining or eliminating grades. The following subsections describe each phase in detail.

Algorithm 5.1 K-Means Reduction Process (KMRP)

Input: X (mechanical properties), T (tramp elements), C (CRMs), $strategy \in \{minimise\_tramp, maximise\_tramp\}$
Output: Reduced grade sets A'_j for each cluster C_j

Phase I: Selecting the optimal number of clusters
for k in $K_{candidates}$ **do**
    Compute WCSS(k), Silhouette(k), and DBI(k)
end for
Select the optimal k^* using Elbow and Silhouette results

Phase II: K-Means clustering
Run K-Means on X with k^* to obtain clusters $C_1, \ldots, C_{k^*}$
for each cluster C_j **do**
    Define $A_j = \{a_1, \ldots, a_{n_j}\}$
end for

Phase III: Reduction process
for each cluster C_j **do**
    Determine property bounds for C_j
    Initialise $A'_j \leftarrow A_j$
    for each grade a_i in A_j **do**
        $A_{temp} = A'_j \setminus \{a_i\}$
                ▷ Coverage check (Step 3)
        if removing a_i changes the min/max property bounds **then**
            continue
        end if
                ▷ Gap detection (Step 4)
        if removing a_i introduces a gap in any property dimension **then**
            continue
        end if
                ▷ Identify overlapping grades (Step 6)
        Identify $G_{overlap} \subseteq A_{temp}$ that fully cover the property range of a_i
        if $G_{overlap} = \emptyset$ **then**
            continue
        end if
                ▷ Compute prioritisation scores (Step 6)
        Compute $S_T(a_i)$, $S_C(a_i)$
        Compute $S_T(G_{overlap})$, $S_C(G_{overlap})$
                ▷ Apply user-defined prioritisation strategy (Step 6)
        if $strategy = minimise\_tramp$ **then**
            if $S_T(a_i) > S_T(G_{overlap})$ **then**
                Remove a_i from A'_j
            else if $S_T(a_i) = S_T(G_{overlap})$ **and** $S_C(a_i) > S_C(G_{overlap})$ **then**
                Remove a_i from A'_j
            end if
        else
            if $S_T(a_i) < S_T(G_{overlap})$ **then**
                Remove a_i from A'_j
            else if $S_T(a_i) = S_T(G_{overlap})$ **and** $S_C(a_i) < S_C(G_{overlap})$ **then**
                Remove a_i from A'_j
            end if
        end if
    end for
end for

Algorithm 5.1: Pseudocode for the KMRP.

5.2.1 Phase I: Selecting the Optimal Number of Clusters

As established in the previous section, K-Means was selected as the most suitable clustering algorithm for this application due to its ability to capture the underlying structure of mechanical property space. Therefore, determining the most appropriate number of clusters in a K-Means method is a very important task to ensuring that steel grades are grouped meaningfully based on their mechanical properties. As discussed in chapter 3, an optimal cluster (k) enables effective pattern recognition by avoiding underfitting (too few clusters, missing variation) and overfitting (too many clusters, capturing noise) (Xie et al., 2022). To achieve this, in this study, two popular methods including the Elbow method and the Silhouette score has been used, and their results compared to determine the optimal number of clusters in the dataset.

5.2.1.1 Elbow Method

The Elbow Method (López-Rubio et al., 2018) was used to estimate the number of clusters K by analysing the WCSS across varying K values. WCSS quantifies how tightly points are grouped around cluster centroids. It is computed as:

$$WCSS = \sum_{j=1}^{k} \sum_{x_i \in C_j} \left\| x_i - \mu_j \right\|^2 \qquad (5.2)$$

In equation 5.2, x_i is the i-th steel grade and μ_j is the centroid of cluster C_j. The squared Euclidean distance measures how well each steel grade fits within its assigned cluster.

To assist in interpreting the results, Yellowbrick visualisation package (Bengfort et al., 2018) was used, which includes distortion score and fit time, to visualise and evaluate the clustering process. The distortion score measures cluster compactness, with a lower score indicating well-defined and compact clusters while fit time reflects the computational efficiency of the clustering algorithm, with shorter times being desirable for larger datasets or real-time applications.

5.2.1.2 Silhouettes Score

In order to validate the result of elbow method, the silhouette score (Rousseeuw, 1987) was applied based on the different K values. This metric considers both intra-cluster cohesion and inter-cluster separation, offering a more robust indicator of how well each steel grade fits its assigned group. Silhouette score for a data point is calculated as:

$$Silhouettes\ Score = (b - a)/max(a, b) \quad (5.3)$$

In equation 5.3, a is the average distance between a steel grade and other similar grades in the same cluster, and b is the average distance to the nearest neighbouring cluster. Scores range from -1 to 1; higher values indicate well-separated, coherent clusters. Using both WCSS and silhouette scores ensures that the final K not only reduces error but also reflects interpretable groupings in property space.

5.2.1.3 PCA

To enhance both the performance and interpretability of the K-Means clustering, PCA was applied as a dimensionality reduction technique (Ding and He, 2004). Since clustering based on mechanical properties can suffer from redundancy and noise due to correlated features, PCA transforms these variables into a smaller set of uncorrelated components while retaining most of the dataset's variance (Xue et al., 2011). This technique not only improves clustering stability but also aids in visualising the distribution of steel grades in a reduced feature space. The transformed dataset was then used in phase II to enable property-driven grouping of grades that reflect mechanical similarity and potential overlapping grades with the same performance but differ in chemical composition.

5.2.2 Phase II: K-Means Clustering

In the second phase of the KMRP framework, steel grades are grouped based on four key mechanical properties, including Brinell hardness, UTS, YS, and elongation. The objective is to identify grades that demonstrate similar performance profiles, regardless of their chemical compositions. By clustering only on properties, the methodology enabled the detection of grades with equivalent performance, which could subsequently be compared and replaced with more sustainable alternatives that offer equivalent performance.

The dataset is defined as $X = \{x_1, x_2, \dots, x_n\}$, where each steel grade x_i is represented as a vector of m mechanical properties. K-Means partitions the dataset into k clusters $C = \{C_1, C_2, \dots, C_k\}$, each associated with a centroid μ_j. The clustering objective follows the WCSS formulation in equation 5.1, which aims to minimise the squared Euclidean distance between each steel grade and the centroid of its assigned cluster.

Rather than re-deriving the mathematical formulation, this section focuses on the operational procedure. K-Means clustering was implemented using the "sklearn" library in Python (Pedregosa et al., 2011) to group steel grades with similar mechanical properties. The algorithm proceeds as follows (Likas et al., 2003):

1. Initialisation: The number of clusters K is set based on the results of the Elbow method and Silhouette score.
2. Centroid selection: K initial centroids are chosen randomly from the dataset.
3. Assignment: Each data point is assigned to the cluster with the nearest centroid, using Euclidean distance as the metric.
4. Centroid update: New centroids are computed by taking the mean of all points assigned to each cluster.
5. Iteration: Steps 3 and 4 are repeated until the cluster assignments converge and the within-cluster variance is minimised.

Upon completion of the clustering process, each steel grade was assigned to a distinct group of mechanically similar grades. These clusters form the foundation for the subsequent phase of the KMRP framework, which involves a targeted reduction strategy. This strategy aims to minimise the number of steel grades within each cluster while preserving the full range of mechanical properties originally represented. The reduction process also incorporates sustainability-driven priorities, including tolerance to tramp elements and minimisation of dependency on critical raw materials.

5.2.3 Phase III: Reduction Process

The reduction process was applied in this phase once the clusters C_j were established in phase II. The objective of this phase is to compare steel grades that contains overlapping mechanical properties within each cluster and reduce the number of grades, while ensuring that the remaining set still covers the full range of original properties. This ensures that only the minimum number of grades is retained without introducing gaps in the defined ranges of UTS and elongation, which are used as indicators of strength and ductility.

To enable flexible decision-making aligned with sustainability goals, the methodology includes a user-defined prioritisation mechanism. This allows the reduction process to favour grades with avoidance/higher-tolerance to tramp elements and lower reliance on CRMs, depending on the selected strategy.

The procedure is iterative and consists of eight sequential steps, each designed to preserve mechanical diversity while reducing redundancy. These steps are detailed below and visually summarised in Figure 5.2, which presents a flowchart of the proposed elimination framework.

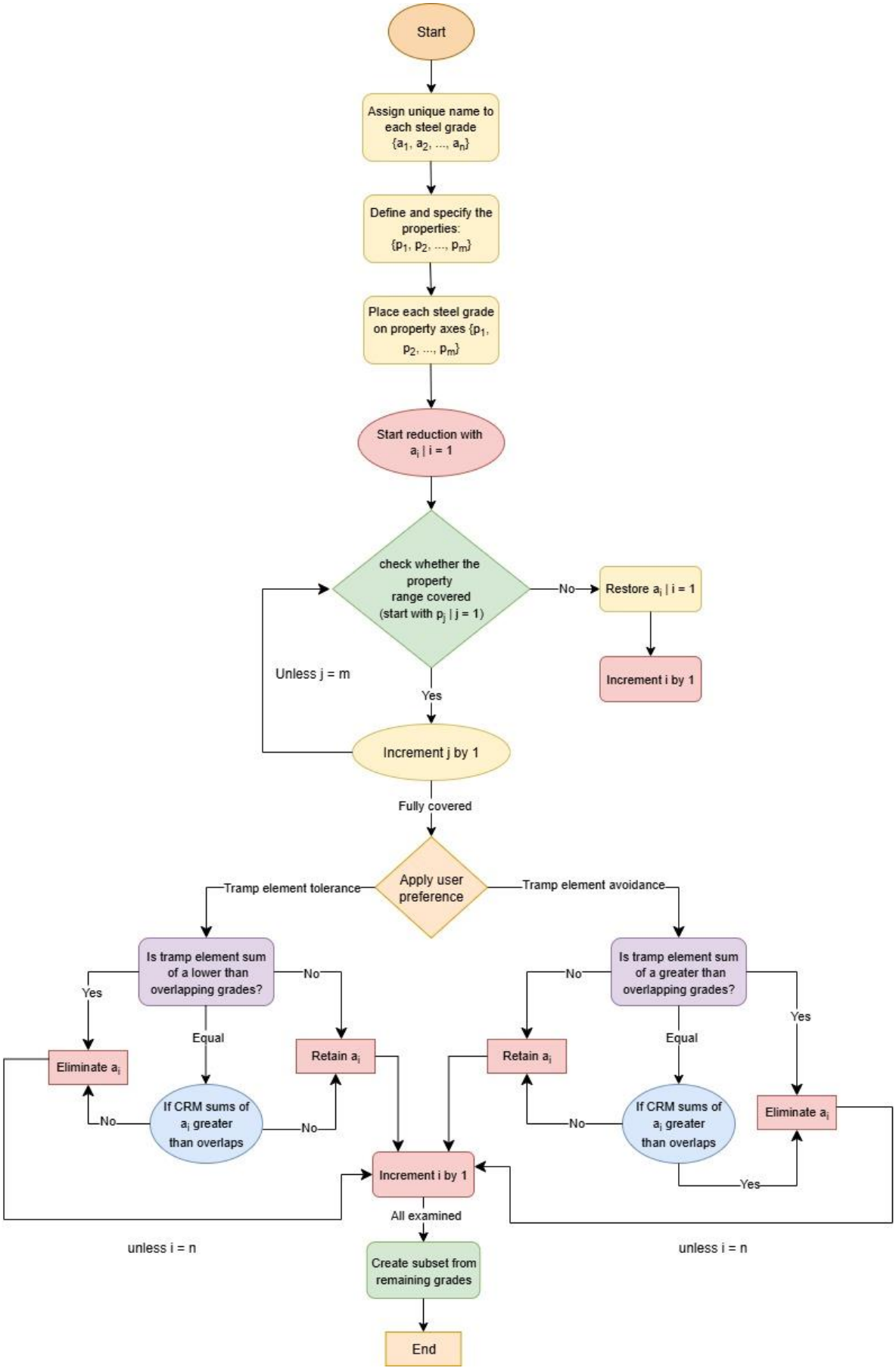

Start
Assign unique name to each steel grade {a1, a2, ..., an}
Define and specify the properties: {p1, p2, ..., pm}
Place each steel grade on property axes {p1, p2, ..., pm}
Start reduction with ai | i = 1
check whether the property range covered (start with pj | j = 1)
No
Restore ai | i = 1
Increment i by 1
Unless j = m
Yes
Increment j by 1
Fully covered
Apply user preference
Tramp element tolerance
Tramp element avoidance
Is tramp element sum of a lower than overlapping grades?
Yes
No
Equal
Eliminate ai
Retain ai
If CRM sums of ai greater than overlaps
No
No
Is tramp element sum of a greater than overlapping grades?
No
Yes
Equal
Retain ai
Eliminate ai
If CRM sums of ai greater than overlaps
No
Yes
Increment i by 1
All examined
unless i = n
unless i = n
Create subset from remaining grades
End

Figure 5.2: Flowchart of the proposed elimination process to reduce the number of steel grades while maintaining the original property space coverage.

Step 1: Define steel grades and properties

Let $A_j = \{a_1, a_2, \ldots, a_{nj}\}$ be the set of grades in cluster C_j, where each grade a_i is associated with a vector of mechanical properties, denoted as $\mathbf{P}_i = (p_{i1}, p_{i2}, \ldots, p_{im})$.

Step 2: Mapping steel grades to property space

In this step, each grade a_i is mapped on the m property space, where $\mathbf{P}_j$ is properties matrix inside each generated cluster:

$$\mathbf{P}_j = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ p_{n_j 1} & p_{n_j 2} & \cdots & p_{n_j m} \end{bmatrix} \tag{5.4}$$

Step 3: Coverage check

In this step, mechanical property coverage was defined quantitatively to ensure that the removal of a grade does not reduce the performance bounds originally present within the cluster. The property coverage is calculated by finding the minimum and maximum values in all grades within the clusters C_j, for each p_k property. A reduced set $A'_j \subset A_j$ is valid only if it preserves both boundaries:

$$min_{pk}(C_j) = \min_{a_i \in C_j} p_{ik} \ , max_{pk}(C_j) = \max_{a_i \in C_j} p_{ik} \tag{5.5}$$

This ensures that eliminating a grade never reduces the original performance envelope of the cluster.

Step 4: Range overlap (gap detection)

In addition to maintaining global coverage, this step verifies that no gaps are introduced in the property space. A gap refers to a range within the mechanical properties where no steel grade provides coverage.

First, $p_k^{sorted} = \left\{p_{1k}^{sorted}, p_{2k}^{sorted}, \ldots, p_{njk}^{sorted}\right\}$ represented the sort values of k property based on all grades in cluster C_j. Then, in steel grades a_i and a_{i+1}, the gap is provided on property k if:

$$p_{(i+1)k}^{sorted} - p_{ik}^{sorted} > \varepsilon_k \quad (5.6)$$

And the gap is not provided on property axis where:

$$p_{(i+1)k}^{sorted} - p_{ik}^{sorted} \leq \varepsilon_k \quad (5.7)$$

where ε_k denotes an allowable tolerance. If a gap is detected, the excluded grade must be reintroduced to preserve coverage.

In this study, a strict threshold of $\varepsilon_k = 0$ was adopted. This ensures that no interior regions of the property range become unrepresented. Any positive gap indicates that the removed grade provided unique intermediate behaviour, in which case the grade must be reintroduced.

The zero-tolerance criterion is justified because the objective of KMRP is to preserve the full range of mechanical performance within each cluster. Allowing a non-zero ε_k would risk discarding grades whose properties are not captured elsewhere, thereby compromising performance.

Step 5: Initiate the reduction process

For each grade $a_i \in A_j$, the system evaluates whether its removal maintains both property coverage and continuity. Two checks are performed:

The reduction process begins for each grade a_i. The system checks whether the remaining grades fully cover the property range of a_i. If removing a_i introduces a new gap, it is retained.

- Coverage check: Confirm that removing a_i does not shrink the cluster's overall mechanical property range.
- Gap check: ensure that removing a_i does not introduce a gap (step 4).

If both conditions are satisfied, the algorithm proceeds to apply prioritisation criteria to determine whether a_i should be eliminated. If either condition fails, the grade is retained and the algorithm skips to Step 7.

Step 6: Prioritisation criteria for reducing the number of options

Let a_i be the selected grade and $G_{overlap}$ be the set of grades that fully cover the same properties range as a_i. The reduction process is based on two user-defined criteria:

1. First criterion: sum of tramp elements $S_T(a_i)$

2. Second criterion: sum of critical raw materials $S_C(a_i)$

where:

$$\boldsymbol{S_T(a_i) = \sum_{k \in T} p_{ik}} \ , \boldsymbol{S_T(G_{overlap}) = \sum_{g \in G_{overlap}} S_T(g)} \tag{5.8}$$

$$\boldsymbol{S_C(a_i) = \sum_{k \in C} p_{ik}} \ , \boldsymbol{S_C(G_{overlap}) = \sum_{g \in G_{overlap}} S_C(g)} \tag{5.9}$$

In equation 7.8 and 7.9, T and C are the sets of tramp elements and critical raw materials, respectively.

Minimisation strategy

If the user chooses to minimise tramp elements, the reduction follows these conditions:

Primary check – Tramp elements:

- If the selected grade's tramp sum is higher than the sum of overlaps ($\boldsymbol{S_T(a_i) > S_T(G_{overlap})}$), then eliminate a_i and proceed to the step 7.
- If the selected grade's sum of tramp elements is lower than the sum of overlapping grades ($\boldsymbol{S_T(a_i) < S_T(G_{overlap})}$), then retain a_i and proceed to the step 7.
- If the selected grade's tramp sum is equal to overlapping grades ($\boldsymbol{S_T(a_i) = S_T(G_{overlap})}$), proceed to the secondary check.

Secondary check – Critical raw materials:

- If the selected grade's critical material sum is higher than the sum of overlapping grades ($\boldsymbol{S_C(a_i) > S_C(G_{overlap})}$), then eliminate a_i and proceed to the step 7.
- If the selected grade's critical sum is lower or equal ($\boldsymbol{S_C(a_i) \leq S_C(G_{overlap})}$), then retain a_i and proceed to the step 7.

Maximisation strategy

If the user chooses to maximise the tramp elements, the elimination follows these conditions:

Primary check – Tramp elements:

- If the selected grade's tramp sum is lower than the sum of overlaps ($\boldsymbol{S_T(a_i) < S_T(G_{overlap})}$), then eliminate a_i and proceed to the step 7.
- If the selected grade's sum of tramp elements is higher than the sum of overlapping grades ($\boldsymbol{S_T(a_i) > S_T(G_{overlap})}$), the grade is retain and proceed to the step 7.
- If the selected grade's tramp sum is equal to overlapping grades ($\boldsymbol{S_T(a_i) = S_T(G_{overlap})}$), proceed to the secondary check.

Secondary check – Critical raw materials:

- If the selected grade's critical material sum is higher than the sum of overlapping grades ($\boldsymbol{S_C(a_i) > S_C(G_{overlap})}$), then eliminate a_i and proceed to the step 7.
- If the selected grade's critical sum is lower or equal ($\boldsymbol{S_C(a_i) \leq S_C(G_{overlap})}$), then retain a_i and proceed to the step 7.

Taken together, Steps 3 to 6 define the quantitative criteria for constructing the minimal grade set within each cluster. A grade a_i is eligible for elimination only if three conditions are simultaneously satisfied: (i) its removal does not change the cluster-level bounds for any mechanical property, i.e. the minimum and maximum values in equation (5.5) remain identical when computed over $A_j' = A_j \setminus \{a_i\}$; (ii) it does not create a new internal gap in the ordered property space, such that the continuity constraint in equations (5.6) and (5.7) with $\varepsilon_k = 0$ remains valid for all properties p_k; and (iii) there exists at least one grade in G_{overlap} that fully covers the same property range and is preferred according to the tramp-element and CRM scores S_T and S_C defined in equations (5.8) and (5.9) under the selected strategy. The algorithm iteratively removes only those grades that meet all three criteria; all remaining grades are therefore necessary either for preserving mechanical coverage and continuity or for satisfying the sustainability-based prioritisation, which ensures that A_j' is the smallest subset for cluster C_j.

Step 7: Iterate over all grades within each generated group

The reduction process repeats for all grades a_i within each generated cluster C_j.

Step 8: Create the final subset of remaining grades

After all iterations, the system outputs a final reduced set $A_j' \subset A_j$ that satisfies the following conditions:

1. Fully covers the original property range of cluster C_j.
2. Introduces no gaps in the property space.
3. Aligns with the user-defined prioritisation criteria.
4. Contains the minimum number of steel grades.

$$A_j' = \left\{a_1', a_2', \ldots, a_{r_j}'\right\} \tag{5.10}$$

This final set ensures full mechanical property coverage while using the smallest possible number of steel grades and supporting the desired sustainability objectives.

5.3 Interpretation and Feature Analysis of KMRP Classification

This section describes the methods applied to evaluate the quality, interpretability, and physical justification of the steel grade groupings generated through the KMRP framework. As unsupervised learning algorithms may produce clusters that are statistically valid but not necessarily meaningful from a metallurgical or sustainability perspective, further analysis is required to assess the internal consistency, domain relevance, and attribute influence for each cluster. The interpretability assessment consists of two components: (1) histogram-based distributional analysis of key mechanical and chemical features to examine intra and inter-cluster variance, and (2) SHAP analysis to quantify the contribution of each input feature to the clustering outcome and to explore the potential assignment of newly collected steel grades to the generated groups. Together, these methods support both the validation and interpretability of the KMRP-based classification.

5.3.1 Histogram Analysis for Evaluating Group Characteristics

To evaluate the clustering results, first the range of different chemical compositions and mechanical properties within each generated group analysed. This helps to understand the characteristics of each cluster and identify overlaps and gaps between different steel grades.

To assess the internal consistency and separability of clusters, histograms were generated for two key mechanical properties (UTS and elongation) as representative indicators of strength

and ductility across all clusters. This approach allowed to evaluate how well the steel grades are grouped, detect potential overlaps between clusters.

The histogram analysis enabled a clearer understanding of the characteristic boundaries of each group generated by the KMRP framework, revealing how mechanical and chemical attributes are distributed within and across clusters.

5.3.2 SHAP Analysis

While K-Means is an unsupervised algorithm, SHAP analysis was applied through a supervised learning model to enable interpretability of the KMRP-generated cluster assignments. A RF classifier was selected and trained to predict these cluster labels using the full input feature set, which included chemical composition, processing indicators, and mechanical properties. The choice of RF was motivated by its robustness to high-dimensional (Chi et al., 2022), heterogeneous data and its ability to handle non-linear interactions between features (Breiman, 2001b), both of which were characteristic of the examined alloy datasets. Furthermore, RF are inherently compatible with SHAP analysis, providing reliable feature attribution through tree-based decomposition methods (Bifarin, 2023). The integration of RF with SHAP enabled the identification of the most influential features contributing to cluster membership and facilitated a critical evaluation of whether the resulting groupings were chemically and metallurgically meaningful. In addition, the trained classifier can be leveraged to assign newly collected steel grades to appropriate KMRP-defined groups, thereby supporting the future deployment of the framework in alloy design and database integration.

The SHAP workflow consisted of the following steps:

1. Training a classifier: A RF model was trained to predict KMRP-assigned cluster labels from the input feature space. Model performance was validated using standard metrics (accuracy, F1-score) to ensure reliable cluster assignment prediction.
2. Computing SHAP values: SHAP values were calculated for each input feature, quantifying its contribution to the prediction of cluster membership on a per-sample basis.
3. Ranking and interpretation: Global feature importance was derived by averaging SHAP magnitudes across all samples. Summary plots were used to visualise the direction and magnitude of each feature's influence on cluster assignment.

Together, the histogram and SHAP analyses serve as interpretability tools to evaluate the KMRP classification. These assessments verify that the clusters not only preserve mechanical diversity but also correspond to meaningful chemical or processing differences.

5.4 Summary

This chapter introduced the KMRP, an application-driven methodology aimed at simplifying the steel grade system by reducing redundancy while maintaining performance and promoting recyclability. The framework consisted of three sequential phases including identifying optimal clusters of mechanically similar grades, applying K-Means clustering to group grades regardless of composition, and proposing an iterative reduction strategy within each group to retain only those grades that preserve full mechanical property coverage. The reduction process included user-defined priorities, such as prioritising grades with higher tolerance to tramp elements and lower dependency on critical raw materials, ensuring alignment with sustainability and circularity objectives. Finally, the chapter outlined an interpretability strategy based on histogram analysis and SHAP values, enabling future validation of the proposed groupings and facilitating the integration of the framework into a simplified steel grade system.

Following the methodological framework for the reclassification of steel grades using the KMRP approach to achieve a reduced set of grades, the following chapter presents the results of these workflows. The results first detail the data analysis and preparation steps undertaken to enable the proposed methodologies.

Chapter 6: Data-driven Analysis of Carbon and Stainless Steel Grades

This chapter presents the results of applying the data preparation and modelling framework described in chapter 4 to carbon and stainless steel grades. The focus is on examining whether the datasets are structurally suitable for clustering, assessing predictive performance of ANN, RF, and XGBoost models, and interpreting the influence of chemical and processing variables on mechanical properties using SHAP. Together, these analyses provide the empirical grounding needed to understand the current classification of steel grades and to inform the KMRP framework developed in chapter 7.

6.1 Carbon Steel: Data Preparation and Modelling

6.1.1 Data Preparation and Clustering Suitability

The dataset on carbon steel contained 237 rows and 43 attributes under four different standards included AISI, ASTM, GB, and EN. It included 117 alloy grades, featuring both dependent variables including mechanical properties, and independent variables, the chemical composition and processing conditions of the steel grades. As described in chapter 4, although carbon steel represents the largest family of steels globally, only a subset of grades could be included in this study because many international designations lack complete mechanical, chemical, or processing information. Therefore, only grades with fully specified and reliable data were retained. In the examined dataset, AISI was the most represented standard with 122 entries, followed by EN with 72 entries, ASTM with 26, and GB being the least represented with only 16.

Figure 6.1 presents the distributions of UTS and YS. As shown in Figure 6.1(a), most grades had UTS values between 350 and 500 MPa, with the distribution skewed to the right, indicating a small number of high-strength steels exceeding 1000 MPa. A similar trend was observed for YS (Figure 6.1b), where the majority of grades ranged from 250 to 400 MPa, with fewer cases above 750 MPa. These findings indicated that most carbon steel grades contained moderate strength characteristics, with only a few high-strength outliers.

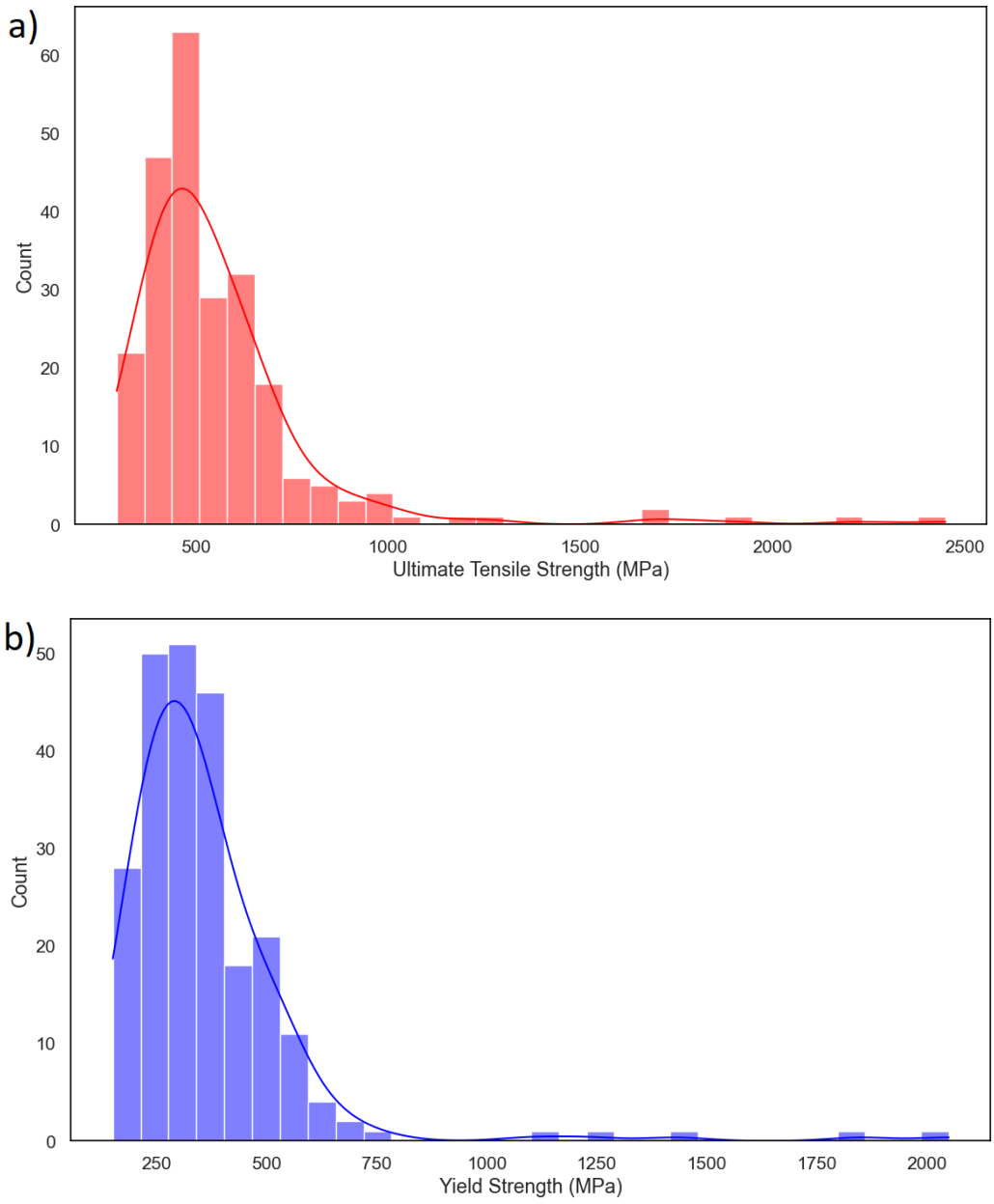


Figure 6.1: Distribution of (a) ultimate tensile strength (UTS) and (b) yield strength (YS) in the examined carbon steel grades.

Further examination of property-property relationships is shown in Figure 6.2, which displays a strong linear correlation between UTS and hardness, consistent with established metallurgical principles. This finding suggested that UTS could be used as an indicator for hardness when classifying steels using clustering.

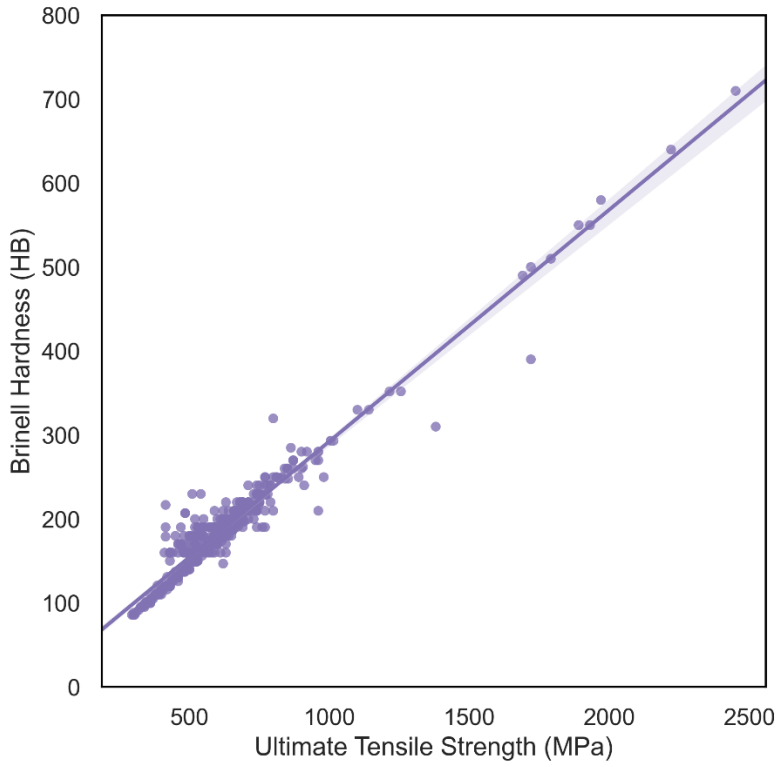


Figure 6.2: Correlation between ultimate tensile strength (UTS) and hardness in the examined carbon steel grades.

Figure 6.3 shows the variation in carbon and manganese content that influence the strength and toughness of carbon steels. Most steels contained between 0.12% and 0.45% carbon, with a few outliers exceeding 0.75%, typical of high-carbon steels (Islam and Rashed, 2019). Although these values appear as outliers in the boxplot, they are not statistical anomalies but genuine high-carbon grades whose mechanical properties fall within the overall performance envelope. Therefore, they are retained in the dataset and evaluated in the KMRP framework based on property coverage rather than chemical extremity. Manganese content ranged from 0.45% to 1.2%, with certain high-Mn grades associated with improved toughness (Zhan et al., 2024). These compositional variations, along with the observed property patterns, support the rationale for property-driven grouping and suggest the presence of redundant grades offering similar mechanical performance despite slight differences in chemical composition.

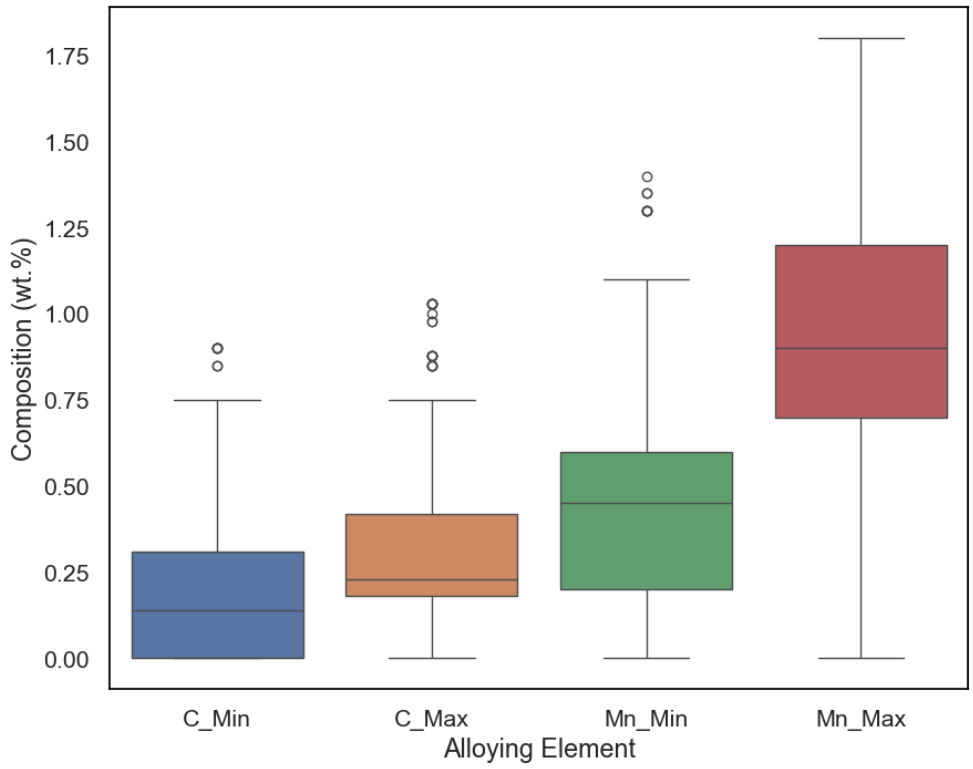


Figure 6.3: Range of carbon and manganese content in the examined carbon steel grades. Outliers indicate grades with exceptionally high levels of either element compared to the majority.

6.1.1.1 Standard Scaling

To ensure fair comparison during clustering and machine learning tasks, standard scaling was applied to all numerical features in the carbon steel dataset. This technique standardises variables by removing the mean and scaling to unit variance, a widely accepted practice for preparing heterogeneous numerical data in materials informatics (Zivic et al., 2025). In the context of steel datasets, this is particularly important because alloying elements such as nickel and chromium can appear in double-digit weight percentages, while others like carbon often remain below 1 wt%. In addition, mechanical properties such as UTS (MPa) and elongation (%) include different numeric ranges and units.

Standard scaling was performed using the "StandardScaler" class from the scikit-learn Python library. In this transformation, each numerical value is adjusted by subtracting the feature's mean and dividing by its standard deviation, allowing the scaled value to reflect how far it lies from the average relative to the natural variability of that feature. As a result, values above the mean become positive, values below the mean become negative, and all numerical features are placed on a comparable scale. This transformation was applied to all chemical composition features (C_Min, C_Max, and Mn_Min) and all the mechanical properties, including UTS and elongation. In contrast, processing condition features (heat treatment, forming method) were

stored as binary-encoded columns (1 for 'applied', 0 for 'not applied'), therefore, did not required scaling. As described in chapter 4, although different processing routes vary in parameters such as temperature, and cooling rate, these quantitative details were not available in the source data. Processing information was provided only in descriptive form, meaning that the presence or absence of a treatment was the only representable feature. Because such binary indicators do not carry magnitude information and scaling would not introduce meaningful structure, they were retained in their original form.

Table 6.1 presents a sample of five steel grades before and after scaling, highlighting how standardisation brought heterogeneous feature magnitudes into a comparable numeric range. Standard Scaling is especially important for distance-based models such as K-Means and DBSCAN, where unscaled features can dominate Euclidean distance and skew clustering results.

Table 6.1: Comparison of selected features before and after standard scaling in the examined carbon steel grades.

C_Min (wt %)	C_Max (wt %)	Mn_Min (wt %)	UTS (MPa)	Elongation (%)	C_Min_scaled	C_Max_scaled	Mn_Min_scaled	UTS_scaled	Elongation_scaled
0.12	0.18	0.6	380	25	-0.14297	-0.54528	0.298774	-0.76952	0.115007
0.0001	0.2	1	630	22	-0.92074	-0.38949	1.494109	0.995429	-0.26835
0.42	0.5	0.6	690	12	1.803087	1.947428	0.298774	1.419016	-1.54621
0.0001	0.14	0.0001	350	25	-0.92074	-0.85687	-1.49393	-0.98131	0.115007
0.17	0.23	0.3	395	36.5	0.181373	-0.15579	-0.59773	-0.66362	1.584543

6.1.1.2 Hopkins statistics

To evaluate the suitability of unsupervised clustering algorithms for the carbon steel dataset, the Hopkins statistic was calculated. This statistical test quantifies the clustering tendency of a dataset by comparing the distribution of real data points to randomly generated points. A value close to 1 indicates a strong clustering structure, whereas a value around 0.5 suggests randomness.

The test was applied to the complete set of standardised numerical features, including both chemical composition (C, Mn, Cr, Ni, Mo) and mechanical properties (Hardness, UTS, YS, elongation), using all 237 carbon steel grades in the dataset. In each evaluation, a random subset of points were selected from the dataset, and the statistic was computed by comparing their nearest-neighbour distances to those of randomly generated points in the same feature space. The calculation was repeated five times using different random seeds, producing values in the range H = 0.84-0.87, demonstrating that the result is stable and not sensitive to the random sample chosen. The representative value reported for this study is H = 0.86, corresponding to the mean of these repeated evaluations, which was close to 1. The result indicated that carbon steels contain non-random structure and natural clusters. The high Hopkins value provided quantitative justification for the application of unsupervised learning algorithms (K-Means, DBSCAN, or hierarchical clustering) in subsequent stages of the analysis to uncover structural patterns that can meaningfully inform the rationalisation of steel grades.

6.1.2 Relationships Between Composition, Processing, and Properties in Carbon Steels

This section presents a comparative analysis of three widely used supervised learning algorithms, including ANN, XGBoost, and Random Forest, to investigate the relationship between chemical composition, processing conditions, and mechanical properties in carbon steels. The input features included scaled values of chemical compositions alongside one-hot encoded indicators of processing conditions. The output variables targeted for prediction were four key mechanical properties, including Hardness (Brinell), UTS, YS, and elongation. The aim of this analysis is to (i) identify the most influential factors affecting mechanical performance which address the application requirement of steel grades and also (ii) evaluate the predictive capability and robustness of each model.

6.1.2.1 Model Evaluation and Comparison

The ANN, XGBoost, and RF algorithms were applied using a five-fold cross-validation approach, where the model was trained on four folds and tested on the remaining one. The MAE scores for both the training and validation phases of each algorithm are presented in Figure 6.4.

As shown in Figure 6.4, the average training MAE was lowest for the ANN model (27.52), followed by XGBoost (30.4) and RF (31.8), suggesting that ANN learned more effectively from the training data. In terms of validation performance, ANN also achieved the lowest average MAE (34.6), followed by XGBoost (38.7) and RF (40.5), indicating better

generalisation. The consistent gap of around 10 units between training and validation MAEs was expected due to the limited size of the dataset.

Although the MAE values in Figure 6.4 represent aggregated errors across all four predicted properties, their magnitude remains physically meaningful. Mechanical test results for carbon steels show variation due to differences in heat treatment, composition; for instance, strength values for nominally similar steels can differ by 20–50 MPa, while elongation may vary by several percentage points. An overall validation MAE in the range of 35–40 therefore falls within the natural scatter observed in industrial data, indicating that the models are operating within realistic engineering uncertainty rather than producing unphysical deviations.

Despite several efforts to improve model performance through hyperparameter tuning, dropout adjustments, and early stopping the prediction errors could not be reduced further, especially for fold 4, which consistently showed the highest validation error across all models. Further investigation showed that this fold included steel grades with unusual chemical compositions and treatment processes. Specifically, elements like titanium, vanadium, molybdenum, and chromium were present which are not typically found in most carbon steel grades. These samples stood out as clear outliers when tested using unsupervised anomaly detection and showed little similarity in composition compared to the rest of the dataset.

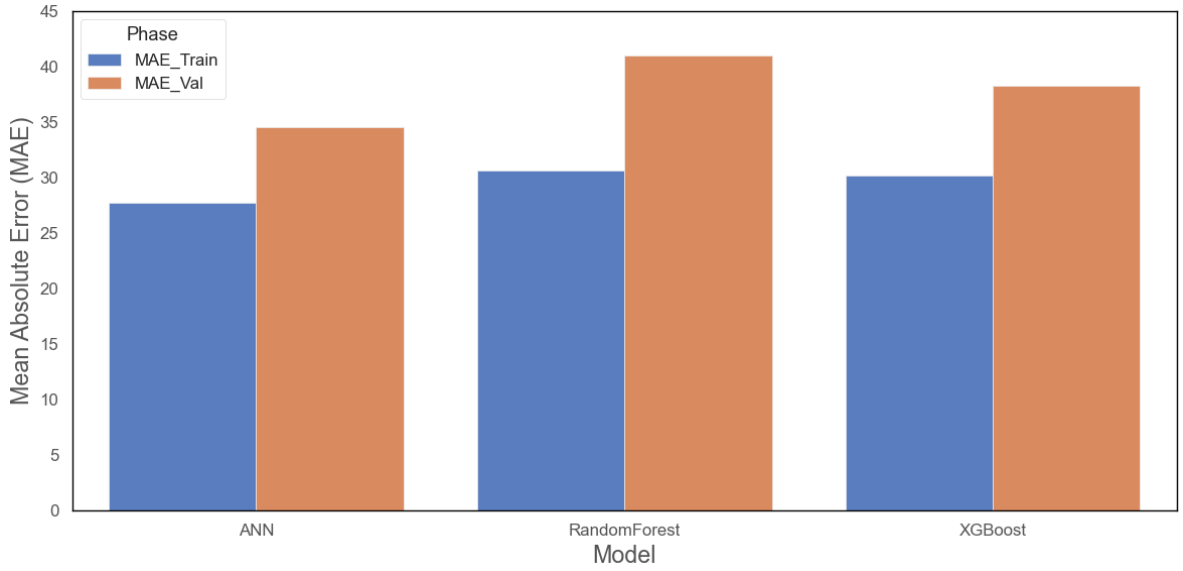


Figure 6.4: Average training and validation MAE across five folds for ANN, XGBoost, and Random Forest models. In five-fold cross-validation, the dataset is divided into five equal subsets; each model is trained on four subsets and evaluated on the remaining one, and this process is repeated five times. Therefore, every subset is used once for validation.

To further assess the predictive reliability of each model and capture the relationships, a per-property evaluation was conducted using four key selected mechanical targets. The results are

presented in Table 6.2, where the performance of each algorithm is summarised in terms of MAE, the actual mean value of each property in the dataset, and the Standardised MAE (S-MAE), which normalises the error by dividing it by the mean of the target variable. For instance, UTS and Hardness have much higher values than elongation, so comparing their MAEs directly would not make sense. The standardised MAE solves this by showing the prediction error as a proportion of the typical value of each property. In other words, it gives us a relative error, like a percentage, that is easier to compare across different targets.

The ANN model achieved the lowest MAE for hardness (19.77), UTS (64.69), and elongation (2.90). Its corresponding S-MAE values were also better in those three properties: 0.12 (Hardness), 0.11 (UTS), and 0.13 (elongation). For YS, ANN had an S-MAE of 0.14, which, although slightly higher than its score for UTS, still outperformed both XGBoost (0.16) and RF (0.17).

Compared to ANN, XGBoost demonstrated good predictive performance, with a similar S-MAE for hardness (0.12) and elongation (0.13), but showed slightly higher error for UTS (0.12) and YS (0.16). RF, while still performing acceptably, contained the highest standardised MAE in all categories, especially for YS (0.17), indicating relatively weaker predictive accuracy and generalisation. These results support the earlier findings presented in Figure 6.4, where ANN also had the lowest average MAE across training and validation folds. While each model captured the general relationships between features and targets, the ANN consistently delivered more accurate and generalisable predictions across both overall performance metrics and individual mechanical properties. The better performance of ANN may be attributed to its ability to model complex and also nonlinear interactions between chemical and processing features, relationships that tree-based methods may fail to capture fully, given the small dataset size.

Table 6.2: Per-property model performance comparison using Mean Absolute Error (MAE), the mean of each target variable, and Standardised MAE (MAE divided by the target's mean).

Model	***Target***	***MAE***	***Mean***	***Standardised MAE***
ANN	Hardness Brinell	19.77	164.18	0.12
ANN	UTS	64.69	558.79	0.11
ANN	YS	53.53	363.84	0.14
ANN	Elongation	2.5	21.89	0.10
XGBoost	Hardness Brinell	20.78	164.18	0.12
XGBoost	UTS	69.64	558.79	0.12
XGBoost	YS	59.94	363.84	0.16

XGBoost	Elongation	2.88	21.89	0.13
RF	Hardness Brinell	22.89	164.18	0.13
RF	UTS	75.41	558.79	0.13
RF	YS	62.81	363.84	0.17
RF	Elongation	2.85	21.89	0.13

It is important to highlight that, while the absolute MAE values in this study may appear higher than those reported in some prior literature (Xie et al., 2021; Zhao and Yang, 2025; Zheng et al., 2010), this is due to the nature of the dataset used. Unlike highly controlled academic datasets, the dataset in this work included a diverse and application-driven collection of standard carbon steel grades, with significant variation in both chemical composition and processing conditions. This real-world complexity introduces additional challenges but also makes the results more practically meaningful. In addition, due to the very small dataset, it was not possible to apply regression models independently to low, medium, and high carbon steels.

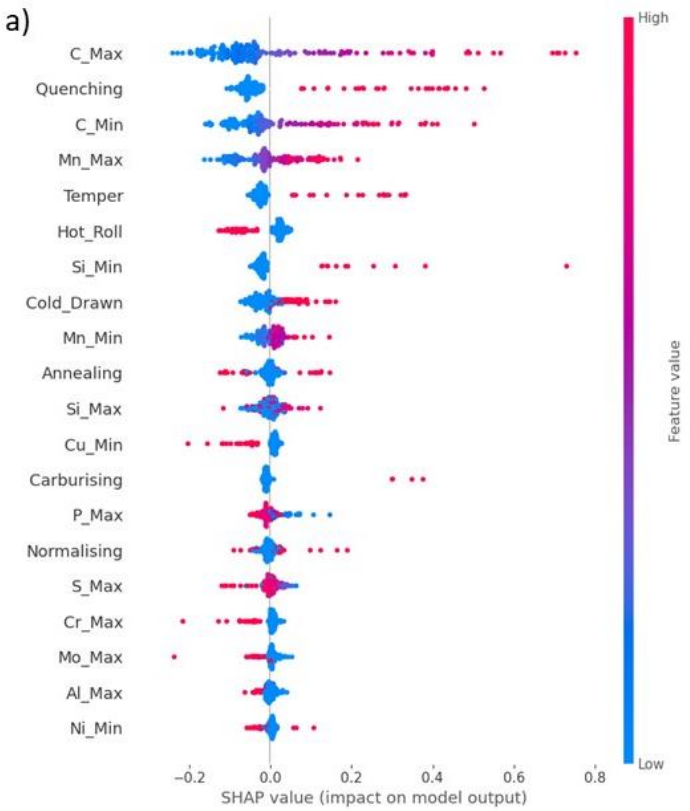
Overall, the ANN model demonstrated better generalisation and predictive accuracy and is therefore selected as the primary modelling tool for relationship analysis tasks. Specifically, this model will be integrated into inverse alloy design introduced in chapter 9 as future direction of this study.

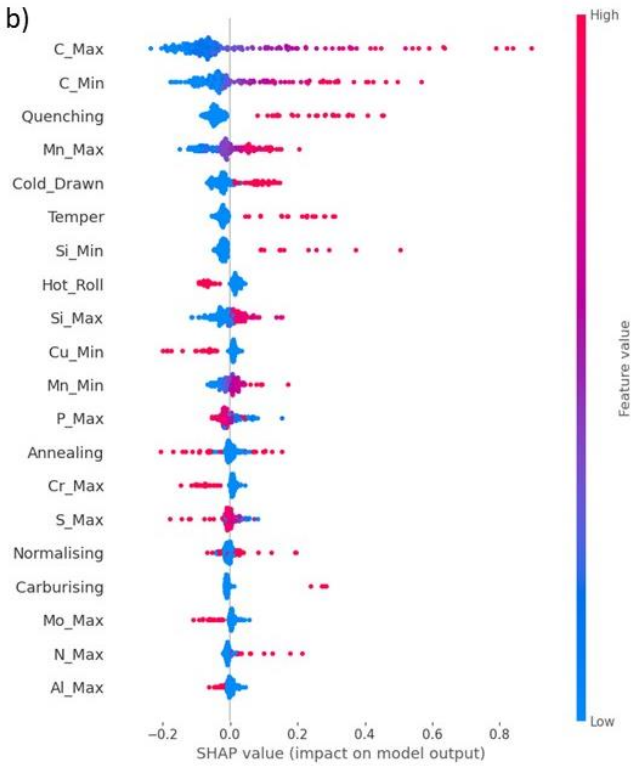
6.1.2.2 ANN SHAP Interpretation

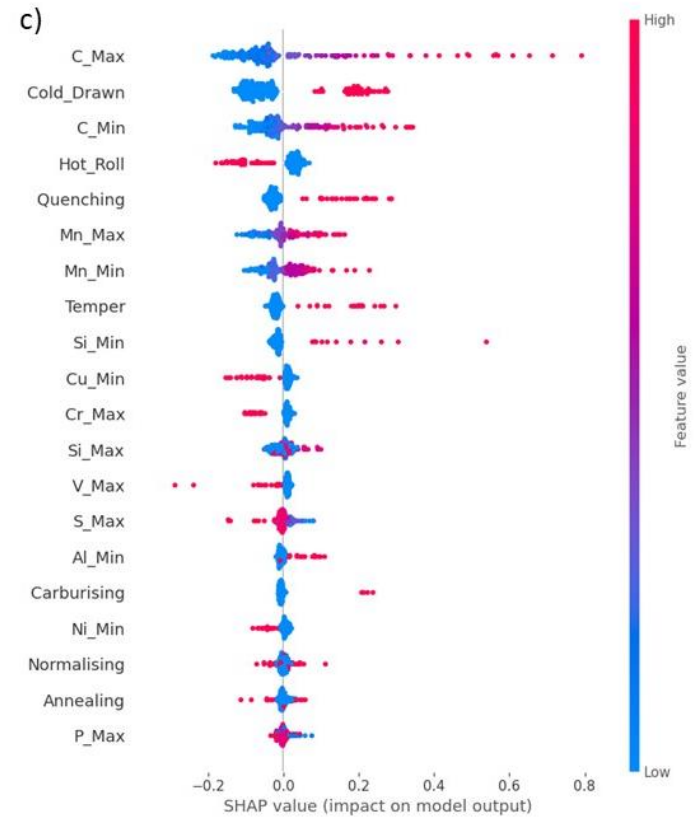
To better understand how chemical composition and processing conditions influence the mechanical properties of carbon steels, SHAP analysis was applied on the best-performing model (ANN). As shown in Figure 6.5(a–c), the SHAP summary plots reveal that carbon and manganese show the largest positive SHAP values for hardness, YS, and UTS. This indicates that increases in these elements have the strongest effect on raising predicted strength within the composition ranges represented in the dataset. This pattern aligns with well-established metallurgical behaviour, carbon enhances strength through solid-solution strengthening and by increasing the fraction of hard phases, while manganese stabilises austenite and improves hardenability, enabling the formation of stronger microstructures during cooling (Keehan et al., 2006). Additionally, lower copper content showed a negative impact on strength, suggesting that higher copper levels as a main tramp element contribute to increased strength in carbon steels. The negative SHAP values for elongation further reflect the expected strength–ductility trade-off, reinforcing that the ANN model has captured relationships consistent with known physical mechanisms.

In addition, the processing features show a similarly coherent pattern. Quenching, cold drawing, and tempering show strongly positive SHAP values for hardness and strength, reflecting their roles in producing martensitic or work-hardened microstructures with elevated dislocation density (Colpaert, 2018). Conversely, hot rolling is associated with negative SHAP contributions to strength, consistent with the formation of softer ferrite–pearlite microstructures and the recovery and recrystallisation processes occurring at elevated temperatures (Zhang et al., 2020).

Figure 6.5(d) presents the SHAP interpretation for elongation, which reflects the classical strength–ductility trade-off, reinforcing that the ANN model has captured relationships consistent with known physical mechanisms. Carbon and manganese both show negative SHAP values for elongation, indicating reduced ductility as these strength-enhancing mechanisms become more dominant. The behaviour of copper is also consistent with metallurgical expectations, at moderate levels Cu can contribute to ductility through relatively solid-solution effects, whereas at higher contents it participates in precipitation-related strengthening that lowers elongation (Salvetr et al., 2023). In terms of processing, cold drawing, hot rolling and tempering all decrease predicted elongation by introducing or retaining harder, less ductile microstructures, while annealing and normalising contribute positively by promoting recovery, recrystallisation and microstructural homogenisation, which together enhance ductility.


a)
C_Max
Quenching
C_Min
Mn_Max
Temper
Hot_Roll
Si_Min
Cold_Drawn
Mn_Min
Annealing
Si_Max
Cu_Min
Carburising
P_Max
Normalising
S_Max
Cr_Max
Mo_Max
Al_Max
Ni_Min
−0.2
0.0
0.2
0.4
0.6
0.8
SHAP value (impact on model output)
High
Feature value
Low


b)
C_Max
C_Min
Quenching
Mn_Max
Cold_Drawn
Temper
Si_Min
Hot_Roll
Si_Max
Cu_Min
Mn_Min
P_Max
Annealing
Cr_Max
S_Max
Normalising
Carburising
Mo_Max
N_Max
Al_Max
−0.2
0.0
0.2
0.4
0.6
0.8
SHAP value (impact on model output)
High
Feature value
Low



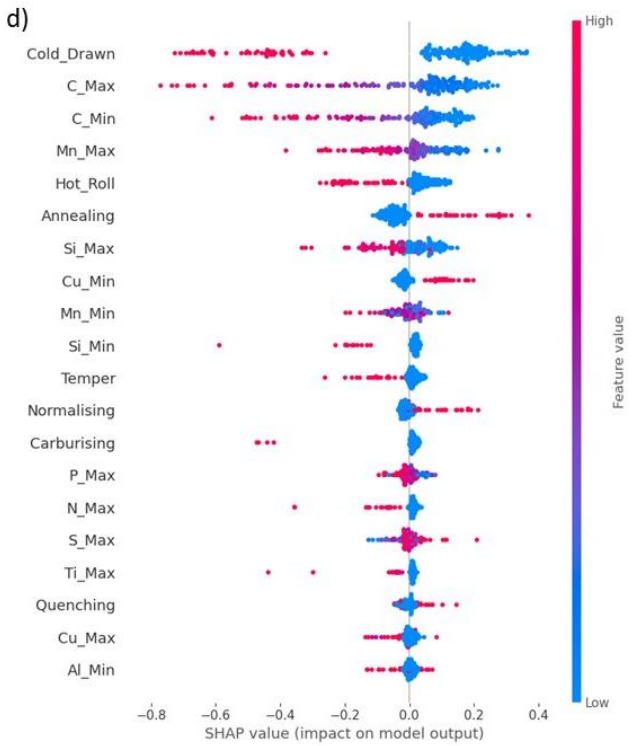


Figure 6.5: SHAP summary plots showing the influence of input features on (a) Hardness, (b) UTS, (c) YS, and (d) Elongation, based on the ANN model. Each plot displays SHAP value distributions per feature, with dots representing individual carbon

steel grades, ordered by feature importance. Positive SHAP values indicate that a feature increases the predicted property, while negative values indicate a decrease; the magnitude reflects the strength of that influence. The colour scale denotes the underlying feature value, with red indicating high values and blue indicating low values.

The results of the ANN model demonstrated strong predictive capabilities and provided a foundation for the subsequent clustering and reduction stages by identifying the chemical elements and processing conditions that most significantly influence mechanical behaviour in carbon steel grades.

6.2 Stainless Steel: Data Preparation and Modelling

6.2.1 Data Preparation and Clustering Suitability

The stainless steel dataset used in this study contained 301 grades sourced from five international standards including AISI, ASTM, UNS, EN and GB. It included both mechanical properties as dependent features, and chemical composition and processing conditions as independent variables. As outlined in chapter 4, although stainless steels include a wide range of international designations, many published grades lack fully specified chemical, processing, or mechanical data. Since such incomplete entries cannot be used for reliable modelling, the 301 records included in this study represent the subset of stainless steels for which complete and internally consistent information was available. Among the examined grades, austenitic steels were the most common with 150 entries, followed by 81 martensitic steels and 70 ferritic steels.

Figure 6.6 shows the distribution of UTS and YS across the examined stainless steel grades. Most grades contained UTS values between 550-650 MPa and YS values between 200-300 MPa, with both distributions being right-skewed. A small subset of steels exceeds 1000 MPa in UTS and YS, indicating a limited number of super high-strength outliers.

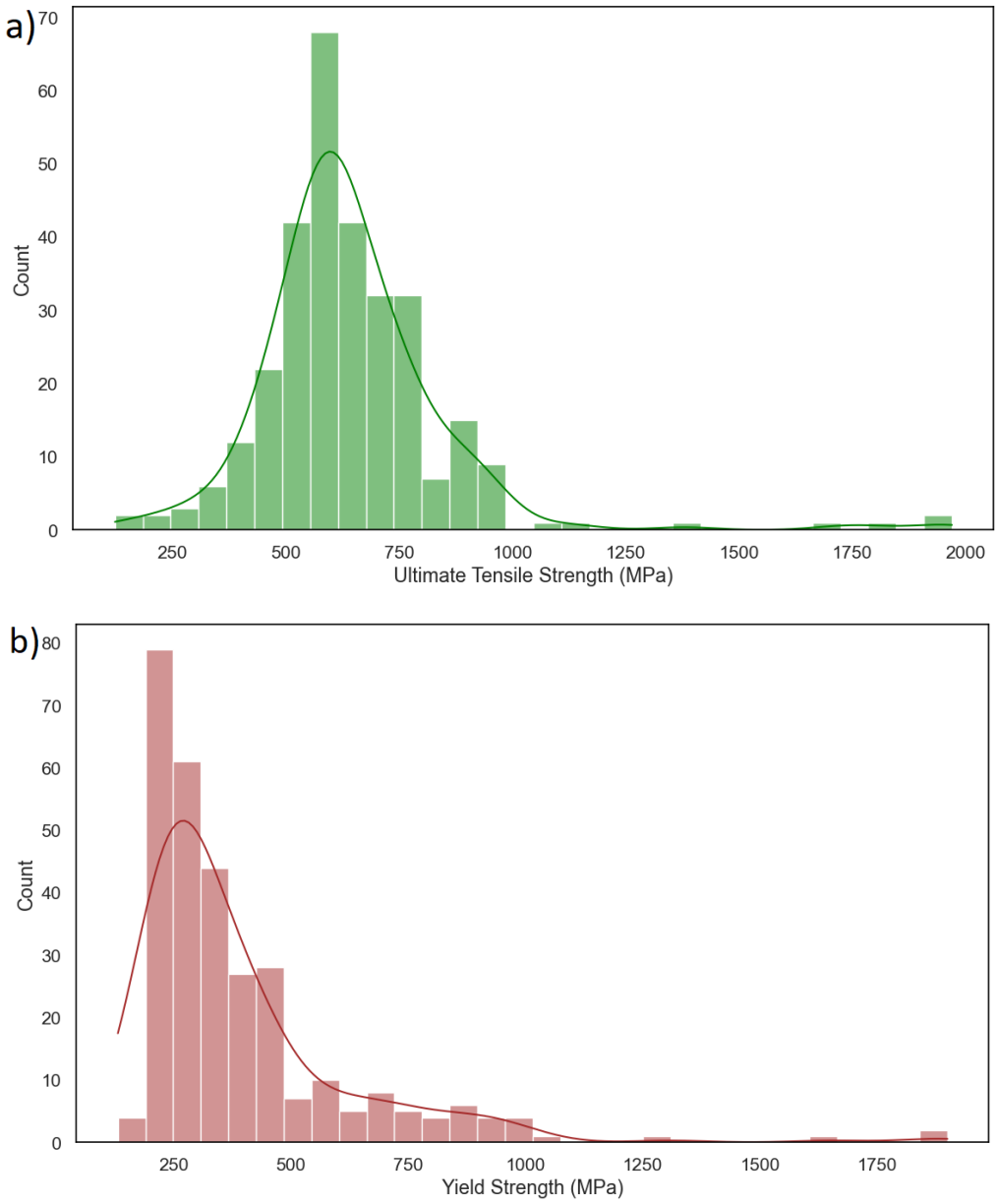


Figure 6.6: Histogram of (a) UTS and (b) YS with 50 MPa bin interval, represents the frequency distribution of stainless steel grades.

To examine chemical composition variation across stainless steel grades, Figure 6.7 illustrates the distribution of two key alloying elements included chromium and nickel. In figure 6.7(a), chromium content mostly ranged from 15% to 20%, with several high-chromium outliers observed in grades designed for enhanced corrosion resistance. Nickel levels were closely associated with steel type which ferritic and martensitic steels typically contained less than 4%, whereas austenitic steels contained concentrations exceeding 10% (Milititsky et al., 2008). Unlike some other elements, nickel has little to no strengthening effect on the yield stress of the austenitic phase which is consistent with the observed data. Austenitic steels in the dataset, despite their elevated nickel content, are primarily favoured for their ductility and corrosion

resistance rather than mechanical strength. This pattern reinforces the role of nickel as a phase stabiliser and corrosion enhancer rather than a strengthening element in stainless steels.

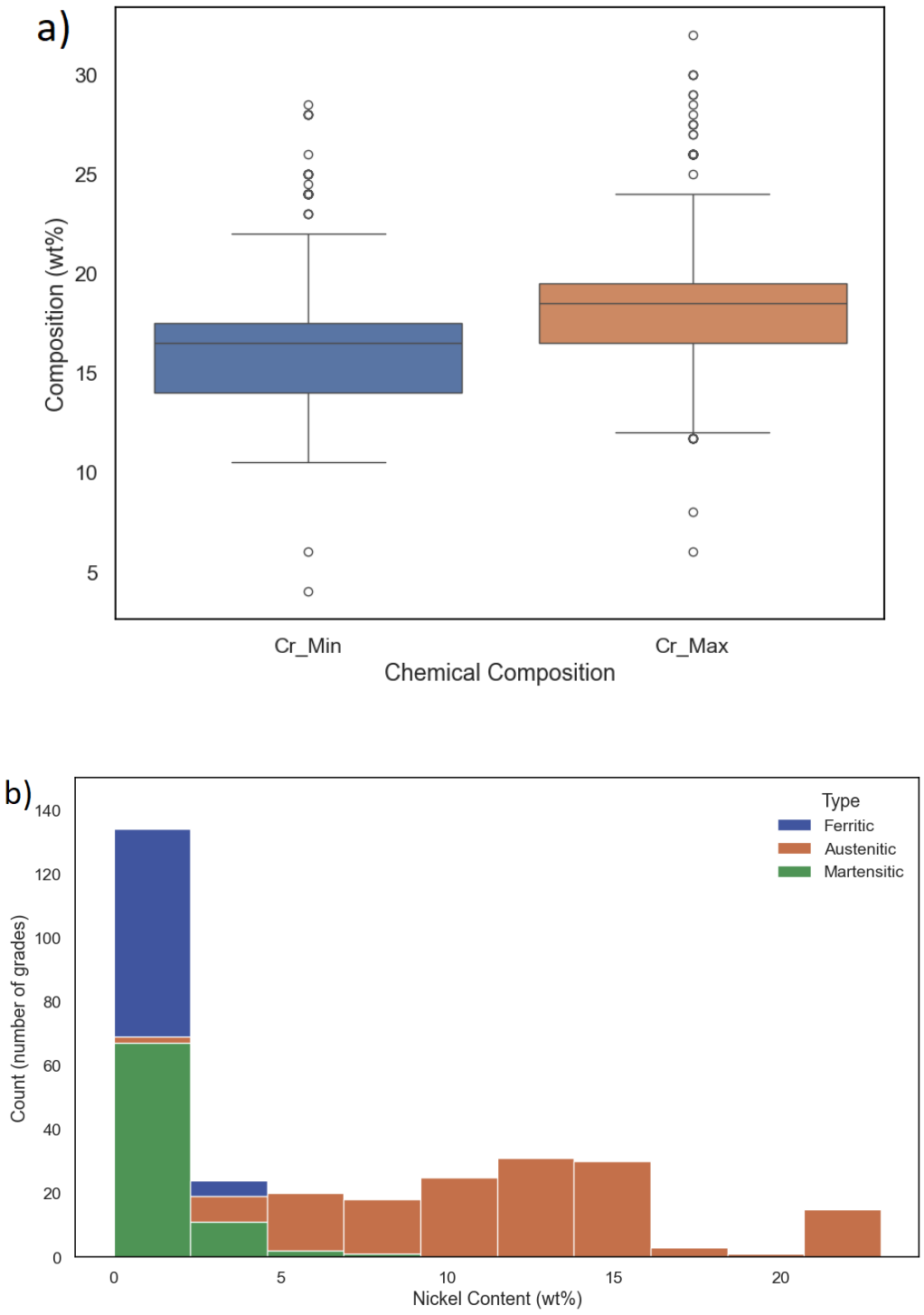


Figure 6.7: Range of chromium content in the examined stainless steel grades. Outliers represent grades with exceptionally high chromium levels compared to the majority.

These compositional patterns highlighted the chemical distinctions between stainless steel families. However, despite these differences, many grades demonstrated overlapping

mechanical properties particularly in UTS and YS as shown earlier in Figure 6.6. This observed overlap provides a strong rationale for applying the KMRP algorithm.

6.2.1.1 Standard Scaling

As previously described in section 6.1.1, standard scaling was applied to all numerical features to normalise the data prior to clustering and machine learning tasks for stainless steels to ensures that all features contribute equally regardless of their original scale or units. The same Standard Scaler procedure used for carbon steels, which involves subtracting each feature's mean and dividing by its standard deviation, was also applied here to place all numerical variables on a comparable scale.

For the stainless steel dataset, scaling was applied to a separate subset of chemical composition and mechanical property variables. The considered features included elements such as chromium and nickel, which are typically present in higher concentrations compared to carbon steels, as well as carbon, which is often tightly restricted in this group. Mechanical properties such as UTS and elongation were also standardised due to their different units. As in the previous group, binary-encoded processing variables were excluded from scaling. To illustrate the effect of the transformation, Table 6.3 presents an example showing a subset of features before and after standard scaling for five randomly selected stainless steel grades.

Table 6.3: Comparison of selected features before and after standard scaling in the examined stainless steel grades.

C_Min (wt %)	C_Max (wt %)	Cr_Min (wt %)	Cr_Max (wt %)	UTS (MPa)	Elongation (%)	C_Min_scaled	C_Max_scaled	Cr_Min_scaled	Cr_Max_scaled	UTS_scaled	Elongation_scaled
0.08	0.18	11.5	14	867	42	2	1.166667	-1.12378	-1.03104	1.819295	0.604919
0.0001	0.08	16.5	18.5	600	40	-0.5	-0.5	0.481621	0.465633	-0.5296	0.42161
0.0001	0.03	18	20	590	46	-0.5	-1.33333	0.963242	0.964526	-0.61758	0.971537
0.0001	0.08	18	20	694	34	-0.5	-0.5	0.963242	0.964526	0.297351	-0.12832
0.0001	0.18	11	13	550	15	-0.5	1.166667	-1.28432	-1.36364	-0.96947	-1.86975

6.2.1.2 Hopkin statistics

To assess whether the stainless steel dataset contains a natural clustering tendency and whether groups of similar grades exist, the Hopkins statistic was calculated following the approach described in section 6.1.1. Like carbon steels, the calculation was performed on the complete set of standardised numerical features, including both chemical compositions and mechanical properties, while excluding binary-encoded process indicators. In each evaluation, a random sample of n = 50 stainless steel grades was drawn from the dataset, and their nearest-neighbour distances were compared with those of randomly generated points in the same feature space. The calculation was repeated five times using different random seeds to ensure stability. The obtained values ranged from H = 0.75 to 0.80, and the representative value reported for this study is H = 0.78, indicating a strong tendency toward clustering. A value closer to 1.0 indicated a highly non-random, clustered distribution in this group. The results of Hopkins statistics justified the application of clustering algorithms, such as K-Means, DBSCAN, or hierarchical clustering, to identify groups of similar grades within this group.

6.2.2 Relationships Between Composition, Processing, and Properties in Stainless Steels

This section follows the same modelling approach applied to carbon steels (section 6.1.2), but focuses on stainless steel grades. The same set of supervised learning algorithms including ANN, XGBoost, and RF were applied to explore how chemical composition and processing conditions influence mechanical properties in stainless steels. The input features and target outputs remain consistent with the previous analysis. However, for stainless steels, one extra feature was added to indicate the type of steel (austenitic, ferritic, or martensitic). Including this information helped the models better understand the differences in behaviour across these steel families, since their properties can vary depending on their phase structure.

6.2.2.1 Model Evaluation and Comparison

Like carbon steels, the ANN, XGBoost, and RF algorithms were applied using a five-fold cross-validation approach, where the model was trained on four folds and tested on the remaining one. As shown in Figure 6.8, the average training MAE was lowest for the ANN model (37.60), followed by XGBoost (48.71) and RF (49.8), suggesting that ANN learned more effectively from the training data. In terms of validation performance, ANN also achieved the lowest average MAE (46.13), followed by XGBoost (54.55) and RF (57.45), indicating

better generalisation. The consistent gap of around 8 units between training and validation MAEs in all the models was expected due to the limited size of the dataset and also more complexity in the stainless steels data.

Further examination revealed that predictions for martensitic stainless steels showed significantly higher MAE compared to other types (austenitic, ferritic), often nearly double. This highlights the challenge of modelling such grades due to their variability in both composition and also small set of data for this type which was only 81 data points out of 301 data point in the stainless steels.

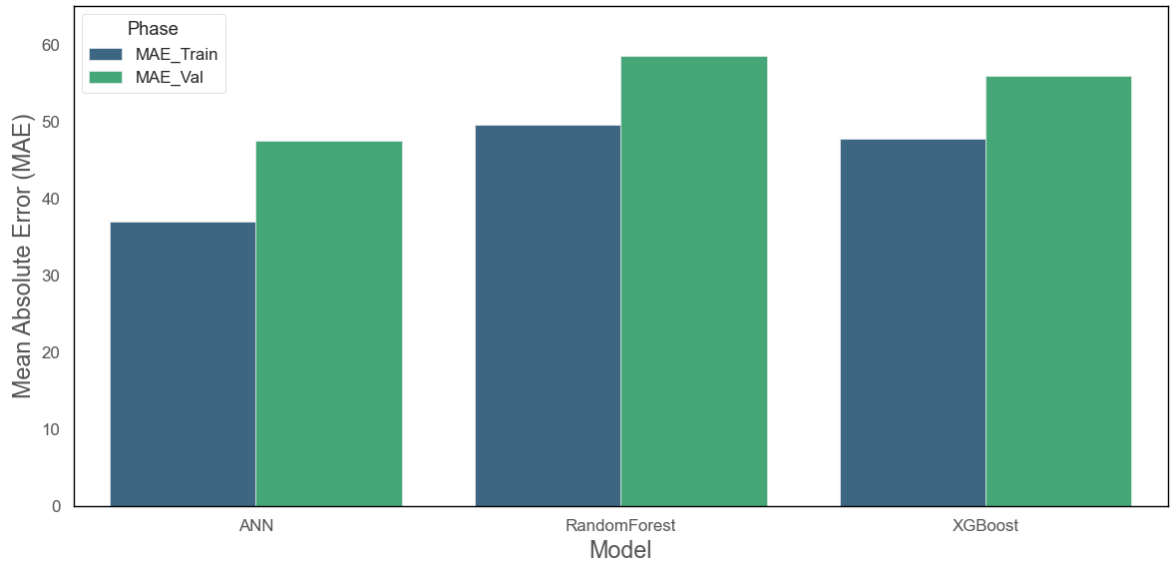


Figure 6.8: Average training and validation MAE across five folds for ANN, XGBoost, and Random Forest models in stainless steel grades.

To further evaluate the predictive performance of each model, a per-property comparison was carried out for the four key mechanical properties including hardness, UTS, YS, and elongation. The results are summarised in Table 6.4, including the MAE, the mean value of each target in the dataset, and the Standardised MAE (S-MAE), which expressed the error as a proportion of the average value for that property. Like in the carbon steel analysis, using S-MAE helps us fairly compare results across different properties.

Among the models, the ANN again achieved the best performance across most targets with the lowest MAE for hardness (26.86), UTS (79.31), and elongation (4.97), with corresponding S-MAE values of 0.127, 0.123, and 0.158. Although the S-MAE for YS (0.20) was higher than for the other properties, ANN still performed better than XGBoost and RF in this category. XGBoost showed competitive results, especially for elongation (S-MAE = 0.153 and MAE

4.8), but its S-MAE was consistently higher across all properties when compared to ANN especially for YS (0.246) and UTS (0.143).

These findings support earlier cross-validation results (Fig. 6.8), confirming ANN's better predictive reliability in this steel group. In addition, the consistently higher prediction errors for YS across all models, especially in tree-based models (RF and XGBoost), may reflect the greater sensitivity of yield strength to processing history or latent microstructural features not captured in the current feature set. This highlights a limitation in using composition and binary processing indicators alone to fully explain yield behaviour, especially in complex stainless alloys.

Table 6.4: Per-property model performance comparison using Mean Absolute Error (MAE), the mean of each target variable, and Standardised MAE (MAE divided by the target's mean).

Model	***Target***	***MAE***	***Mean***	***Standardised MAE***
ANN	hardness Brinell	26.86	211.29	0.12
ANN	UTS	79.31	643.48	0.12
ANN	YS	78.97	391.50	0.20
ANN	elongation	4.96	31.43	0.15
XGBoost	hardness Brinell	30.93	211.29	0.14
XGBoost	UTS	91.91	643.48	0.14
XGBoost	YS	96.23	391.50	0.24
XGBoost	elongation	4.81	31.43	0.15
RF	hardness Brinell	31.51	211.29	0.14
RF	UTS	95.48	643.48	0.14
RF	YS	101.86	391.50	0.26
RF	elongation	4.95	31.43	0.15

From a practical standpoint, the ANN model's ability to generalise across both carbon and stainless steel datasets positions it as the most suitable tool for capturing composition-property-processing relationships in steel applications.

6.2.2.2 ANN SHAP Interpretation

To gain insight into the influence of individual input features on the predicted mechanical properties of stainless steels, SHAP analysis was applied to the trained ANN model.

Figure 6.9(a) presents the SHAP summary plot for hardness predictions. Martensitic stainless steels, along with processing routes such as quenching, tempering, and the hardened condition, showed strong positive contributions to hardness. Among the alloying elements, carbon had

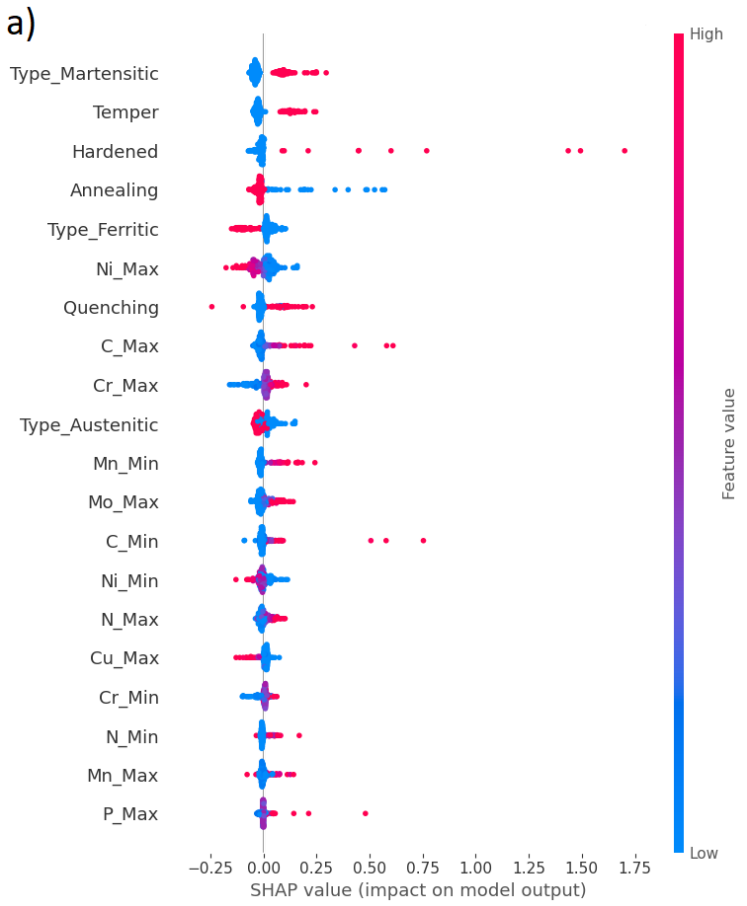
the most significant impact, with higher contents associated with increased hardness, reflecting its strong solid-solution and microstructural strengthening effects. Chromium also contributed positively in several cases, especially when present alongside hardening treatments, consistent with its role in carbide strengthening. In contrast, austenitic stainless steels and annealing treatments were negatively correlated with hardness, indicating their softening effect. This aligns with expectations from metallurgical literature (Wang et al., 2024). Additionally, higher nickel content, characteristic of austenitic grades, was also linked to reduced hardness, consistent with its known role in enhancing ductility rather than strength and its role in stabilising ductile FCC structures rather than strengthening (section 2.2.3).

Figure 6.9(b) summarises SHAP values for UTS prediction. As expected, annealing had a strong negative effect on UTS, while the hardened condition, quenching, and martensitic structure were dominant positive contributors. For the chemical composition perspective, both nitrogen and carbon enhanced UTS, aligning with their known strengthening mechanisms in stainless steels via solid-solution hardening and increased work-hardening capacity (Kim et al., 2018). Manganese showed a modest positive influence, likely due to its effect on hardenability. Nickel and austenitic classification, on the other hand, were consistently associated with reduced UTS, again reflecting their trade-off between strength and corrosion resistance.

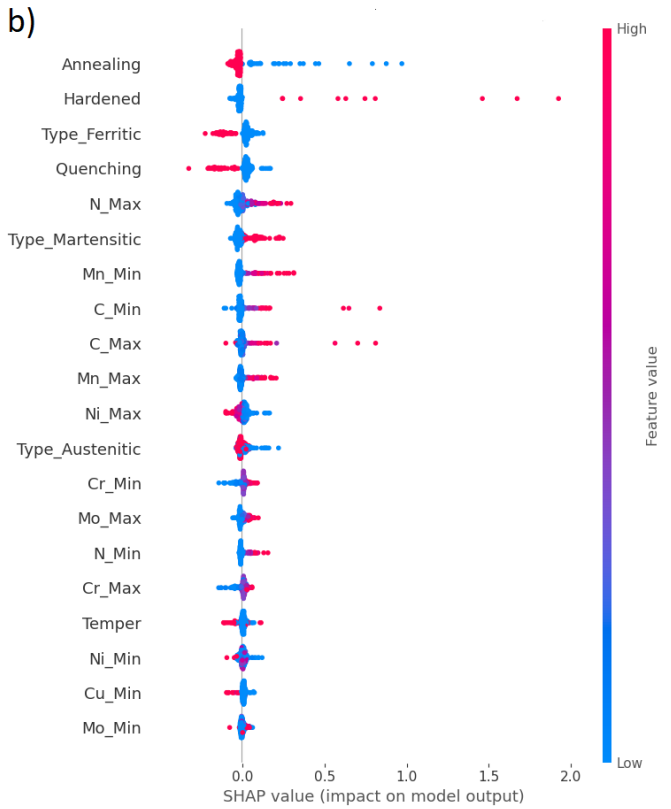
Figure 6.9(c) shows the SHAP results for YS. Similar to hardness and UTS, martensitic grades, quenching, tempering, and hardening all positively influenced YS. Carbon remained the most impactful chemical contributor, with chromium offering moderate support. Conversely, nickel, annealing, and austenitic structure showed negative SHAP values, reinforcing their softening roles in YS reduction.

Finally, Figure 6.9(d) presents the SHAP summary for elongation, highlighting the factors that influence the ductility of stainless steels. As expected, austenitic stainless steels and higher nickel content were strongly associated with increased elongation, consistent with their more ductile face-centered cubic structure and higher work-hardening capacity. In contrast, martensitic and ferritic types, known for their higher strength but lower ductility, showed a clear negative impact on elongation. In addition, treatments such as quenching and the hardened condition were associated with reduced elongation. On the other hand, annealing had a positive influence on elongation. Among the chemical elements, carbon showed a mild negative contribution and molybdenum was also associated with a slight decrease in

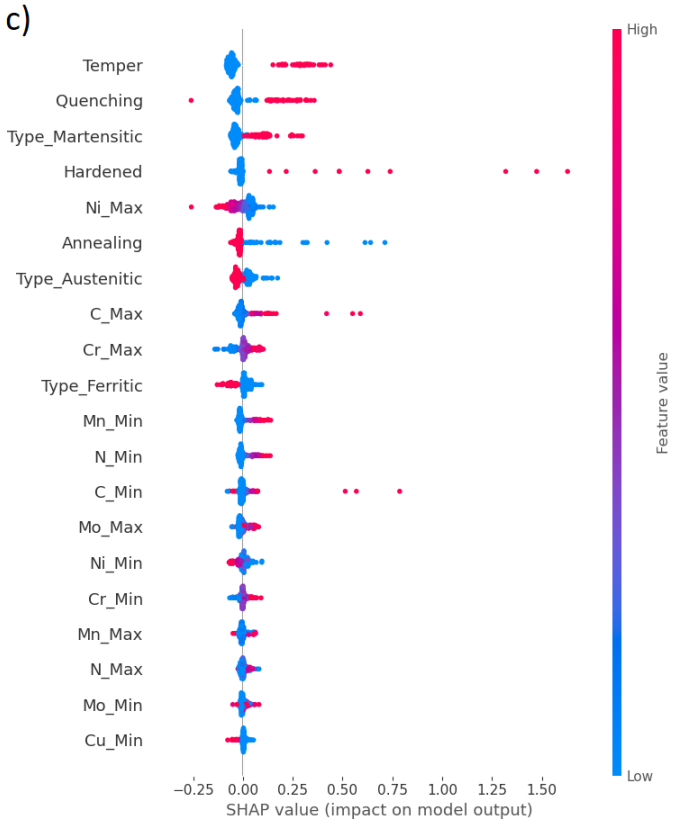
elongation, which aligns with its role in increasing strength through solid-solution and carbide-forming mechanisms.



b)


Annealing
Hardened
Type_Ferritic
Quenching
N_Max
Type_Martensitic
Mn_Min
C_Min
C_Max
Mn_Max
Ni_Max
Type_Austenitic
Cr_Min
Mo_Max
N_Min
Cr_Max
Temper
Ni_Min
Cu_Min
Mo_Min
0.0
0.5
1.0
1.5
2.0
SHAP value (impact on model output)
High
Feature value
Low

c)


Temper
Quenching
Type_Martensitic
Hardened
Ni_Max
Annealing
Type_Austenitic
C_Max
Cr_Max
Type_Ferritic
Mn_Min
N_Min
C_Min
Mo_Max
Ni_Min
Cr_Min
Mn_Max
N_Max
Mo_Min
Cu_Min
−0.25
0.00
0.25
0.50
0.75
1.00
1.25
1.50
SHAP value (impact on model output)
High
Feature value
Low

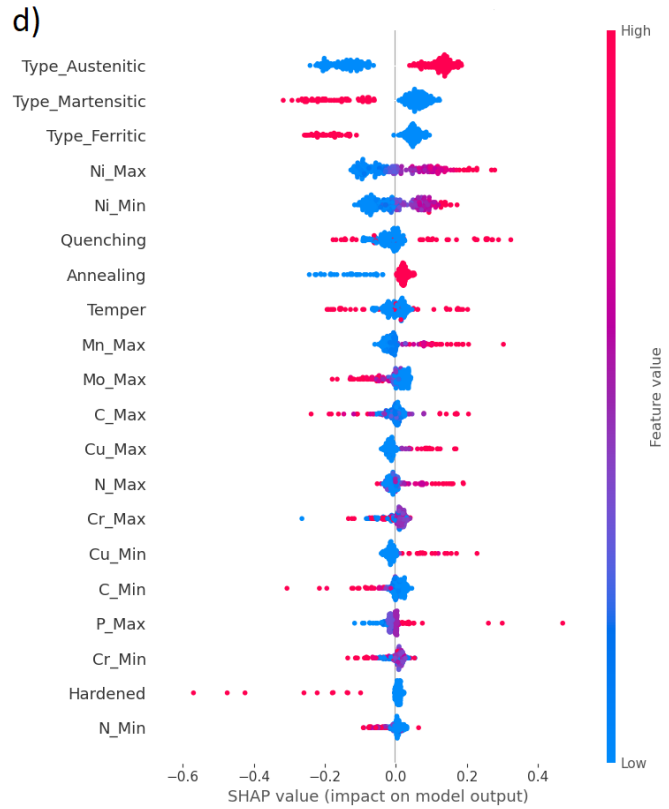


Figure 6.9: SHAP summary plots showing the influence of input features on (a) Hardness, (b) UTS, (c) YS, and (d) Elongation, based on the ANN model. Each plot displays SHAP value distributions per feature, with dots representing individual stainless steel grades, ordered by feature importance. Positive SHAP values indicate that a feature increases the predicted property, while negative values indicate a decrease; the magnitude reflects the strength of that influence. The colour scale denotes the underlying feature value, with red indicating high values and blue indicating low values.

These findings demonstrated how the combination of chemical elements, heat treatments, and steel types shapes the mechanical behaviour of stainless steels (objective 2). The ANN model not only captured these relationships with strong predictive accuracy but also successfully reproduced known metallurgical trends. By understanding how each feature influences properties like hardness, YS, UTS, and elongation, the analysis gives a clearer view of how these steels perform. This understanding also lays the groundwork for future inverse alloy design, which is discussed as a direction for future work in this study.

6.3 Summary

This chapter presented a data-driven analysis of both carbon and stainless steel datasets to prepare them for grade rationalisation. Standard scaling was applied to normalise chemical composition and mechanical property features, ensuring fair comparison and enabling distance-based machine learning methods. The Hopkins statistic indicated strong clustering tendencies in both datasets (H = 0.86 for carbon steels and H = 0.78 for stainless steels), confirming the presence of a meaningful internal structure and supporting the application of

unsupervised learning in later stages. To quantify the influence of composition and processing conditions on performance, three predictive models (ANN, XGBoost, and RF) were trained for each steel type. Across both datasets, the ANN consistently outperformed the other models in generalisation and accuracy, making it the preferred tool for property prediction. SHAP analysis of the ANN models revealed domain-aligned insights, where in carbon steels, strength was driven by carbon and manganese, with quenching enhancing strength and annealing improving ductility. In stainless steels, martensitic grades and hardening treatments increased strength, while austenitic structures and high nickel content favoured ductility. These results not only reinforced metallurgical knowledge but also exposed how tramp elements and critical alloying choices affect mechanical trade-offs. Together, the findings addressed objective 2 of this study and established a solid foundation for rationalisation workflows developed in the following chapters.

Chapter 7: Steel Reclassification using KMRP

This chapter presents the results of applying the KMRP to rationalise the number of existing steel grades. Carbon and stainless steel datasets were merged into a unified dataset to enable global clustering based on mechanical properties. The goal of this global mapping was not to imply direct substitutability between steel families, but to construct a unified performance landscape that captures how steels behave functionally, independent of traditional standards or alloy classifications. While grades were grouped globally based on shared mechanical characteristics, the rationalisation process remained metallurgically valid by applying reduction only within steel types (e.g., carbon-to-carbon or stainless-to-stainless).

The results are presented phase by phase, corresponding to the stages of the KMRP framework. Together, these steps demonstrate how the proposed method simplifies the steel system while preserving the full coverage of mechanical property ranges, prioritise the grades with tolerance to tramp elements where applicable, and reducing reliance on CRMs.

7.1 Clustering Algorithm Selection

The objective of this section is to benchmark commonly used clustering algorithms, including HC, DBSCAN, and K-Means and select the one that most effectively groups steel grades with similar mechanical behaviour. These groups form the foundation for the subsequent reduction phase, ensuring that any rationalisation is grounded in functional equivalence across selected mechanical properties.

7.1.1 Hierarchical Clustering Result

Agglomerative hierarchical clustering was initially applied due to its interpretability and ability to reveal nested data structures without requiring a predefined number of clusters. Visual inspection of the dendrogram suggested four clusters, as four major branches emerge before any substantial increase in linkage distance, indicating the most natural cut in the hierarchical tree. However, these clusters showed moderate separation. In addition, the Silhouette score was 0.41 and the DBI was 0.84, which indicated overlapping clusters and limited separation between groups. These values indicate moderate intra-cluster similarity and considerable overlap between clusters. PCA projections confirmed this, particularly for clusters 2 and 3, which showed significant overlap in UTS and YS space. This can be attributed to the limitations of agglomerative hierarchical clustering, which applies irreversible merging without global optimisation. Such rigidity makes it unsuitable for datasets with continuous and

overlapping distributions, such as mechanical property data (Vichi et al., 2022). While the method helped explore the hierarchical structure of the data, it was unable to capture the well-separated performance-based groups for the examined steel grade data. The dendrogram and corresponding PCA visualisation are provided in appendix A.

7.1.2 DBSCAN Result

DBSCAN was selected for evaluation due to its ability to detect non-spherical clusters and exclude outliers as noise, unlike centroid-based methods such as K-Means which assume convex clusters. It performs better for datasets with irregular geometries and varying densities (Ester et al., 1996b). Given the potential variability in mechanical properties across different steel grades and types, this feature was initially hypothesised to be beneficial.

To assess the suitability of DBSCAN, the algorithm was applied to the selected scaled mechanical property features, with tuning of its two key hyperparameters, including eps (radius of neighbourhood) and min_samples (the minimum number of points to form a dense region). The k-distance plot using 5-nearest neighbours revealed a transition between dense and sparse regions around eps = 0.68. A grid search across selected eps around the 0.68 and min_samples combinations was conducted to evaluate clustering quality using Silhouette Score and DBI. The best configuration of eps = 0.8 and min_smaple = 4 achieved a Silhouette Score of 0.47 and a DBI of 1.54 (Table 7.1). The observed improvement in Silhouette score indicated that DBSCAN was more effective at identifying compact and well-separated clusters in the denser regions of the mechanical properties feature space (medium-strength steels), which can be attributed to its density-based formulation that does not rely on global distance metrics (Fahim, 2023).

However, the high DBI value indicates that DBSCAN struggled to maintain compactness and separation uniformly across the entire dataset. In regions of the feature space where mechanical properties were more gradually distributed or sparsely populated, the algorithm failed to establish clear density-based boundaries, particularly between mid-strength and high-strength steel grades due to insufficient neighbourhood density. This behaviour reflects a known limitation of DBSCAN in handling datasets with uneven distributions, as documented by (Schubert et al., 2017).

PCA analysis (appendix A) further illustrated this issue, that although DBSCAN identified three valid clusters, it also misclassified some grades, particularly those with superior and distinct mechanical properties, as noise (cluster -1). This was a group of grades previously

recognised as a valid cluster in hierarchical clustering (cluster 2), but in DBSCAN, it was entirely detected as noise. This result highlighted the limited application of the DBSCAN algorithm in this study, as these grades represent outliers not due to noise or error, but rather to unique and valid mechanical performance. Their elimination risks ignoring the valuable alloy design candidates with specialised application potential. The corresponding k-distance curve and PCA projections are provided in appendix A.

Table 7.1: Evaluation of DBSCAN clustering performance under different combinations of eps and min_samples values.

eps	min_samples	n_clusters	n_noise	Silhouette score	Davies Bouldin Index
0.65	3	4	59	0.40	1.60
0.65	4	3	66	0.45	1.57
0.65	5	3	67	0.45	1.55
0.7	3	4	57	0.40	1.63
0.7	4	3	62	0.45	1.60
0.7	5	3	64	0.45	1.60
0.75	3	5	52	0.38	1.64
0.75	4	4	56	0.40	1.78
0.75	5	3	61	0.46	1.61
0.8	3	4	50	0.44	1.48
0.8	4	3	55	0.47	1.54
0.8	5	3	57	0.46	1.59

7.1.3 K-Means Result

K-Means algorithm was applied due to its efficiency in handling numerical data, ease of interpretation, and strong internal validation performance. Similar to the other algorithm, it was applied to the scaled mechanical property dataset to identify similar performance-based clusters of steel grades. One of the limitations of K-Means, as reported in the literature (Chiang and Mirkin, 2010; Ikotun et al., 2023), is its sensitivity to the user-defined number of clusters, or K, which can significantly influence the outcome. To address this, the Elbow method and Silhouette score were used to identify the optimal number of clusters. The result showed that the transition at K = 4 started, and further increases in K did not improve the cluster compactness (detailed discussion is provided in the following section). This result suggests the presence of four dominant groupings of steel grades based on mechanical performance within the dataset.

As shown in Table 7.2, K-Means achieved the highest Silhouette score (0.58) and the lowest DBI (0.65) among the evaluated algorithms. These results indicated a better intra-cluster

cohesion and inter-cluster separation compared to HC (Silhouette = 0.41, DBI = 0.85) and DBSCAN (Silhouette = 0.47, DBI = 1.54).

The algorithm's improved performance can be attributed to the underlying structure of the mechanical property space, where hardness, UTS, and YS showed strong linear correlations (Fig. 6.2). These correlations cause the data to align along principal directions, which led to compact and directionally consistent groupings. After scaling, such patterns approximating convex regions in the feature space. In this context, K-Means is particularly well-suited because it partitions the space based on Euclidean distance to cluster centroids, as its distance-based assignments effectively capture variance in low-dimensional and convex groupings (Fahim, 2021). Furthermore, its ability to iteratively refine centroid positions enabled local optimisation of cluster assignments, which resulted in tighter and more meaningful groupings (Fränti and Sieranoja, 2019). In contrast, as seen in DBSCAN, it underperformed in sparsely populated regions where local density was insufficient to form clusters, and HC lacked centroid adjustment mechanisms, leading to suboptimal separation.

Table 7.2: Comparison of clustering algorithm performance on scaled mechanical property data using Silhouette Score and Davies-Bouldin Index.

Algorithm	*No. of Clusters*	*Silhouette Score*	*Davies-Bouldin Index*
Hierarchical	4	0.41	0.85
DBSCAN	3	0.47	1.54
K-Means	4	0.58	0.65

These results indicate that K-Means achieved better intra-cluster cohesion and inter-cluster separation compared to the other algorithms, making it more effective at identifying groups of steel grades with similar mechanical performance. These finding supports its selection as the foundational clustering method within the proposed KMRP reduction framework. Detailed validation metrics and supporting PCA visualisations are provided in following section.

7.2 K-Means Reduction Process (KMRP) Result

Building upon the benchmarking analysis in previous section, this section presents the application of the proposed KMRP framework to the full mechanical property dataset. The complete set of retained and eliminated steel grades generated by this process has been archived in an open-access repository to support transparency and reproducibility (Jalalian, 2026). The analysis proceeds in three phases: (i) identifying the optimal number of clusters, (ii) grouping

steel grades based on mechanical performance characteristics, and (iii) reducing grade variety using criteria related to recyclability and critical raw material content.

7.2.1 Phase I: Finding Optimal Number of Clusters Result

The first phase of the KMRP framework focuses on determining the optimal number of clusters (K) based on the selected scaled mechanical property space (hardness, UTS, YS, elongation). Selecting an appropriate K is critical for capturing meaningful performance-based variability among steel grades because poorly defined clusters could lead the rationalisation process to compare and eliminate grades with fundamentally different mechanical behaviours, which could risk the loss of critical performance characteristics in some applications.

To guide K selection, the elbow method was applied across a wide range of cluster numbers (K = 2 to 20). As shown in Figure 7.1(a), the distortion score, which is defined as the within-cluster sum of squared distances to the centroid, decreased sharply as K increased from 2 to 5. The rapid decline indicated a significant improvement in intra-cluster compactness, suggesting that the data structure is well captured up to this point. However, the improvement became very small after K = 5, and the curve started to flatten, which suggested that adding more clusters did not significantly enhanced grouping compactness. Therefore, K = 5 was initially considered the optimal choice and adding additional splits may begin to reflect minor variance or noise. However, Figure 7.1(b) shows that when focusing on a narrower range of K values (2 to 10), the elbow method suggested K = 4 as the optimal number of clusters, which indicated that the optimal number may vary depending on the selected K range in this method. The shift occurred was because the elbow method relies on visual interpretation, and the apparent "elbow" can appear differently depending on the width of the K range. In wider ranges, early changes can appear less noticeable, whereas a narrower range can exaggerate them. Although the fit time is lower at K = 5, indicating greater computational efficiency but fit time is not a consideration in this study. To strengthen the selection of an appropriate cluster count, silhouette analysis was applied to compare the quality of the K = 4 and K = 5 configurations identified by the elbow method. While both configurations yielded relatively high average Silhouette Scores, closer inspection revealed important structural differences.

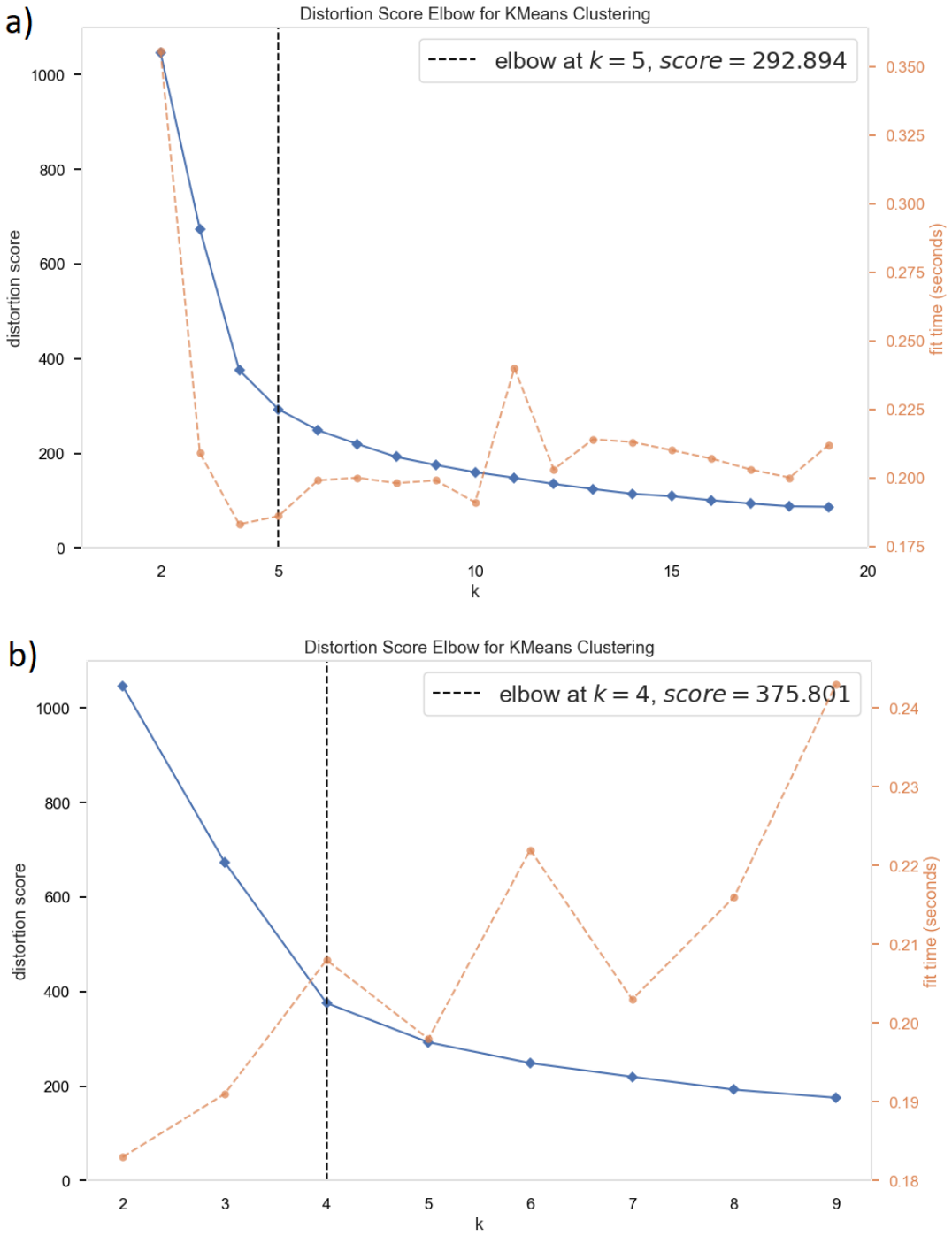


Figure 7.1: Elbow method score to identify the optimum number of clusters based on (a) K=20 and (b) K=10 on the examined carbon and stainless steel grades. The blue line shows the distortion score and the orange dashed line shows the fitting time.

Figure 7.2 demonstrates the Silhouette Scores for each clustering configuration. Despite the average Silhouette values for both configurations were close, the plot for K=5 revealed one cluster containing several points with scores below the overall average (indicated by the red dashed line) that suggested poor fit and potential boundary overlap. In contrast, the clustering result for K = 4 showed better internal cohesion, with most samples achieved Silhouette Scores above the mean. Therefore, although K = 5 showed slightly lower within-cluster distortion, the

K = 4 configuration achieved a more interpretable and stable clustering structure, and it suggests that in the examined steel grades there are 4 group of steel grades that each shared the distinct property characteristics. Nevertheless, overlaps in property ranges within clusters indicated the presence of intra-cluster redundancy, justifying further reduction in subsequent stages of the KMRP framework.

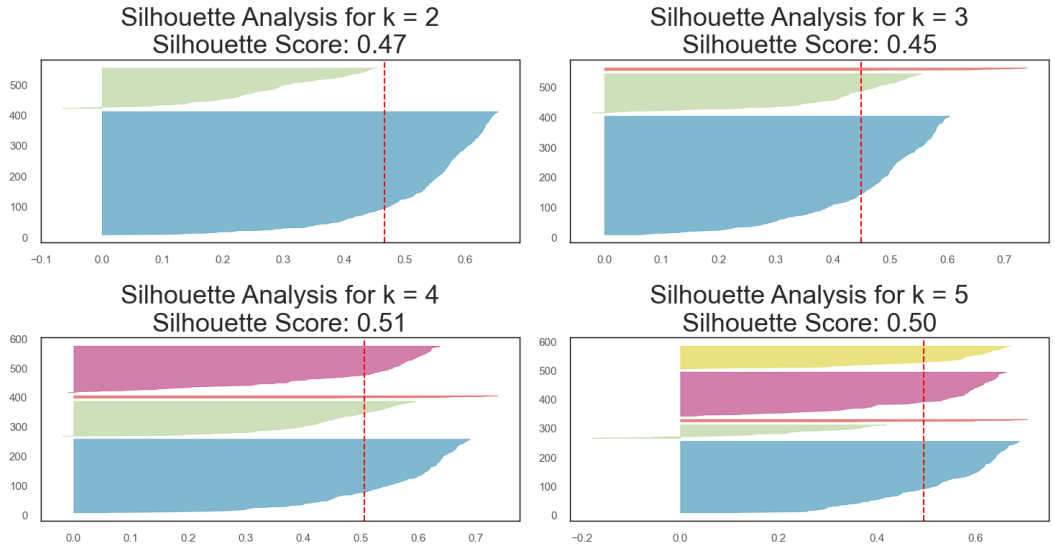


Figure 7.2: The silhouette scores for clustering the examined carbon-stainless steel grades based on different cluster configurations. The red dashed line indicates the average silhouette score, while the width of the silhouette bar represents the number of grades within each cluster.

After identifying the optimal number of clusters as K=4, PCA was applied to reduce the dimensionality of the scaled mechanical property space and improve the interpretability of cluster structures. This step was particularly important due to the high degree of correlation observed between hardness, UTS, and YS (as previously shown in Figure 6.2). In clustering algorithms such as K-Means, which rely on Euclidean distance, highly correlated features can introduce geometric distortions. Specifically, when multiple features convey similar information, they effectively overweight that dimension in the distance calculation, leading to biased clustering outcomes and reduced sensitivity to variation in less dominant directions. PCA addresses this by transforming the correlated variables into a new set of uncorrelated principal components through an orthogonal rotation of the original feature space, obtained by computing the eigenvectors of the covariance matrix, each capturing a unique axis of variation in the data (Hastie et al., 2009; Jolliffe and Cadima, 2016).

As shown in Table 7.3, the first principal component (PC1) shows similar positive component loadings for hardness, UTS and YS (0.544, 0.545 and 0.537, respectively), with a smaller

negative loading for elongation (-0.341). This indicates that PC1 predominantly represents a combined strength dimension within the dataset. The second principal component (PC2) is dominated by a strong positive loading from elongation (0.939), with only minor contributions from the strength-related variables, and therefore captures variability associated primarily with ductility, independent of strength. Together, PC1 and PC2 explain approximately 85% of the total variance (based on the PCA output), meaning that the two-dimensional PCA space retains more than 80% of the mechanical variability relevant to this study. This decomposition is consistent with the classical strength–ductility trade-off in steels and provides a physically meaningful basis for visualising and interpreting the clusters.

Table 7.3: Contribution of different properties to the positioning of the data points in the PCA space.

	Hardness Brinell	***UTS***	***YS***	***Elongation***
PC1	0.544	0.545	0.537	-0.341
PC2	0.215	0.212	0.161	0.939

As a result, four distinct performance-based groups identified in this phase using the elbow method and silhouette analysis, and data projected into a two-dimensional PCA space, the next phase applied K-Means clustering to group the steel grades and examine the characteristics of each cluster for identifying internal patterns and preparing the dataset for subsequent grade rationalisation based on mechanical similarity.

7.2.2 Phase II: K-Means Clustering Result

After determining the optimal number of clusters, the K-Means algorithm was applied to the scaled mechanical property dataset, and the results were visualised using PCA. Figure 7.3 illustrates the distribution of the steel grades within the reduced PCA space, where PC1 represents the principal direction associated mainly with strength-related properties (hardness, UTS, YS), while PC2 captures variation dominated by elongation and therefore reflects ductility. Each point corresponds to an individual steel grade, and the "×" markers indicate the cluster centroids.

Clusters 0 and 3 showed as compact and well-defined clusters along the PC1-PC2 axes, indicating that the steels within these clusters showed consistent strength–ductility characteristics. Cluster 1 demonstrated a wider spread but remained concentrated within a continuous region of the PCA space, which suggested a broader yet coherent range of mechanical responses. In contrast, the grades assigned to cluster 2 were dispersed across the

PCA projection and occupied regions that lay outside the dominant variability patterns captured by PC1 and PC2. This dispersion reflected the distinct mechanical property ranges of these grades and indicated that their behaviour differed from the main clusters. As a result, K-Means separated them into a small, independent cluster rather than merging them with the larger groups, which would have reduced internal cohesion and weakened the overall clustering structure.

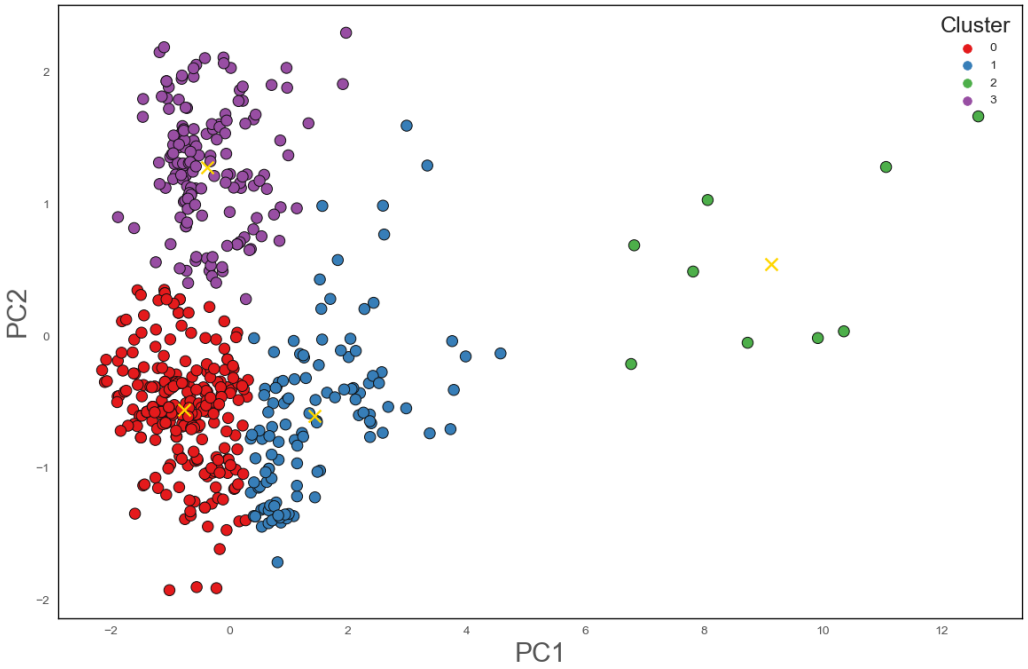


Figure 7.3: Distribution of the examined grades of carbon steel grades on the PCA values. PC1 primarily represents variation associated with strength-related properties (hardness, UTS, YS), while PC2 reflects variation dominated by elongation. Each point corresponds to a grade and the "×" markers represent the centroid of each cluster.

To interpret the K-Means clustering results, the distribution of grades by cluster, standard, and steel type was analysed (Fig. 7.4). Cluster 0 included the highest number of grades (n = 186), with a majority being carbon steels, particularly from the AISI and EN standards. Cluster 3 (n = 148) was contained almost entirely of austenitic stainless steels. Cluster 1 (n = 91) showed a relatively balanced distribution across standards and steel types, although stainless steels were slightly more frequent. Cluster 2 (n = 9) had the fewest grades and included both carbon and stainless steels. The detail discussion about the characteristics of the generate clusters has been discussed in chapter 8.

The presence of both carbon and stainless steels within the same clusters suggests that specific grades from different alloy families showed overlapping performance when assessed only on strength and ductility metrics. Although these materials differ fundamentally in microstructure (e.g., ferritic vs. austenitic phases) and alloying elements (e.g., Cr, Ni, Mo), as well as service

requirements such as corrosion resistance and weldability, these distinctions are not reflected in the selected mechanical property space used for clustering (Dundu, 2018).

These finding suggests that performance-based classification offers an alternative view of steel grade rationalisation that is driven by application requirements rather than conventional taxonomy. However, it also points to a fundamental limitation that if the clustering were extended to include additional features such as corrosion resistance, thermal properties, or weldability, the current group boundaries would likely shift. Expanding the dimensionality of the feature set would provide a more holistic view and may reinstate distinctions that mechanical properties could not capture. Therefore, while four distinct clusters were formed based on mechanical similarity, the subsequent reduction phase directly aligns with metallurgical boundaries by enforcing comparisons within pre-defined alloy categories to preserve functional and compositional validity.

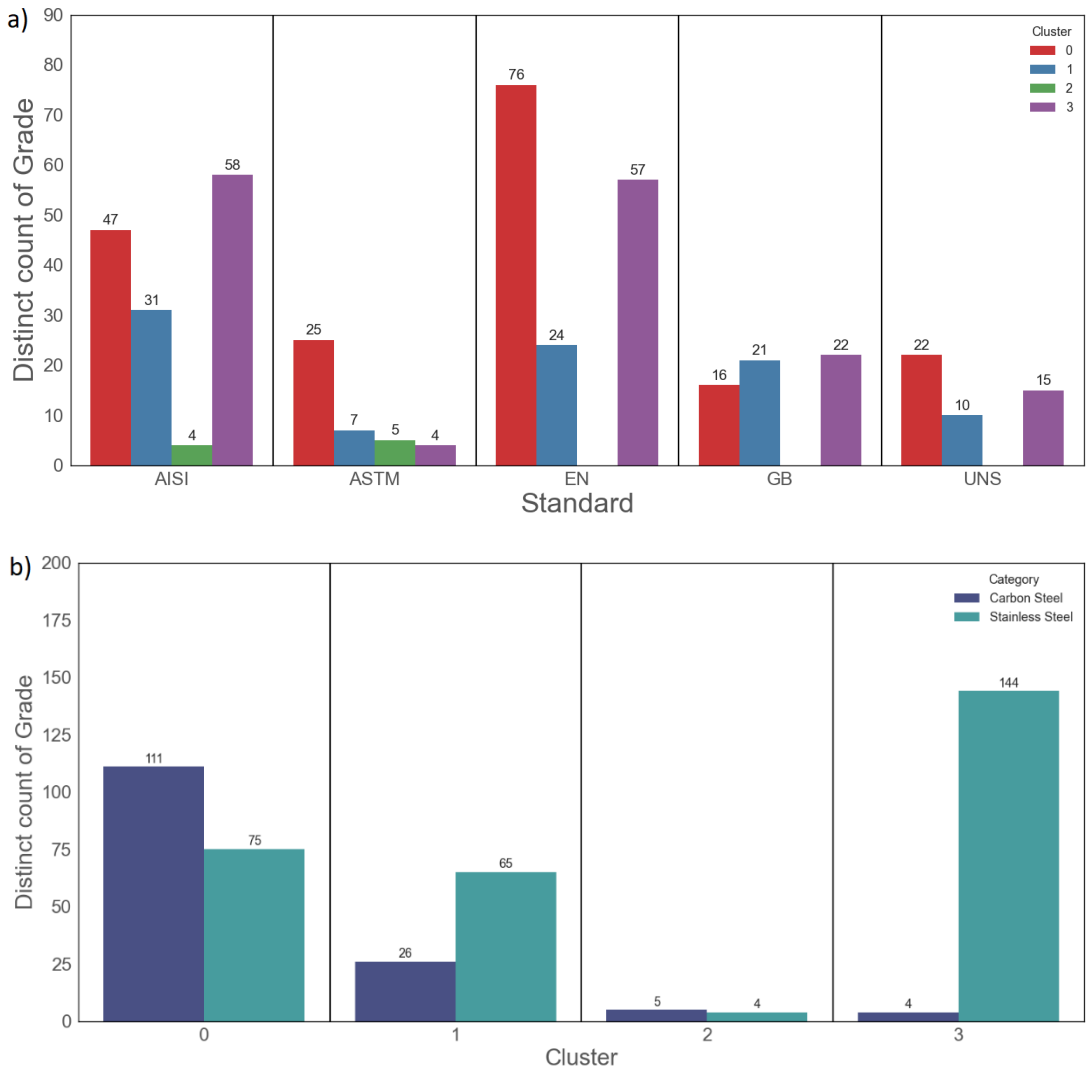


Figure 7.4: The alloy grades distribution based on (a) various standards (AISI, ASTM, UNS, EN and GB) and (b) different steel type. The x-axis represents the cluster number, while the y-axis represents the count of alloy grades within each cluster.

To validate whether performance-based clustering aligns with known compositional groupings, the average carbon content of each cluster was analysed (Fig. 7.5). Clusters 0 and 3 showed low carbon content, which was consistent with the low-carbon and austenitic stainless steels in these clusters, respectively. Cluster 1 showed the highest average carbon content, aligning with the expected properties of high-carbon steels (increased strength but reduced ductility). In contrast, cluster 2 showed a moderate carbon content, indicating grades that may represent hybrid performance classes. The alignment between mechanical clustering and chemical variation, strengthens the metallurgical validity of the clustering results. In addition, it suggests that mechanical properties alone can perform as effective representatives for compositional grouping when alloy chemistry is unavailable or incomplete.

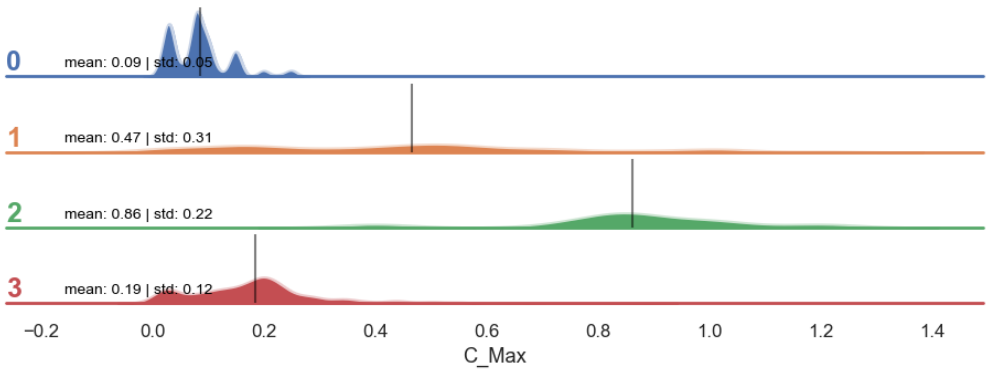


Figure 7.5: Distribution of carbon content within each cluster of steel grades grouped based on their mechanical properties. The black vertical line indicates the mean carbon content in each group. The x axis represents the range of carbon content, while the y axis represents the cluster identifier.

These results demonstrated that the K-Means phase effectively prepared performance-based clusters and grouped the grades that shared similar mechanical property coverage but differed in chemical composition and processing conditions (objective 3), and it could successfully capture not only performance similarities but also indirectly aligned with traditional carbon-based classification. The next step is to simplify the grade system by comparing the grades within each cluster and reducing the grades by prioritising those with higher recyclability and lower content of critical raw materials.

7.2.3 Phase III: Reduction Process Result

This phase supports a more streamlined and recyclable steel system by identifying a minimal set of representative grades per cluster.

To ensure that grade reduction did not compromise functional performance, mechanical property preservation was quantified explicitly at the cluster level. For each cluster, the minimum and maximum values of the target mechanical properties (UTS and elongation) were evaluated before and after reduction. A reduced grade set was accepted only if it fully preserved the original property range of the cluster. In addition, for every eliminated grade, at least one retained grade (or combination of retained grades) was required to fully cover its mechanical property range. This validation was applied independently within each cluster and for both reduction strategies, ensuring that redundancy was removed without excluding any achievable strength–ductility combinations present in the original dataset.As described in chapter 5, the reduction process is guided by a user-defined prioritisation strategy based on two perspectives: (1) minimisation of tramp elements, which is a conservative approach that favours grades with the lowest concentrations of copper and tin, and prioritising clean-input alloys suitable for high-purity applications. While such grades may offer superior performance in early cycles, they are less tolerant to tramp element accumulation and thus less robust in repeated recycling. This strategy aligns with traditional metallurgical thinking that seeks to minimise impurity thresholds to preserve input quality and minimise hot shortness or surface defects during processing (Bell et al., 2006; Owen et al., 1976), and (2) maximisation of tramp elements in the steel system, which this approach favours grades that already tolerate higher levels of copper and tin when other grades cover their property range. As discussed in section 2.1.3, these elements are unavoidably accumulated in steel recycling systems; thus, by retaining these grades, the framework supports closed-loop recycling, where scrap is returned to make the same or functionally equivalent products.

In this study, the tramp-tolerant strategy is the primary focus, as it reflects realistic future recycling challenges and aligns with circular economy goals. However, results from the tramp-avoidance strategy are also reported, to demonstrate the flexibility of the KMRP algorithm and validate that it can adapt to different user or industry objectives.

In both cases, the algorithm prioritises alloy designs with minimum dependency on critical raw materials when other grades could cover the property range. For critical raw materials, the selected elements were molybdenum, nickel, titanium, and vanadium, based on their economic value and supply risk. The following subsections present the results of the reduction process under each scenario.

7.2.3.1 Minimum Tramp Element (Tramp Element Avoidance)

In this scenario, the reduction process was applied to each cluster individually. The algorithm identified sets of overlapping grades, meaning grades that collectively covered the same property ranges within each category (carbon or stainless steel). From these sets, it selected the grade or grades with the lowest total tramp element content (copper and tin) to retain.

In cases where multiple overlapping candidates have equal tramp levels, the algorithm applied a secondary reduction filter by comparing the total content of CRMs (Mo, Ni, Ti, V), thus promoting resource and cost-efficient, low-risk steel alternatives.

The results showed that in cluster 0 the number of grades was reduced from 186 to 93, while fully preserving the UTS-elongation range, as confirmed by unchanged minimum and maximum values and the absence of new gaps in the original property distributions. Both carbon and stainless grades were retained, confirming that the algorithm avoided cross-type substitutions. In cluster 1, the reduction was modest, reflecting fewer overlaps. In cluster 2, no grades were reduced, as each of the nine grades contributed unique mechanical coverage. In cluster 3, the number of grades was reduced from 148 to 73. The algorithm identified significance redundancy in this group, where many steels shared similar UTS-elongation profiles. Redundant grades were eliminated by retaining only those with cleaner chemical compositions.

In conclusion, the minimum tramp element strategy successfully reduced grade numbers where overlaps existed, while preserving complete mechanical ranges. By prioritising steels with lower copper and tin contents, the algorithm ensured compatibility with recycling systems sensitive to tramp accumulation. Although effective in demonstrating the flexibility of the framework, this avoidance strategy is more aligned with traditional metallurgy. In the context of circular economy, the tramp-tolerant approach (presented in the following section) offers greater practical relevance.

7.2.3.2 Maximum Tramp Element Tolerance Strategy

Due to the increase in the use of scrap in steel production, controlling copper content in recycled steel becomes increasingly difficult. Rather than aiming to avoid copper, this strategy assumes that future recycling systems must be able to adapt to higher levels of tramp elements (Gao et al., 2025). As a result, the algorithm selected the grade or grades with the highest copper content from each overlapping set, based on the assumption that these grades are more

compatible with production environments that rely heavily on scrap material. In addition, this strategy is particularly suitable when the recycled steel is intended for applications that can tolerate higher levels of copper, such as construction products and reinforcing bars. In these cases, copper concentrations of up to 0.4 weight percent or more are considered acceptable (Daehn et al., 2017). The complete list of retained and eliminated grades under this strategy is available in the accompanying open-access dataset (Jalalian, 2026).

Cluster 0

The reduction process for cluster 0 was applied under the maximum tramp element strategy. As shown in Figure 7.6, the number of grades was reduced from 186 to 89 in this strategy. This reduction was performed automatically by the algorithm by prioritising grades with higher copper tolerance, while strictly enforcing full preservation of the original property coverage within the cluster. A grade was eliminated only if its complete mechanical property range was fully covered by at least one retained grade within the same cluster.

The resulting reduction indicates that 89 grades were sufficient to cover the required property range, allowing greater tolerance in tramp element levels while minimising critical raw material usage if existed.

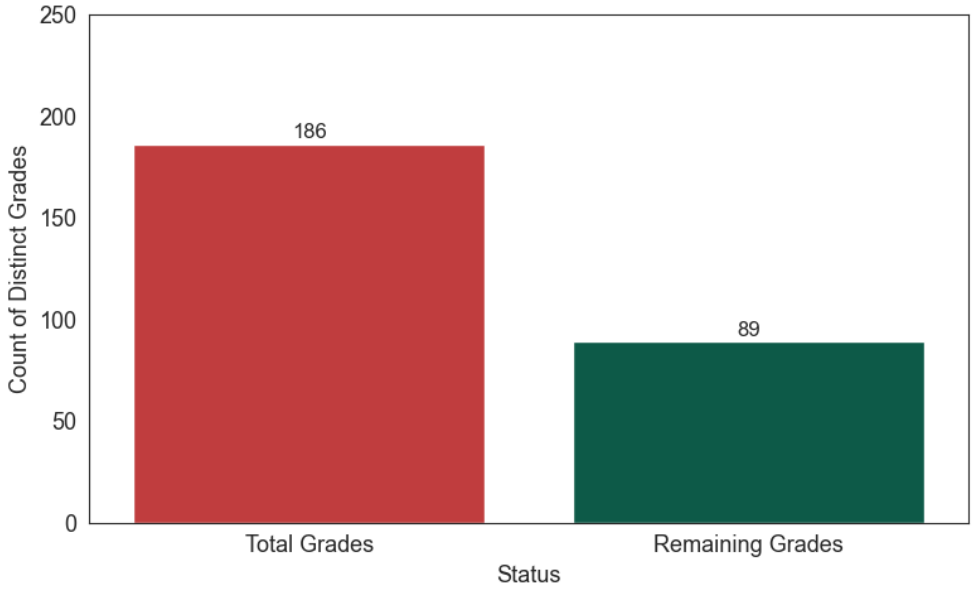


Figure 7.6: Count plot of steel grades before and after the reduction process under maximum tramp element strategy in cluster zero.

To illustrate how the algorithm functions under this strategy, grades with a UTS of 600 MPa and elongation of 23% were examined. Eight steel grades met these criteria included six stainless and two carbon steels. Table 7.4 shows their copper and critical raw material contents

(Mo, V, Ni, Ti). Among the stainless steels, UNS S44625 was retained over UNS S44800 and UNS S44700 due to its higher copper content (0.2%) and lower total critical material content. Within the carbon steels, ASTM A372 Grade B was selected over AISI 1020 (Carburised); as both grades had identical elemental content, either could have been retained based on property coverage.

The rationalisation outcomes are summarised in Table 7.4, which presents the elemental profiles of all grades in this mechanical window. The final retained set includes one stainless steel and one carbon steel, each representing its category while maintaining tolerance to tramp elements and minimising reliance on strategic alloying additions.

Table 7.4: Tramp element (Cu) and critical raw material (Mo, V, Ni, Ti) content for the status of the nine steel grades that cover a UTS of 600 MPa and Elongation of 23% in cluster zero using maximum tramp element strategy.

Cu_Max (wt.%)	Mo_Max (wt.%)	V_Max (wt.%)	Ni_Max (wt.%)	Ti_Max (wt.%)	UTS (MPa)	Elongation (%)	Type	Grade	Status
0	0	0	0	0	600	23	Stainless Steel	EN 1.4749	Eliminated
0.2	1.5	0	0.50	0	600	23	Stainless Steel	UNS S44625	Remaining
0.15	4.2	0	0.15	0	600	23	Stainless Steel	UNS S44700	Eliminated
0.2	4.2	0	2.50	0	600	23	Stainless Steel	UNS S44800	Eliminated
0	0	0	0.75	0	600	23	Stainless Steel	EN 1.4006 (Annealed)	Eliminated
0	0	0	0	0	600	23	Stainless Steel	EN 1.4024	Eliminated
0	0	0	0	0	600	23	Carbon Steel	AISI 1020 (Carburised)	Eliminated
0	0	0	0	0	600	23	Carbon Steel	ASTM A372 Grade B	Remaining

Cluster 1

The reduction process for cluster 1 followed the same approach. As shown in Figure 7.7, the number of steel grades was reduced from 91 to 71. The smaller reduction rate supports the

earlier observation (Section 7.2.2) that cluster 1 contains many grades that contribute unique combinations of UTS and elongation, thereby limiting opportunities for rationalisation.

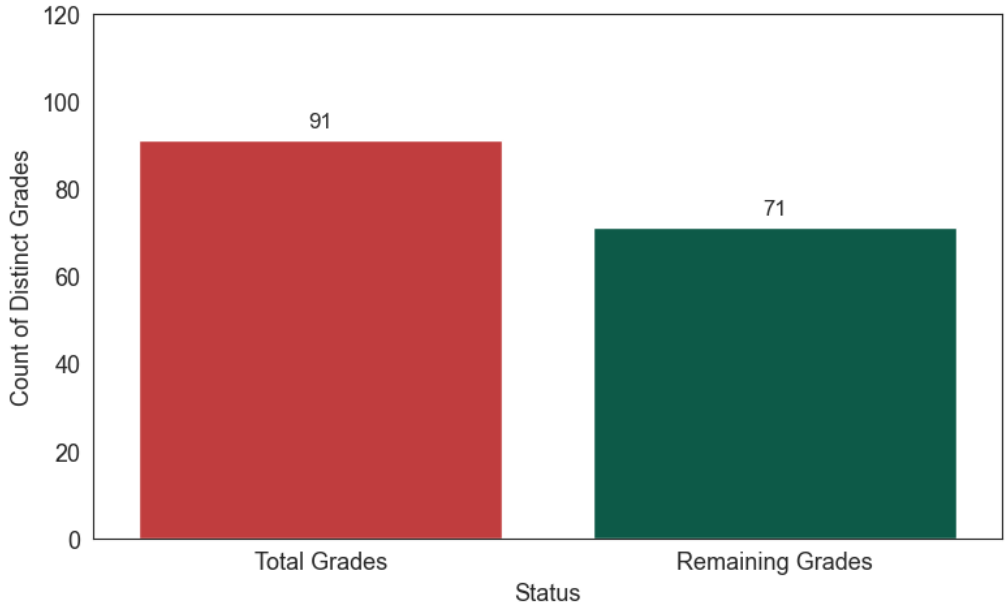


Figure 7.7: Count plot of steel grades before and after the elimination process under maximum tramp element strategy in cluster one.

To demonstrate the practical operation of the strategy in this cluster, steel grades with an elongation of 25% were extracted. In total, two grades met this condition, both of them were carbon steels. As summarised in Table 7.5, both grades provided overlapping mechanical performance, making them interchangeable in terms of UTS–elongation coverage. GB/T Q215 was retained because it had a higher copper content (0.3%). Although ASTM 1013 also covered the same mechanical property range and had slightly lower levels of critical raw materials, the primary objective of this strategy was to prioritise the grades with higher tramp element tolerance. Therefore, the algorithm selected the grade with the higher copper content, aligning with the principles of the maximum tramp element strategy.

Table 7.5: Tramp element (Cu) and critical raw material (Mo, V, Ni, Ti) content for the two steel grades status that cover a UTS between 880 MPa and Elongation 25% in cluster one using maximum tramp element strategy.

Cu_Max (wt.%)	Mo_Max (wt.%)	V_Max (wt.%)	Ni_Max (wt.%)	Ti_Max (wt.%)	UTS (MPa)	Elongation (%)	Type	Grade	Status
0.0	0.00	0.00	0.0	0.0	880	25	Carbon Steel	ASTM 1013	Eliminated
0.3	0.00	0.00	0.3	0.0	880	25	Carbon Steel	GB/T Q215	Remaining

Cluster 2

As previously discussed in Section 7.2.2, cluster 2 included steel grades with no overlap in mechanical properties, making them essential for preserving the full performance range of the group. Consequently, the application of the maximum tramp element strategy did not change the outcome of the reduction.

However, this outcome also highlighted a potential opportunity for further optimisation. By extending the dataset to incorporate additional grades used in high-strength applications such as hardened spring wires, medical instruments, or high-hardness cutlery, it is possible to introduce overlapping candidates in the future and reduce the number of grades within these applications.

Cluster 3

As shown in Figure 7.8, the number of steel grades was reduced from 148 to 57, confirming that more than half of the grades shared overlapping mechanical property ranges. All four carbon steel grades in this group were retained, as they contained distinct property characteristics with no internal overlap.

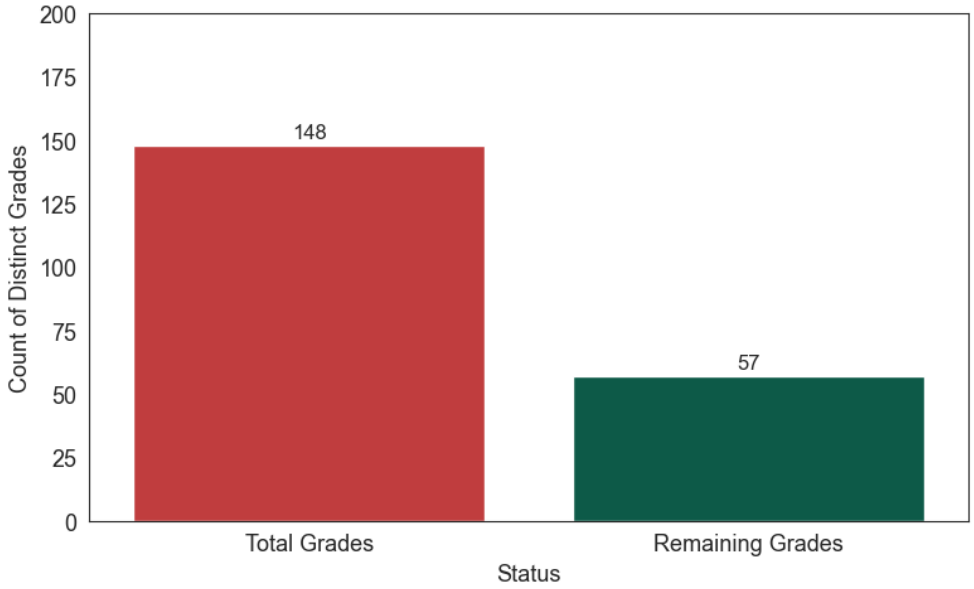


Figure 7.8: Count plot of steel grades before and after the reduction process in cluster three.

To better understand how the algorithm's decision-making performs on this cluster, steel grades with an elongation of 40% and a UTS of 610 MPa were selected. A total of four grades met these criteria. Table 7.6 lists these grades along with their tramp element (Cu) and critical raw material content (Mo, V, Ni, Ti). As shown in the table, EN 1.4570 (X6CrNiCuS18-9-2) was retained due to the highest copper content among the grades within this property range.

This result confirms that the reduction process effectively prioritises the grade with the highest tramp element content, consistent with the objectives of the maximum tramp element strategy.

Table 7.6: Tramp element (Cu) and critical raw material (Mo, V, Ni, Ti) content for the status of the six steel grades that cover a elongation 50% in cluster three.

Cu_ Max (wt.%)	Mo_ Max (wt.%)	V_Max (wt.%)	Ni_Max (wt.%)	Ti_Max (wt.%)	UTS (MPa)	Elongation (%)	Category	Grade	Status
1.0	0.0	0.0	10.0	0.0	610	40	Stainless Steel	EN 1.4305 (X8CrNiS18-9)	Eliminated
1.8	0.6	0.0	10.0	0.0	610	40	Stainless Steel	EN 1.4570 (X6CrNiCuS18-9-2)	Remaining
0.0	0.0	0.0	12.0	0.0	610	40	Stainless Steel	EN 1.4912 (X7CrNiNb18-10)	Eliminated
0.0	0.0	0.0	13.0	0.0	610	40	Stainless Steel	AISI 347LN (S34751)	Eliminated

In conclusion, the maximum tramp element strategy successfully retained grades by prioritising the higher copper tolerance. Besides that, the algorithm maintained full coverage of mechanical property ranges across all clusters and ensured low dependency on critical raw materials. This strategy offered an alternative reduction pathway for enhancing adaptability in future recycling systems.

7.3 Summary

This chapter presented the application of the KMRP algorithm to reclassify steel grades. The process involved three phases: identifying the optimal number of clusters, applying K-Means with PCA to group grades by mechanical properties, and reducing the number of grades within each cluster. Four clusters were identified, which not only aligned with mechanical performance similarities but also reflected traditional metallurgical categories, reinforcing the validity of the approach.

The reduction phase was carried out under two strategies. The minimum tramp element strategy prioritised grades with lower Cu and Sn content and reduced CRM dependence, while the maximum tramp element strategy favoured grades with higher tramp tolerance, supporting scrap-compatible recycling. Both strategies supported objectives 3 and 4 by demonstrating that

redundant grades can be systematically reduced while maintaining property coverage and enhancing sustainability.

Overall, the results confirm that the KMRP framework achieved the study's aim of developing an application-driven reclassification system that simplified grade variety, preserves performance, and supports recyclability. The next chapter interprets these outcomes in the context of mechanical properties and application requirements, and compares the K-Means clustering with the refined output produced by the KMRP algorithm. Although the algorithm developed in this study is capable of incorporating user preferences to guide the classification process, the following section reports only the results obtained under the maximum tramp element strategy, which is more aligned with the aim of this study.

Chapter 8: Interpreting the KMRP Results using Maximum Tramp Element Strategy

This section presents a comparative interpretation of the results obtained from applying the KMRP algorithm under the maximum tramp element strategy (section 7.2.3). While K-Means was used within Phase II of KMRP to group steel grades based on mechanical property similarity, the key enhancement of KMRP lies in Phase III which evaluates overlap within each cluster and selectively retains grades with lower dependency on CRMs and higher tolerance to tramp elements.

The analysis begins by quantifying the reduction in grade count within each generated cluster, focusing on how overlapping grades were identified and filtered. Grades showing redundant property ranges were reduced unless they demonstrated better recyclability characteristics and lower dependency on CRMs. Subsequently, each cluster is examined before and after the application of KMRP to assess changes in grade composition and to verify whether the original mechanical property ranges (UTS and elongation) were preserved in the context of traditional steel classifications (e.g., carbon vs. stainless).

Finally, the effectiveness of the maximum tramp element strategy is assessed in relation to sustainability objectives by analysing changes in tramp element content and critical raw material dependency in the proposed steel system.

8.1 Grade Reduction Comparison

This section evaluates the number of steel grades reduced in each cluster after applying the KMRP algorithm under the maximum tramp element strategy. This analysis provides evidence of how the algorithm systematically addressed the study's aim by simplifying the steel system.

Table 8.1 summarises the existing number of steel grades in the clusters before and after reduction under the maximum tramp element strategy. The most significant reduction occurred in cluster 3, where the number of grades decreased by 61.4%, followed by cluster 0 with a 52.1% decrease. These results suggest that both clusters contained a large number of overlapping grades, which enabled KMRP considerable simplification. In contrast, cluster 1 showed a minor reduction of 21.9%, indicating fewer overlaps and a greater proportion of unique grades. Cluster 2 showed no reduction, confirming that all grades in this cluster were required to maintain full coverage of the mechanical property range, with no detectable overlap.

By examining the results from the minimum tramp element strategy (appendix A) and comparing them with those from the maximum tramp element strategy, it is evident that the maximum strategy yielded a slightly higher overall reduction in the number of grades. This was especially noticeable in cluster 3, where the reduction was 61.4% compared to 50.7%, and in cluster 0, where it was 52.1% compared to 50%. These differences in reduction rate suggest that most steel standards have low tolerance to copper and high CRM dependency in the examined data, which led to the reduction of a larger number of grades in the maximum tramp element strategy.

Table 8.1: Comparison of the number of steel grades in each cluster before and after the reduction process using the maximum tramp element strategy.

Cluster	***Original grades (K-Means)***	***Remaining grades (KMRP)***	***% Reduced***
0	186	89	52.1
1	91	71	21.9
2	9	9	0
3	148	57	61.4

Moreover, Figure 8.1 demonstrates the distribution of steel grades by traditional classification (carbon and stainless) before and after applying the reduction process based on the maximum tramp element strategy in the KMRP algorithm. Figure 8.1(a) represents the clustering results from the K-Means phase of KMRP (before reduction), while Figure 8.1(b) shows the final outcome after applying the reduction process (Phase III).

The plots highlight a significant reduction in carbon steel grades in cluster 0, where almost 50 of the original grades were reduced. Similarly, around 91 stainless steel grades were reduced in cluster 3. This suggests a high level of redundancy and overlap in terms of selected properties among grades in these regions of the performance space. The effective reduction of grades in these clusters is particularly significant from a sustainability standpoint. In both clusters, the retained grades possessed higher tolerance to tramp elements (especially copper), which is a common contaminant in recycled scrap. The result revealed that copper-tolerant alternatives already exist for many conventional grades, and their identification through KMRP allowed for simplification without functional loss.

In addition, it is worth noting that the number of carbon steel grades in cluster 3 remained unchanged. As discussed in section 7.2.3, no overlap was detected in either UTS or elongation among these grades. For this reason, the algorithm retained all of them because each grade played a distinct role in fulfilling the mechanical performance range of this cluster.

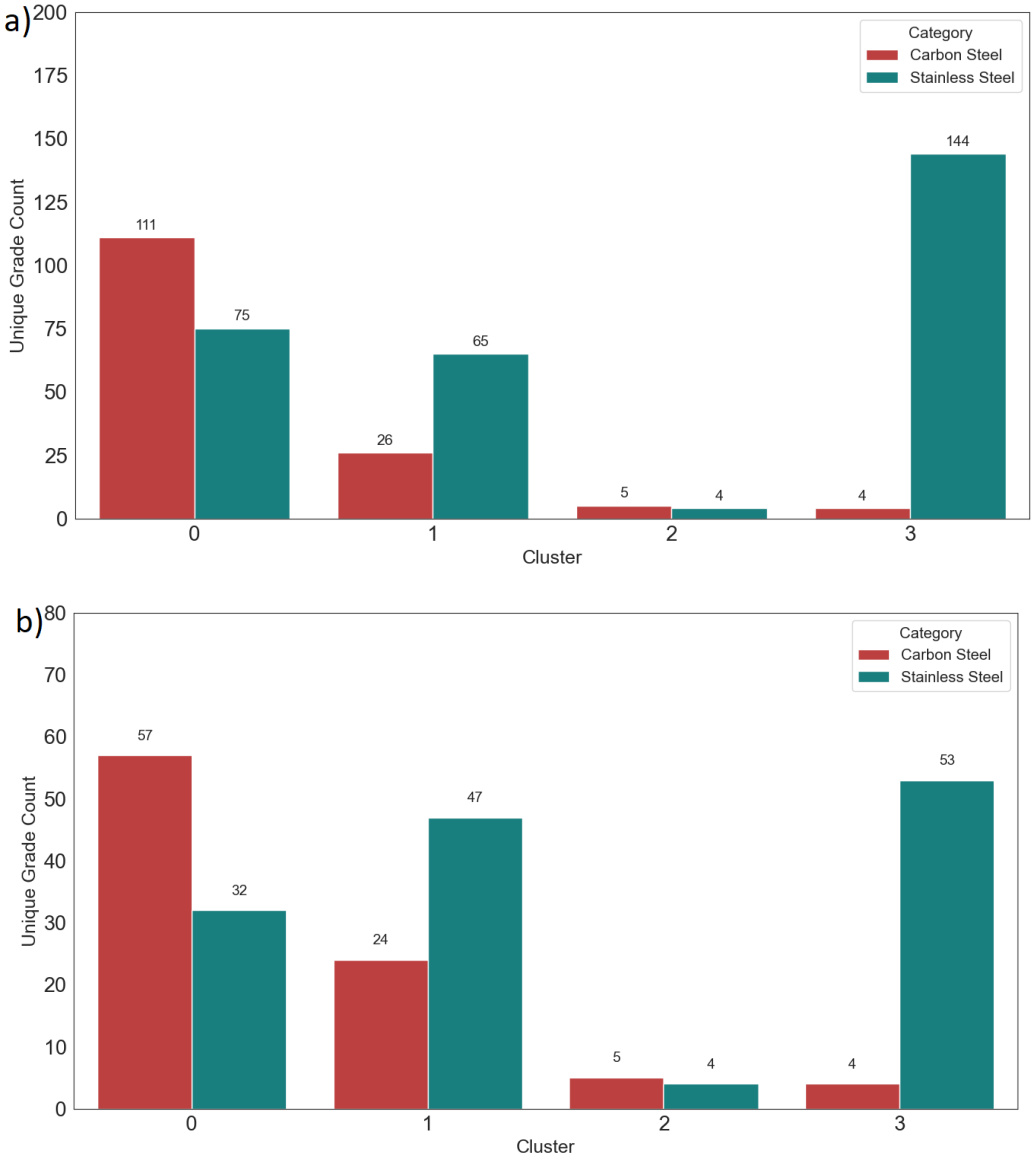


Figure 8.1: Distribution of steel grades across different cluster by categories before (K-Means result) and after (KMRP result) the reduction process using the maximum tramp element strategy.

To assess the broader implications of the KMRP algorithm, the output under the maximum tramp element strategy was compared against the conventional classification of steels (carbon and stainless families). Table 8.2 presents the number of grades before and after reduction, along with their mechanical property ranges (UTS and elongation). This comparison excluded similarity-based groupings to isolate the direct effect of KMRP on traditional classification.

The results showed that KMRP could reduce the total number of steel grades by 47.9%. The preserved UTS (120 to 2450 MPa) and elongation (2 to 55%) ranges confirm that the retained grades covered the entire original mechanical properties. This finding challenges the assumption embedded in conventional steel standardisation systems such as EN, ASTM, or AISI that each grade is uniquely required (Hall and Fekete, 2017).

By examining at traditional material categories, the current steel ecosystem did not adequately capture functional redundancy or recycling compatibility and is under-optimised. Grades grouped by compositional standards rather than mechanical behaviour, leading to an inflated variety with similar utility. Within this context, the maximum tramp element strategy reduced stainless steels by 52.8% and carbon steels by 38.4%. These results indicated that a significant proportion of stainless steel grades showed overlap in mechanical performance and nearly half of the stainless steel grades could be reduced without sacrificing the required property coverage.

In conclusion, the KMRP approach addressed the challenge of excessive steel grade diversity by favouring grades that are not only functionally robust but also more compatible with the constraints of secondary steel production.

Table 8.2: Overview of steel grade counts and mechanical property ranges before and after applying the KMRP algorithm using the maximum tramp element strategy, presented for carbon steels, stainless steels, and the full dataset.

Category	***Original grades***	***Original UTS range (MPa)***	***Original elongation range (%)***	***KMRP grades***	***KMRP UTS range (MPa)***	***KMRP elongation range (%)***	***% Reduced***
Carbon Steel	146	295 - 2450	7 – 41	90	295 - 2450	7 - 41	38.4
Stainless Steel	288	120 - 1970	2 – 55	136	120 - 1970	2 - 55	52.8
All Grades	434	120 - 2450	2 – 55	226	120 - 2450	2 - 55	47.9

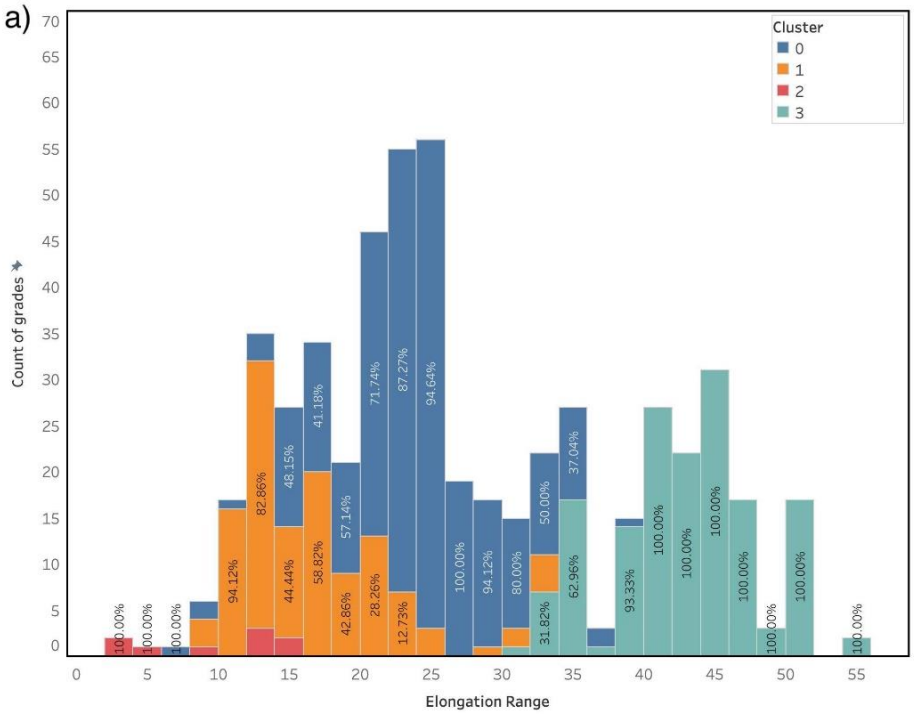
To better understand the changes introduced by KMRP within each cluster, the next section examines how the mechanical property coverage (UTS and Elongation) has been affected before and after reduction. It also identifies the unique characteristics and boundaries of each generated cluster.

8.2 Preservation of Mechanical Property Coverage

This section evaluates the characteristics of the generated clusters and whether the maximum tramp element strategy preserved the mechanical performance of the examined steel dataset after the reduction process. The focus is on UTS and elongation, which represent strength and ductility, respectively. Figures 8.2 and 8.3 illustrate the distributions of UTS and elongation across the four clusters generated by K-Means (phase II) and their corresponding reduced forms after applying Phase III of the KMRP algorithm.

As shown in Figure 8.2(a), cluster 3 contained most of the grades with high elongation values (above 30%) that is indicating excellent ductility and capacity for deformation. Figure 8.2(b) shows that although the number of high-elongation grades was reduced in phase III but the full elongation range was retained. As discussed in section 7.2.2, nearly all the grades in cluster 3 were stainless steels, with only four carbon steels. This aligns with findings from section 6.2.2, which showed that austenitic stainless steels and processing conditions including annealing or normalising, tend to higher elongation, especially due to their nickel content. Therefore, the grades in this cluster are particularly suitable for applications requiring superior ductility, corrosion resistance such as orthopaedic applications (Nouri and Wen, 2021), or high elongation carbon steels.

In contrast, cluster 0 concentrated in the moderate elongation range, where most of the grades fell between 20% and 30%. Clusters 1 and 2 represent grades with lower elongation values, typically below 15%. In the 20–22% interval, which initially had overlap between clusters 0 and 1, the distribution became more balanced after phase III. This suggests that overlap grades from cluster 1 were reduced, especially those with lower tolerance to tramp elements, and that cluster 0 retained the key representatives.



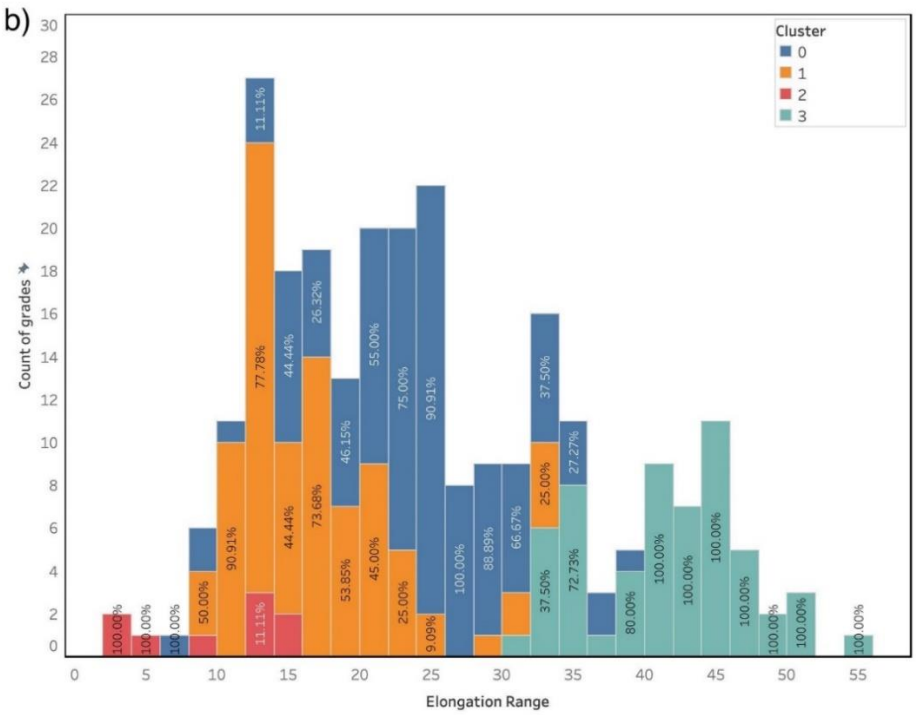
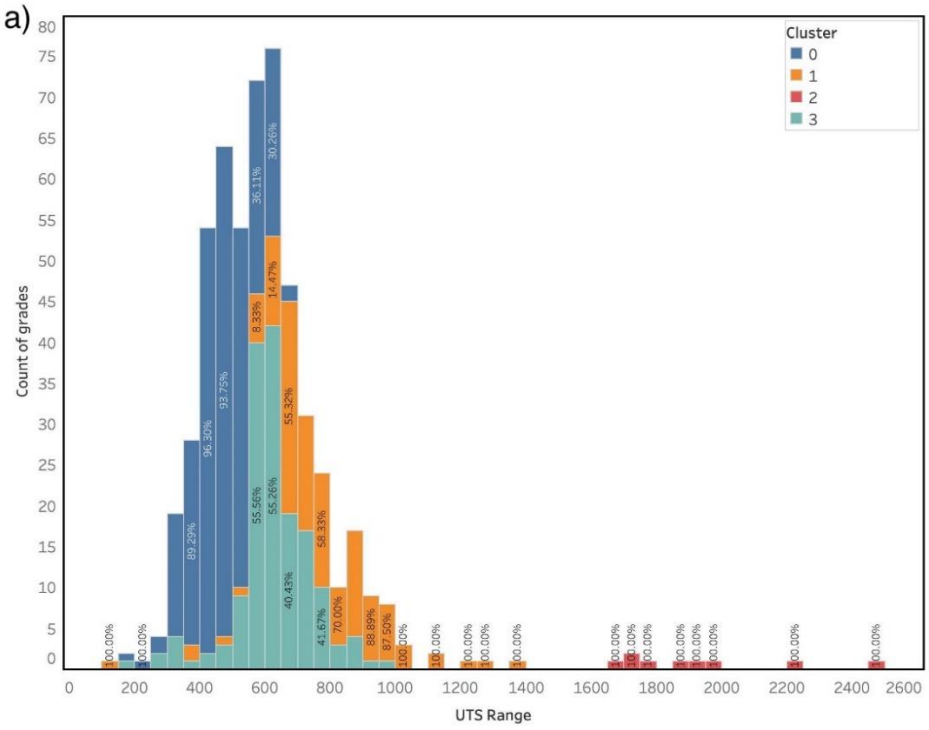


Figure 8.2: Distribution of the elongation (%) using (a) K-Means and (b) KMRP algorithms under maximum tramp element strategy.

To extend the analysis, the distribution of steel grades was also examined based on their UTS after applying the maximum tramp element strategy. Figure 8.3(a) demonstrates the UTS distribution across the four clusters, grouped in 50 MPa intervals. Prior to reduction, overlapping UTS ranges were present between clusters 0, 1, and 3. However, when combined with elongation data, clearer distinctions emerged. Cluster 3 contained grades with moderate UTS (550 to 650 MPa) and high elongation, while cluster 1 consisted of grades with medium to high UTS and lower elongation values (10 to 15%). Cluster 2 contained the highest strength steels, suggesting suitability for demanding structural or industrial applications. In contrast, cluster 0 included low-strength steels with moderate ductility, indicating potential applications in construction (e.g., buildings, bridges), pressure vessels, ships, and off-highway vehicles (Islam and Rashed, 2019). The stainless steel grades in this cluster also be suitable for automotive exhaust systems, trims, and hot water tanks (Nouri and Wen, 2021).

As shown in Figure 8.3(b), cluster 0 remained dominant in the lower to mid-strength range, particularly between 400 and 650 MPa. This reflects the cluster's continued association with grades that provide moderate strength and ductility. Cluster 1 had become more concentrated in the higher UTS intervals between 750 and 1000 MPa, a range that was previously dominated by grades from cluster 3. This shift suggested that the algorithm retained stronger grades in cluster 1 and reduced many grades from cluster 3 that contained lower levels of copper.

Cluster 3 still includes grades in the medium UTS range and retains its general profile, though with significant grade reduction, around 62 percent as observed in section 8.1. Cluster 2 continued to consist of grades with the highest UTS values, confirming its relevance for high-performance applications. Despite the reduction of the number of grades, the full UTS range remained covered across all clusters. These results indicated that the maximum tramp element strategy successfully reduced redundancy while maintaining original UTS and elongation.



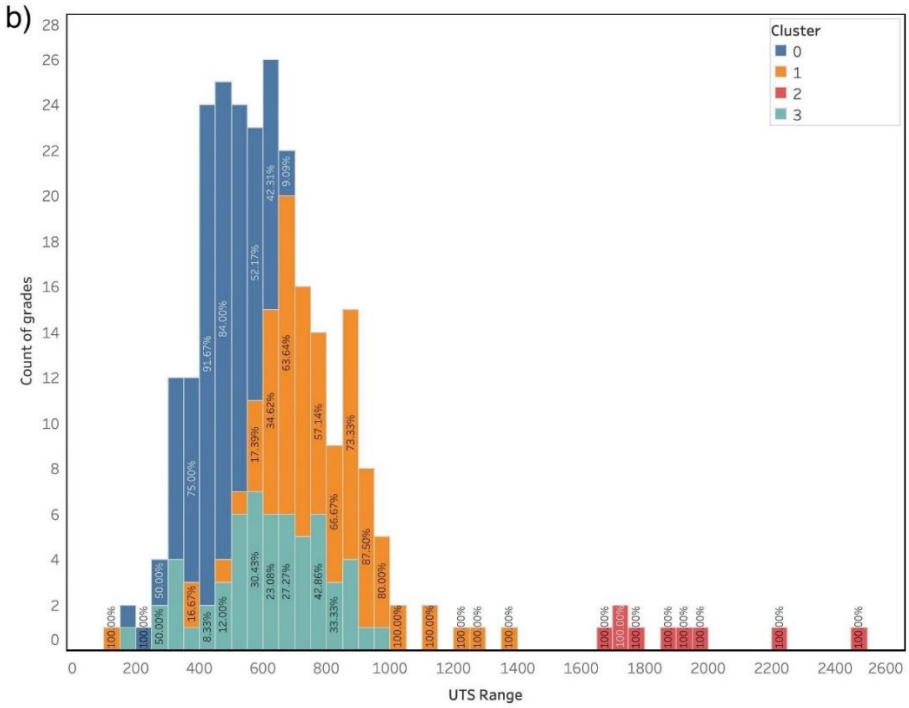


Figure 8.3: Distribution of the UTS (MPa) using (a) K-Means and (b) KMRP algorithms under maximum tramp element strategy.

From a metallurgical perspective, the KMRP outcomes are consistent with established correlations between composition, processing, microstructure, and mechanical response in carbon and stainless steels. In the low- to medium-strength region (cluster 0), the low-carbon steels differ primarily through modest variations in Mn, Si, and residual Cu. Following similar processing, these grades develop ferrite–pearlite microstructures with overlapping UTS-elongation combinations. Within this cluster, variations in Cu content do not fundamentally alter the ferrite–pearlite phase balance or the dominant strengthening mechanisms, as the Cu remains in solid solution or as fine-scale precipitates.

A similar redundancy is evident in cluster 3, where many austenitic stainless steels with closely comparable Cr–Ni–N levels and similar annealed conditions showed almost identical combinations of moderate UTS and high elongation. In this cluster, the KMRP algorithm retained a subset of grades such as EN 1.4570 (X6CrNiCuS18-9-2), which preserve the conventional austenitic microstructure but include deliberate Cu additions, while structurally similar low-Cu austenitic grades with overlapping strength–ductility behaviour were reduced. Because these substitutions always occur within the same austenitic family, and because modest variations in Cu at these levels do not change the underlying deformation mechanisms of the austenitic matrix (solid-solution strengthening with high work-hardening capacity governed by dislocation slip, and where relevant TWIP/TRIP), replacing low-residual grades

with Cu-containing equivalents that reproduce the same UTS-elongation range is metallurgically reasonable in this cluster.

By contrast, KMRP retained all grades in cluster 2 and all carbon-steel grades in cluster 3 because no other steels matched their combinations of very high strength and limited ductility. These groups correspond to high-carbon martensitic and/or heavily work-hardened steels in which alloy design and processing are tightly coupled to application-critical microstructures, such as quenched-and-tempered martensite. In these cases, simply increasing tramp tolerance or lowering CRM levels would require substantial changes in composition or heat treatment that would alter the microstructural basis of strength, so the grades were correctly identified as non-redundant within the examined property space.

Independent of cluster assignment, the broader effect of the KMRP algorithm was examined on the mechanical property distribution of traditionally classified steel categories (carbon and stainless). A summary of UTS and elongation coverage before and after KMRP reduction is provided in appendix A.

These results demonstrate that the streamlined steel grade system produced by the KMRP strategy successfully preserved the full mechanical property coverage of the traditional classification, using only a minimal subset of grades. This confirms the method's potential to enhance material efficiency while maintaining functional sufficiency. The following section examines the resulting shifts in tramp element concentrations and CRM dependencies following the application of KMRP.

8.3 Tramp and Critical Raw Materials Summary

This section presents the analysis of tramp elements and CRMs in the retained steel grades following application of the KMRP algorithm under the maximum tramp element strategy. The focus is on evaluating recyclability, environmental impact, and production cost by comparing the ranges of these elements before and after reduction. In addition, the effect of KMRP on the overall dataset is examined by assessing how tramp element and CRM distributions change in current classification system once redundant grades are eliminated.

8.3.1 Copper Content

As discussed in chapter 2, copper is identified as a main tramp element in steel recycling. While copper content below 0.1% are generally considered non-problematic, levels above this

threshold can lead to defects such as hot shortness during processing and are typically avoided when scrap is used (Daehn et al., 2017). The maximum tramp element strategy in this study aims to challenge this traditional paradigm of avoiding tramp elements in steel production by prioritising grades that can tolerate higher copper content without compromising mechanical performance (Mehta et al., 2025). This strategy aligns with the realities of secondary steel production, where tramp elements are often introduced unintentionally and are difficult to eliminate entirely. By selecting tolerant grades over less tolerant ones when they offer similar mechanical property coverage, the KMRP algorithm supports improved recyclability and material circularity. Although this study did not explicitly expand the copper tolerance range, the algorithm remains valuable in guiding decision-making toward robust, impurity-tolerant grades that preserve functional requirements while reducing grade complexity.

Table 8.3 summarises the distribution of copper content across each cluster generated by the KMRP system, based on K-Means clustering (Phase II) and after the subsequent reduction process (Phase III).

Table 8.3: Summary of copper content before and after KMRP reduction in generated clusters under maximum tramp element strategy.

Cluster	*Grades < 0.1% Before*	*Grades > 0.1% Before*	*Cu Range Before (%)*	*Cu Range After (%)*	*Grades < 0.1% After*	*Grades > 0.1% After*
0	141	45	0 – 1.3	0 – 1.3	65	24
1	87	4	0 – 3	0 – 3	67	4
2	9	0	0	0	9	0
3	126	22	0 – 3.6	0 – 3.6	47	10

In cluster 0, the copper content ranged from 0 to 1.3%. out of the 186 grades originally presented in this cluster, 141 contained below 0.1% copper (77 carbon, 64 stainless steels), and 45 contained above 0.1% copper (34 carbon, 11 stainless). After reduction, the number of low-copper grades decreased to 65 (41 carbon, 24 stainless), and grades above 0.1% dropped to 24 (16 carbon, 8 stainless). These patterns indicate that many of the removed grades were low-Cu steels whose UTS-elongation combinations were already represented by other low-Cu grades, whereas only a smaller subset were replaced by more Cu-tolerant alternatives that offered comparable mechanical performance while also reducing overall dependence on critical alloying elements.

In cluster 1, copper content ranged from 0 to below 3%. Initially, 87 grades had below 0.1% copper (25 carbon, 62 stainless), while only 4 grades had above 0.1% (3 stainless, 1 carbon).

After reduction, 67 grades remained with below 0.1% copper (23 carbon, 44 stainless), and the same four higher-copper grades were retained. The minimal reduction in carbon steel grades suggests that most were essential for full mechanical property coverage.

In cluster 2, all nine grades were retained, as no property overlap was found (section 7.2.3). All grades had copper content, so no changes were made.

In cluster 3, the copper content ranged from 0 to just below 3.6%. Before reduction, 126 grades had below 0.1% copper (all carbon steels had zero copper), while 22 stainless steels exceeded this threshold. After KMRP, low-copper grades were reduced to 47, while high-copper grades dropped to 10. In addition, all carbon steels were retained due to their unique property coverage, while redundant austenitic stainless steels were reduced.

These outcomes demonstrate that although the copper content range remained unchanged across all clusters, the number of low-copper grades was significantly reduced. This reduction occurred because many of these grades, especially those in cluster 3 and 0, shared similar strength and ductility characteristics with those that had higher copper tolerance or lower dependency on CRMs.

In addition, to better understand the overall impact of the KMRP algorithm on copper content, Table 8.4 presents the distribution of grades above and below the 0.1% copper threshold, based on the traditional carbon and stainless steel classification. In the original dataset, 71 grades exceeded the 0.1% copper threshold (containing 36 stainless steels and 35 carbon steels), while the remaining 363 grades fell within the traditionally acceptable limit. After applying the KMRP algorithm, the number of low-copper carbon steel grades ($\leq 0.1\%$) was reduced from 111 to 73 (34% reduction), while in stainless steels, it dropped from 252 to 115 (54% reduction). This reduction rate, particularly in stainless steels, suggests that the traditional classification contained significant redundancy, where more copper-tolerant alternatives could replace many low-copper grades without compromising mechanical performance or application suitability.

These results confirm that the KMRP algorithm effectively reduced the number of grades by prioritising those with higher copper content where overlaps in mechanical properties were identified.

Table 8.4: Comparison of copper content in the full dataset before and after applying the KMRP algorithm under maximum tramp element strategy.

Dataset	*Copper Content Range (%)*	*Carbon steels below 0.1%*	*Carbon steels above 0.1%*	*Stainless steels below 0.1%*	*Stainless steels above 0.1%*	*Total Grades Below 0.1%*	*Total Grades Above 0.1%*
Original Data	0 - 4	111	35	252	36	363	71
After KMRP	0 - 4	73	17	115	21	188	38

8.3.2 Molybdenum Content

Although the maximum tramp element strategy prioritises grades with higher tolerance to tramp elements, this section also considers molybdenum as a CRM identified by the European Union (Blengini et al., 2020). Molybdenum plays an important role in improving properties such as hardenability, strength, toughness, and corrosion resistance. Kishor et al., (2025) reported that including 2 to 3% molybdenum in stainless steel significantly improves pitting and corrosion resistance. However, high concentrations can lead to increased material cost and resource concerns due to its limited availability on earth. Therefore, it is beneficial to minimise its usage where possible without compromising performance.

Table 8.5 summarises the distribution of molybdenum across the performance-based clusters generated during phase II (K-Means clustering) and after applying the phase III (reduction process) of the KMRP algorithm. A threshold of 0.1% was used to distinguish low and high molybdenum content, consistent with conventional alloying practices, where molybdenum is typically present in the range of 0.1% to 0.3% (Uranga et al., 2020).

Table 8.5: Summary of molybdenum content before and after KMRP reduction in generated clusters under maximum tramp element strategy.

Cluster	*Grades < 0.1% Before*	*Grades > 0.1% Before*	*Mo Range Before (%)*	*Mo Range After (%)*	*Grades < 0.1% After*	*Grades > 0.1% After*
0	156	30	0 – 4.3	0 – 4.3	73	16
1	51	40	0 – 4.6	0 – 3.1	43	28
2	5	4	0 – 0.8	0 – 0.8	5	4
3	94	54	0 – 7	0 – 4.1	34	29

Cluster 0 included 186 grades, where 156 contained molybdenum below 0.1% (106 carbon steels, 50 stainless steels), and 30 grades contained molybdenum above 0.1% (5 carbon steels, 25 stainless steels). After reduction, the number of low-Mo grades decreased to 73 (54 carbon

and 19 stainless), and the high-Mo subset declined to 16 (3 carbon and 13 stainless). The overall molybdenum range remained unchanged (up to 4.3%). By changing the reduction process to the minimum tramp element strategy, the molybdenum range in this cluster narrowed. However, by considering the tolerance to tramp element strategy, this range was preserved due to the retention of a high-molybdenum stainless steel grade that also demonstrated high copper tolerance. Nevertheless, the number of grades in the upper molybdenum band (4.0 to 4.3%) fell from three to one. Overall, this indicates that more than half of the high-molybdenum grades shared similar mechanical properties, suggesting that CRM content could be significantly reduced without affecting the mechanical property coverage of either steel category (section 8.2).

In cluster 1, 51 grades contained molybdenum below 0.1% (25 carbon, 26 stainless), and 40 contained molybdenum above 0.1% (39 stainless, one carbon). After reduction, the low-Mo subset decreased to 43 (23 carbon, 20 stainless), and high-Mo grades to 28 (27 stainless, one carbon). The molybdenum range narrowed from 0 to 4.6% to 0 to 3.1% as both UNS S44635 and UNS S44660 have been reduced due to the overlap of their properties with those of other stainless steels and higher dependency on their Mo (3-4%). These findings suggest that, with molybdenum at up to 3.1%, the property range of this cluster is still covered. The carbon steel grades in this cluster remained unaffected, indicating that internal redundancy was higher within stainless steels, which enabled greater CRM optimisation.

In addition, cluster 2 contained 9 grades where 5 high-carbon steels with molybdenum below 0.1%, and 4 stainless steels with molybdenum above 0.1%. As mentioned in section 7.2.3, no reduction occurred due to the absence of overlapping mechanical properties in this cluster. As a result, all nine grades, including those stainless steels with high molybdenum content is necessary to preserve the mechanical property coverage in this cluster.

However, cluster 3 offered the clear example of CRM minimisation. This cluster initially included 54 grades with molybdenum content above 0.1%, all of which were austenitic stainless steels. After applying the reduction process, this number dropped to 29, and the molybdenum range decreased from 7% to 4.1%. This suggests that many high- molybdenum stainless steel grades were interchangeable with alternatives that contained higher tramp element tolerance while still covering the same mechanical property range. In addition, all four carbon steels in this cluster were retained due to their unique combination of elongation and moderate UTS within this category, as discussed in section 7.2.3. Overall, the number of high-

molybdenum stainless steel grades was reduced by nearly 50%, while the cluster's full mechanical property coverage was preserved.

In addition, Table 8.6 summarises the overall effect of the KMRP algorithm on molybdenum content within the context of traditional steel classification.

Table 8.6: Comparison of molybdenum content in the full dataset before and after applying the KMRP algorithm under maximum tramp element strategy.

Dataset	*Molybdenum Content Range (%)*	*Carbon steels Below 0.1%*	*Carbon steels Above 0.1%*	*Stainless steels below 0.1%*	*Stainless steels above 0.1%*	*Total Grades Below 0.1%*	*Total Grades Above 0.1%*
Original Data	0 – 7	140	6	166	122	306	128
After KMRP	0 – 4.3	86	4	67	69	153	73

In the original dataset, 128 grades had molybdenum content above 0.1 percent, while 306 grades fell below this threshold. After applying the KMRP algorithm under the maximum tramp element strategy, the number of grades exceeding 0.1 percent was reduced to 73. Furthermore, the overall molybdenum content range was narrowed from 0-7% to 0-5%. In terms of the steel grades based on their category, the number of high-molybdenum stainless steel grades decreased from 122 to 69. In contrast, 4 out of 6 high-molybdenum carbon steel. While two grades were reduced due to overlapping mechanical properties with lower-molybdenum alternatives, the remaining four lacked sufficient overlap and were therefore preserved to maintain the full performance range.

These results demonstrated that the KMRP algorithm, despite of enabling reclassification of steels with smallest number of grades, it could also reduce the range of molybdenum content as a secondary optimisation criterion. The narrowing of the overall molybdenum range (from 7% to 4.3%) and the 43% reduction in high-molybdenum grades support the method's ability to align steel rationalisation with broader sustainability goals.

8.3.3 Vanadium Content

This section evaluates the distribution of vanadium, a critical raw material as identified by both the European Union and the United States (Jammulamadaka and Pisupati, 2023). Vanadium contributes to improving several mechanical properties of steel, including hardness, strength, toughness, wear resistance, and fatigue resistance (Yang et al., 2021). However, high concentrations of vanadium can significantly increase material costs and raise concerns due to

its limited natural availability. Aside from its limited availability, vanadium also raises serious environmental and health concerns. Studies have identified vanadium as one of the most concentrated trace elements released into the atmosphere by human activity, and it is also one of the most increased in rivers worldwide (Yang et al., 2022). Therefore, the World Health Organisation (WHO) has classified vanadium as highly toxic, with known harmful effects on both people and the environment (Organization, 1984).

In this context, vanadium content was evaluated to assess its overall presence within the proposed application-driven classification (Table 8.7). This analysis examines whether the KMRP algorithm, primarily focused on tramp element tolerance, also contributed to a reduction in vanadium-containing grades and a narrowing of the vanadium content range within traditional steel classifications.

Table 8.7: Summary of vanadium content before and after KMRP reduction in generated clusters under maximum tramp element strategy.

Cluster	***Grades < 0.01% Before***	***Grades > 0.01% Before***	***V Range Before (%)***	***V Range After (%)***	***Grades < 0.01% After***	***Grades > 0.01% After***
0	163	23	0 – 0.16	0 – 0.08	83	5
1	80	11	0 – 1.6	0 – 1.6	63	8
2	9	0	0	0	9	0
3	145	3	0 – 0.88	0 – 0.88	54	3

Cluster 0 initially included 186 grades that 163 grades were vanadium-free (88 carbon steels, 75 stainless steels), and 23 carbon steels grades contained vanadium content. After reduction, the vanadium-free grades dropped to 83 (51 carbon, 32 stainless), while high-vanadium grades declined to 5 (all carbon steels). The fact that only 5 out of 23 high-vanadium grades were retained implies that vanadium was not essential for achieving the required mechanical behaviour of carbon steel in this cluster. As a result, the overall vanadium range narrowed from 0-0.16% to 0-0.08%, indicating that the mechanical property range of this cluster could still be fully represented within a lower vanadium content threshold.

Cluster 1 had a broader vanadium range (0 to 1.6%) before reduction, with 80 vanadium-free grades (25 carbon, 55 stainless) and 11 vanadium-containing grades (10 stainless, one carbon). After reduction, the number of vanadium-free grades decreased to 63 (23 carbon, 40 stainless), and vanadium-containing grades to 8. Although three high-vanadium stainless steel grades were reduced due to overlapping mechanical properties, the upper vanadium threshold of 1.6% of this cluster was retained in the final set. This outcome illustrates a trade-off between tramp

element tolerance and the minimisation of CRMs. In certain cases, high-vanadium grades were retained not for their resource efficiency, but because they provided essential mechanical property coverage and showed high tolerance to tramp elements. These results suggest that optimal recyclability cannot always be achieved by minimising critical raw materials alone, but requires balancing multiple sustainability and performance constraints.

Cluster 2 as discussed in section 7.2.3, showed no property overlap and no reduction occurred. Further analysis revealed that none of the grades in this cluster contained vanadium.

In addition, cluster 3 showed a narrow vanadium range (0-0.88%) across 148 grades, of which 145 were vanadium-free. In addition, all four carbon steels in this cluster contained no vanadium. Only three stainless steel grades contained vanadium, and all were retained after the reduction process due to their higher tolerance to tramp elements.

To evaluate the broader effect of the KMRP algorithm on vanadium content, an analysis was conducted on the entire traditional steel classification, independent of the generated clusters. Table 8.8 summarises the vanadium content range and the number of grades falling below and above the 0.01% threshold, which serves as a marker distinguishing vanadium-free from vanadium-containing steels. Among the 434 steel grades in the dataset, only 37 contained vanadium (24 carbon steels and 13 stainless steels).

Table 8.8: Comparison of vanadium content in the full dataset before and after applying the KMRP algorithm under maximum tramp element strategy.

Dataset	***Vanadium Content Range (%)***	***Carbon steels below 0.01%***	***Carbon steels above 0.01 %***	***Stainless steels below 0.01%***	***Stainless steels above 0.01%***	***Grades Below 0.01%***	***Grades Above 0.01%***
Original Data	0 – 1.6	122	24	275	13	397	37
After KMRP	0 – 1.6	83	7	50	10	133	17

Despite the algorithm's primary focus on maximising tramp element tolerance, it also led to a considerable reduction in vanadium-containing grades. Under the traditional steel classification, 37 grades, including 24 carbon steels and 13 stainless steels, were implicitly considered necessary to preserve the full mechanical property range while accommodating vanadium content. However, the proposed application-driven classification approach demonstrated that only 17 grades were sufficient to retain this coverage. In addition, the reduction was more noticeable in carbon steels, where the number of high-vanadium grades

was reduced from 24 to 7. This outcome suggests that many vanadium-containing carbon steels in the traditional system were redundant, as alternatives could match their mechanical behaviour with lower or zero vanadium content.

In conclusion, the KMRP algorithm demonstrated promising results in reclassifying the existing steel grade into a significantly smaller and more efficient set. By prioritising grades with higher tolerance to tramp elements (where available), the approach supports enhanced recyclability. At the same time, in clusters where mechanically equivalent alternatives existed, many high-Mo and high-V grades were removed, leading to a substantial narrowing of the overall CRM content ranges; in other clusters, selected high-CRM grades were retained because they provided essential property coverage or combined unique mechanical behaviour with high tramp-element tolerance. Overall, the framework therefore enables reductions in CRM use and grade complexity while preserving the essential mechanical property ranges required for application performance, but it also makes explicit where recyclability and CRM minimisation reinforce each other and where they must be balanced as competing design objectives.

The next section examines the characteristics of the final steel groups generated by KMRP and evaluates the robustness of the proposed classification system by assessing its generalisability to new steel grades.

8.4 SHAP Analysis

To assess the interpretability and structural coherence of the clusters generated by the KMRP algorithm, SHAP analysis was conducted. The clustering was initially performed using only mechanical properties, and the subsequent reduction stage incorporated recyclability/sustainability-oriented criteria such as tramp element tolerance and CRM minimisation. SHAP was used to identify the dominant chemical and processing features associated with each cluster. This clarified the metallurgical basis of the groupings and confirmed that the clusters were not only statistically coherent but also meaningful in terms of material characteristics. In addition, the analysis provided guidance on how new or emerging steel grades could be positioned within the proposed classification framework, supporting its generalisability and practical application.

A RF classifier was trained to predict cluster membership from chemical composition and processing features, achieving an F1 score of 0.81 and an accuracy of 0.79. This demonstrated

that the mechanically defined clusters are consistently recoverable from chemistry and processing variables, confirming their structural coherence. To improve interpretation, SHAP values were calculated to find the most significant features that influenced cluster assignment. This analysis highlighted the key chemical elements, processing parameters, and traditional steel types that defined each group. It effectively changed the KMRP framework from a black-box clustering model to a clear and understandable alloy classification system.

Figure 8.4(a) shows the SHAP result for cluster 0. Cluster 0 was predominantly composed of ferritic stainless steels, as evidenced by strong negative SHAP values for the Type_Austenitic and Type_Martensitic indicators. These indicators clearly defined a group that was compositionally and structurally distinct from both austenitic and martensitic steels. In addition, carbon, nickel, and chromium all showed negative SHAP values, suggesting that low concentrations of these elements significantly contributed to group membership. This finding aligned with the histogram analysis (Fig. 7.5), which confirmed low carbon levels across all grades in this cluster. The copper-related features (Cu_Min and Cu_Max) did not appear among the top SHAP contributors. This suggested that, despite the presence of many carbon steels with higher copper content, the overall copper range in this group remained lower than in other clusters (section 8.3.1), and thus had minimal impact on the group's classification. In contrast, Mo_Min and Mo_Max appeared with negative SHAP values, indicating that grades with higher molybdenum content were less likely to belong to cluster 0. This finding suggested that molybdenum, despite its potential to enhance strength (Feng and Gs, 2024), was not essential in this group, and similar mechanical properties could be achieved through other alloying elements or processing conditions. This is consistent with the ferritic stainless steel, where solid-solution strengthening from Cr and other ferrite stabilisers dominates and high Mo additions are not required to reach the moderate strength–high ductility balance observed in this cluster (Kolotyrkin et al., 1994).

Figure 8.4(b) presented the SHAP summary plot for cluster 1. It was dominated by medium-carbon steels and martensitic stainless steels, as indicated by the positive SHAP values for C_Min, C_Max, and Type_Martensitic. However, the main difference in this strategy was the appearance of Cu_Min among the contributing features, with a negative SHAP value, indicating that a lower copper content increased the likelihood of a grade being assigned to this cluster. This was consistent with the group's composition, which showed that only three grades exceeded the 0.1% copper threshold, while the remaining grades remained below this limit. The absence of copper as a positive contributor suggested that for applications requiring

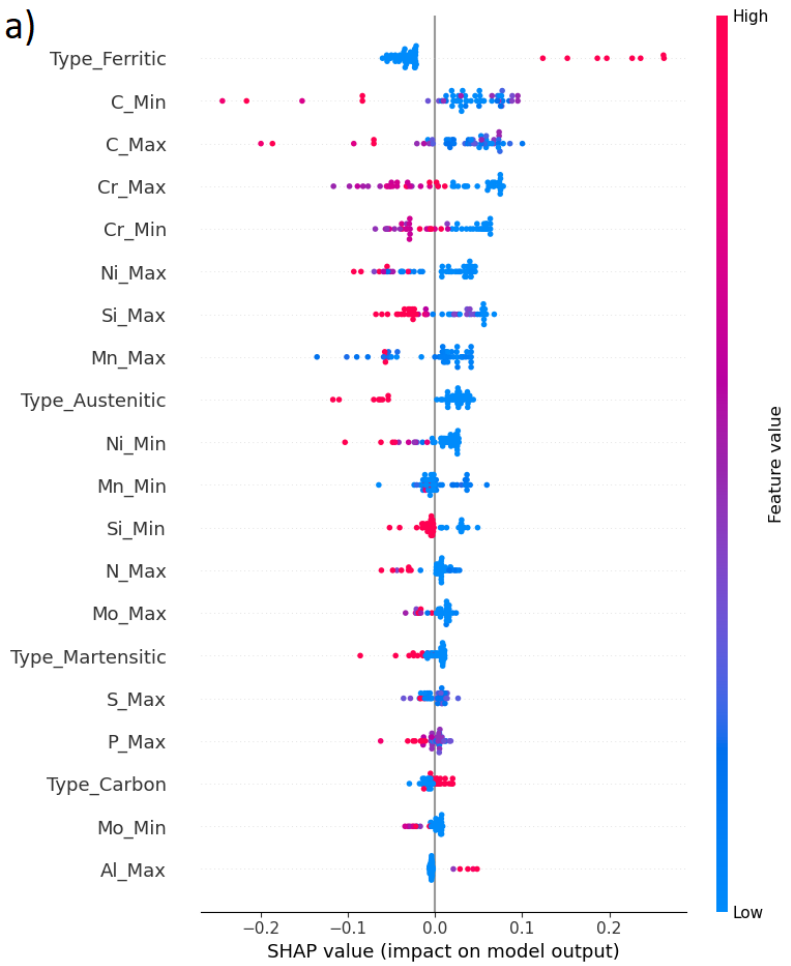
medium-to-high UTS and low elongation, copper was not a critical alloying element. Instead, other compositional and microstructural features dominated performance contributions in this regime. Here, the large positive SHAP contributions from carbon are in line with its dominant role in martensitic strengthening via increased carbon-in-solution and carbide formation, which raise hardness and UTS at the expense of ductility (Harwarth et al., 2021).

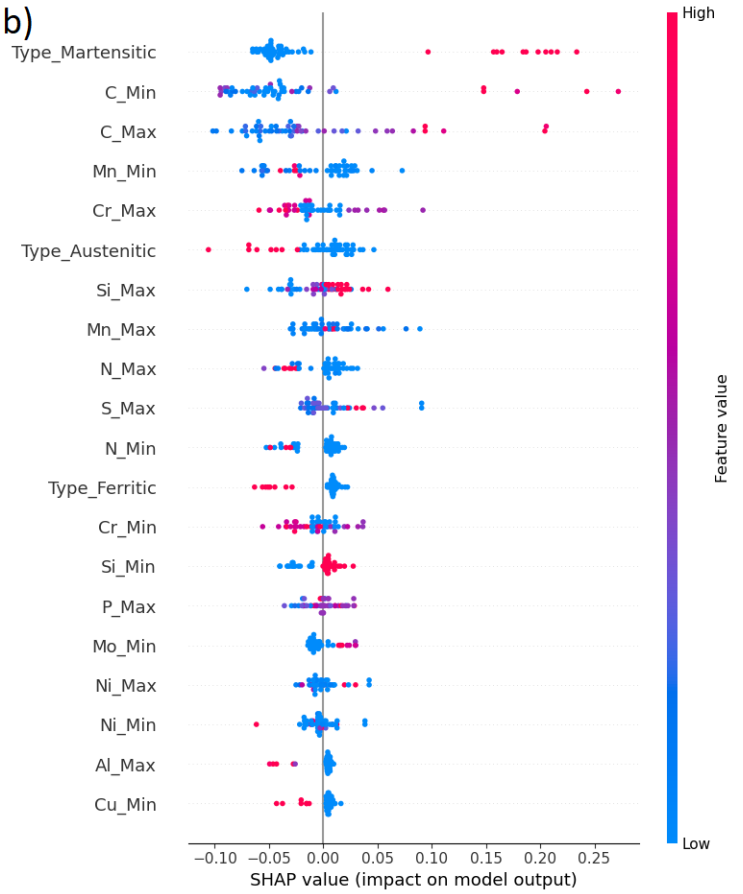
Figure 8.4(c) displayed the SHAP summary plot for cluster 2, where the top-ranked features included C_Max, Si_Min, and C_Min, indicating that both the upper and lower contents of carbon and silicon played a role in defining cluster membership. This implied that the steels in this cluster were generally high-carbon steels, and this aligned with the Type_Carbon indicator, which showed positive in this group. The high SHAP scores for C and Si are therefore consistent with their known roles in increasing dislocation density and promoting carbide and secondary phase formation, which explains the very high UTS and reduced elongation observed for these grades. Additionally, stainless steels present in this group were exclusively of the martensitic type, as indicated by the positive SHAP values for Type_Martensitic. These findings aligned with the metallurgical identity of the cluster, which was dominated by high-strength, low-ductility steels (Fig. 8.2 and 8.3). The association between increased carbon and silicon content and elevated UTS was well established in the literature (Serajzadeh and Karimi Taheri, 2002), confirming the validity of this cluster's structural coherence.

Figure 8.4(d) presented the SHAP summary plot for cluster 3. This cluster was clearly defined by the Type_Austenitic feature, along with high contributions from Ni_Min, Ni_Max, N_Max, and Cr_Max, which were all typical alloying elements in austenitic stainless steels. As discussed earlier (section 8.2), grades in this group showed the highest elongation values across all clusters, which was consistent with the known ductile behaviour of austenitic steels. A key difference in this SHAP result was the appearance of Cu_Min and Cu_Max as important features, both showing positive SHAP values. This meant that grades with higher copper content were more likely to be placed in cluster 3, supporting the goal of including tramp-tolerant grades under this strategy. This also suggested that the copper content had become more influential in defining the group, partly due to the imbalance in the original dataset, where high-copper steels were underrepresented compared to their low-copper counterparts. As the KMRP algorithm prioritised tramp element tolerance during reduction, it disproportionately eliminated low-copper grades, which leaded for boosting the representation and influence of high-copper grades within the cluster. In austenitic grades, the strong positive SHAP contributions from Ni, N and Cu mirror their metallurgical functions; Ni and N stabilise the

austenite phase and enhance work-hardening capacity (Huang et al., 2022), while Cu can contribute to precipitation strengthening (Huang et al., 2022), jointly underpinning the high ductility and adequate strength that characterise this cluster.

At the same time, molybdenum (Mo_Min and Mo_Max) continued to show a negative impact on group membership, suggesting that grades with higher molybdenum content were still deprioritised, even under the high tramp element constraints. Metallurgically, this is consistent with the fact that Mo is primarily introduced to enhance localised corrosion resistance (Sugimoto and Sawada, 1976), while in the austenitic grades in this cluster the required corrosion performance is largely provided by higher Cr–Ni–N levels (with some contribution from Cu), and the dominant design drivers are the strength–ductility combinations captured by these elements rather than by Mo. This highlighted the continued influence of KMRP's CRM minimisation objective.




b)
Type_Martensitic
C_Min
C_Max
Mn_Min
Cr_Max
Type_Austenitic
Si_Max
Mn_Max
N_Max
S_Max
N_Min
Type_Ferritic
Cr_Min
Si_Min
P_Max
Mo_Min
Ni_Max
Ni_Min
Al_Max
Cu_Min
−0.10 −0.05 0.00 0.05 0.10 0.15 0.20 0.25
SHAP value (impact on model output)
High
Feature value
Low

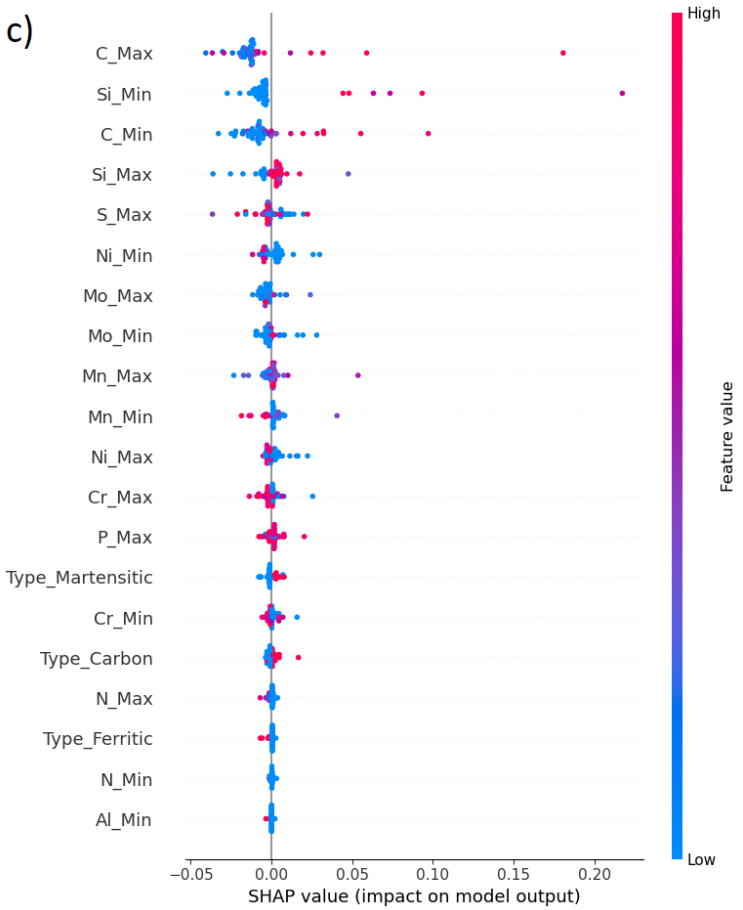

c)
C_Max
Si_Min
C_Min
Si_Max
S_Max
Ni_Min
Mo_Max
Mo_Min
Mn_Max
Mn_Min
Ni_Max
Cr_Max
P_Max
Type_Martensitic
Cr_Min
Type_Carbon
N_Max
Type_Ferritic
N_Min
Al_Min
−0.05 0.00 0.05 0.10 0.15 0.20
SHAP value (impact on model output)
High
Feature value
Low

Figure 8.4: SHAP summary plots for (a) Cluster zero, (b) Cluster one, (c) Cluster two, and (d) Cluster three, generated by KMRP algorithm under the maximum tramp element strategy. Positive SHAP values indicate that a feature increases the predicted cluster, while negative values indicate a decrease; the magnitude reflects the strength of that influence. The colour scale denotes the underlying feature value, with red indicating high values and blue indicating low values.

Overall, SHAP analysis strengthened the scientific credibility of the application-driven KMRP framework by validating that retained grades were not only mechanically representative but also chemically coherent within each cluster. It revealed how specific alloying elements particularly tramp elements and CRMs influenced group membership, confirming that the rationalisation process aligned with sustainability objectives. By revealing the dominant features that define each group, the SHAP analysis facilitated future decisions on where newly developed or recycled grades should be assigned to the proposed system.

8.5 Comparison with Industrial Steel Simplification Initiatives

The rationalised grade sets produced by the KMRP methodology are consistent with the broader industrial objective of reducing material complexity, but they differ fundamentally from existing simplification initiatives in both scope and approach. In several industrial sectors, most notably automotive manufacturing, simplification strategies have primarily focused on chemistry-driven unification. Concepts such as UniSteel propose replacing multiple body-in-

white steel grades with a single lean base chemistry that is differentiated through controlled processing routes to deliver a wide range of strength–ductility combinations (Lu et al., 2025; Raabe, 2023). These studies demonstrate that substantial reductions in alloy variety are technically feasible and can simplify forming, joining, and end-of-life recycling within a narrowly defined application domain.

Such initiatives, however, are inherently forward-looking alloy-design strategies. They introduce new material concepts and are optimised for specific product families or manufacturing environments. By contrast, the KMRP framework operates within the existing landscape of standardised steels (EN, ASTM, AISI, GB) and addresses a complementary but distinct question: whether the current population of steel grades is technically necessary to preserve mechanical performance. Rather than proposing new chemistries or unified alloys, KMRP quantitatively interrogates existing standards to identify where multiple designated grades already deliver indistinguishable strength–ductility behaviour and therefore constitute functional redundancy.

When applied to the examined steel grades, KMRP demonstrated that 47.9% of the 434 examined grades could be reduced without loss of mechanical coverage, while fully preserving the UTS range of 120–2450 MPa and elongation range of 2–55%. Crucially, this rationalisation was not driven solely by mechanical overlap. Within mechanically equivalent sets, grades were further filtered using recyclability-relevant criteria, prioritising those with higher tolerance to tramp elements and reduced dependency on critical raw materials. This differentiates KMRP from purely performance-based consolidation approaches and directly links grade reduction to circular-economy constraints.

From a recycling perspective, as mentioned in chapter 2, studies increasingly identify tramp-element accumulation, particularly copper, as a structural limitation on higher scrap utilisation. Scrap-tolerant steel concepts therefore seek to accommodate unavoidable tramp elements through alloy and processing strategies rather than relying on increasingly pure scrap streams. In this context, KMRP plays a complementary role: it does not redesign alloys to improve impurity tolerance, but instead demonstrates where existing grades already show equivalent mechanical performance while differing in residual tolerance or CRM content. Where such redundancy exists, grades with more favourable recyclability characteristics can be prioritised; where it does not, grades are correctly retained as mechanically indispensable.

Overall, industrial simplification initiatives articulate why reducing grade variety is desirable, while the KMRP outputs provide a data-driven, portfolio-level assessment of how far rationalisation can proceed using today's standards. The methodology therefore complements alloy-design and recycling-oriented research by offering a systematic means of identifying which existing steel grades are genuinely necessary and which contribute primarily to unnecessary complexity within the current standardisation system.

8.6 Summary

In the maximum tramp element strategy, the KMRP algorithm was applied and its outcomes interpreted to validate its role as a rationalisation tool. Rather than redesigning new alloys, KMRP retained a minimal set of existing international grades (e.g., ASTM, AISI, EN). These retained grades were analytically confirmed to preserve the required mechanical property coverage within each cluster while tolerating higher levels of tramp elements.

In Cluster 0, ferritic stainless steels and low-carbon steels both occupied the same performance range of low-to-mid strength and moderate elongation. KMRP showed that 52% of grades were redundant, and by removing them, the group retained full property coverage while reducing unnecessary duplication between two separate classifications. Reduction process remained sensitive to existing classifications (carbon to carbon, stainless to stainless). This also reduced reliance on molybdenum, which SHAP confirmed was not essential for the group's performance.

In cluster 1, medium-carbon steels and martensitic stainless steels both addressed high-strength, low-elongation applications. KMRP revealed that nearly a quarter of grades could be reduced, with copper-tolerant grades retained to strengthen recyclability. Reduction favoured copper-tolerant steels while narrowing molybdenum and vanadium ranges, demonstrating that rationalisation can reduce CRM reliance while retaining functional equivalence.

In cluster 2, steels providing the highest UTS and lowest elongation formed a distinct group where no redundancy was possible. All grades were retained, showing that when performance ranges are non-overlapping, rationalisation cannot proceed without property loss.

In cluster 3, austenitic stainless steels and a subset of low-carbon steels overlapped in covering the high-elongation, medium-to-high strength range. KMRP reduced the number of stainless steel grades by almost two-thirds, retaining copper-tolerant grades and lowering CRM usage

while preserving the property space. This confirmed that within ductile applications, large redundancies exist.

Overall, the results demonstrated that out of the 146 existing carbon steel grades, only 90 were sufficient to fully cover the selected mechanical property ranges (UTS and elongation). These retained grades represent more sustainable alternatives, as they contain higher tolerance to the recyclability-limiting elements and lower dependency on the critical raw materials. A similar outcome was observed for stainless steels, where just 136 out of the original 288 grades were needed to preserve the complete mechanical performance range, while substantially reducing the reliance on environmentally and economically costly compositions.

8.7 Broader Impact and Relevance

This reduction of grade set has wider implications beyond data analysis and engineering. It supports more efficient recycling, sustainable production, and informed decision-making.

As discussed in Chapter 1, global crude steel production reached approximately 1,951 million tonnes in 2021, of which only 32% (around 630 million tonnes) was derived from recycled scrap (World Steel Association, 2021). The remaining 68% came from primary production using resource and energy-intensive methods such as the BF–BOF route. The steel industry is among the largest industrial CO_2 emitters globally, accounting for roughly 7-9% of total emissions (De Ras et al., 2019). Steel produced via scrap-fed EAFs emits approximately 300 kg CO_2 per tonne, compared to 1,200 kg CO_2 per tonne from the BF–BOF process (Bailera et al., 2021). Therefore, shifting more steel production to the EAF route could reduce emissions by up to 900 kg CO_2 per tonne.

In this study, the number of examined carbon steel grades was reduced from 146 to 90, and stainless steels from 288 to 136, while still covering the full UTS and elongation ranges. The refined subset of grades contained fewer CRMs and included steels with higher copper tolerance, making them more compatible with future recycling challenges.

To contextualise the potential impact the following example is considered. If each of the 56 replaced by grades that contains lower CRM and higher tolerant to tramp element that are routed through EAF instead of BF–BOF, and each accounted for 1 million tonnes of production annually, this could result in a saving of 50.4 million tonnes of CO_2 per year (56 × 900 kg CO_2/t). Similarly, the reduction of 152 stainless steel grades could potentially save 136.8 million tonnes of CO_2 per year if produced via EAF.

Importantly, the KMRP output should not be interpreted as an alternative standard or a replacement for existing designation systems. Instead, it provides an analytical means of questioning whether the current population of standard grades is technically necessary. Within mechanically defined performance windows, the results show that multiple EN, ASTM, or AISI grades often deliver indistinguishable strength–ductility behaviour, despite being treated as distinct designations. From an industrial perspective, this indicates that grade proliferation, rather than mechanical necessity, is a limiting factor for recyclability and system efficiency.

A similar opportunity exists on the recycling side. The impurity and CRM distributions quantified in Chapter 8 define natural bands in terms of allowable Cu, Mo and V levels for each KMRP cluster (e.g. Cu 0–0.4% and Mo ≤3.1% for the majority of medium-strength steels). These bands could form the basis of recyclability-oriented classes or digital product categories, in which groups of standard grades are treated as a single family. Integrating such KMRP-derived families into future European scrap standards or recycling codes would therefore not replace existing grade specifications, but would add an application and recyclability-based layer that helps steelmakers and recyclers to manage tramp elements and CRMs more systematically while remaining within the established EN 10027 framework.In summary, the KMRP-based reclassification system supports a cleaner and more manageable steel ecosystem. It creates a foundation for improving recycling practices, reducing emissions, lowering complexity, and making informed choices at both policy and application levels.

Chapter 9: Conclusion and Future Work

9.1 Conclusion

This study proposed and validated a data-driven framework for rationalising the diversity of existing steel grades with the goal of enhancing recyclability without compromising mechanical performance or production cost. The motivation originated from the excessive variety of over 3,500 steel grades, many of which shared overlapping performance characteristics but differ in chemical composition or processing conditions, which ultimately complicates recycling, causes environmental damage, and generates a lot of waste.

To address this, objective (1) was achieved by constructing and analysing a multi-source steel database integrating chemical composition, processing routes, and mechanical properties from major international standards (AISI, ASTM, EN, GBT, UNS). This study focused on carbon and stainless steels as representative families of steel. As a foundational step, objective (2) was achieved by quantifying the relationship between chemical composition, processing, and mechanical properties using supervised learning models. Comparative modelling revealed that ANNs consistently performed better than tree-based algorithms (XGBoost, RF) in predicting key mechanical properties such as UTS, YS, and elongation in both steel families. This evaluation was crucial for two reasons: (1) ensuring that subsequent clustering and rationalisation were grounded in accurate property understanding, and (2) establishing a robust baseline for potential future integration with neural generative models for inverse designing.

At the core of the methodology, objective (3) was achieved through the development of the KMRP to facilitate application-driven reclassification, aiming to simplify steel grade systems and reduce variety. This hybrid algorithm combined unsupervised clustering with grade reduction and sustainability-aware filtering. In its first phase, K-Means clustering was applied to selected mechanical property data (hardness, UTS, YS, and elongation) to segment grades into functionally distinct groups. Four groups were identified based on selected mechanical properties, regardless of predefined alloy families or chemical compositions. The segmentation approach enabled the identification of grades with similar performance but differing chemical compositions, allowing for direct comparison and ultimately reducing grade variety by selecting options with higher tolerance of tramp element content and fewer CRMs, in accordance with sustainability constraints. The results demonstrated that some steel families shared almost the same performance for the mechanical properties. In addition, further analysis

of chemical composition in each group revealed alignment with conventional metallurgical classifications of low, medium, and high carbon steels. The proposed method also demonstrated its utility as a powerful technique for compositional grouping when alloy chemistry is unavailable or incomplete. Cluster 0 consisted of low-carbon steels and ferritic stainless steels (low strength, moderate elongation); Cluster 1 included medium-carbon steels and martensitic stainless steels (moderate to high strength, low elongation); Cluster 2 formed high-carbon steels and martensitic stainless steels (very high strength, low ductility); and Cluster 3 was dominated by austenitic stainless steels with a subset of low-carbon steels (moderate strength, high elongation).In the second phase, objective (4) was achieved by applying a rule-based overlap-detection and reduction strategy within each cluster, retaining only the minimum number of grades required to cover the full mechanical property range of that group under the existing steel classification. Unlike conventional clustering methods, KMRP introduced second layer of sustainability constraints, allowing users to prioritise grades based on tramp element content and CRM levels; this refinement achieved objective (5) by simplifying the alloy system in a way that improves recyclability (via tramp-element tolerance) and reduces dependence on critical raw materials while preserving application-relevant mechanical performance.

Under the tramp element tolerance strategy, which prioritised recyclable but copper-tolerant steels, cluster 0 achieved a 52.1% reduction (from 186 to 89 grades) while preserving performance by retaining higher copper-tolerant compositions. Cluster 1 was reduced from 91 to 71 grades, showing a shift toward copper while still deprioritising high vanadium content. Cluster 2 remained unreduced due to the uniqueness of its performance profile, and no overlap was detected. Cluster 3 achieved the highest reduction, with only 57 of the original 148 grades retained. In the broader context of traditional predefined categories, this strategy revealed that only 90 of the 146 carbon steel grades and 136 of the 288 stainless steel grades were required to maintain the original performance coverage of these groups. Furthermore, the concentration ranges of critical raw materials were significantly narrowed. For instance, the molybdenum range was reduced from 0–7% to a maximum of 4.3%, while still covering the original range of mechanical performance and intended applications.

Overall, this research not only demonstrated a practical method for alloy system simplification but also delivered a flexible, explainable, and sustainability-conscious tool that could inform both industrial alloy selection and policy-led standardisation. The generalisability of the KMRP framework further suggested potential applications in other domains such as pharmaceutical

compound reduction or supply chain optimisation, where rationalising complex option sets is essential.While the results above demonstrate that the KMRP framework can reduce steel grade variety while preserving mechanical performance and improving recyclability-related characteristics, their significance extends beyond methodological validation. On this basis, the following points summarise the main industrial implications of the findings, followed by key directions for future development of the framework.

Industrial implications

- The results provide quantitative evidence that a large share of the current grade landscape across EN/ASTM/AISI is mechanically redundant within the examined strength–ductility space, meaning that the size of the existing number of steel grades are not fully justified by performance requirements.
- The retained sets can be used in practice as a rationalisation reference, where mechanically interchangeable grades exist, by favouring options with higher tramp-element tolerance and lower CRM dependency while remaining within the existing standards.
- For scrap-rich production routes, the outcomes support more robust decision-making because the reduction step explicitly prioritises grades that are less sensitive to tramp element constraints, which are increasingly unavoidable as scrap utilisation increases, when viable alternatives exist.
- For standardisation and product portfolio management, the findings do not introduce new steel grade designations; instead, they provide a quantitative basis for identifying cases where multiple existing grades show overlapping mechanical performance and could therefore be rationalised, without compromising the mechanical coverage.

Future outlook

- Extending the analysis beyond mechanical properties to include additional properties such as corrosion resistance, fatigue performance, weldability, and high-temperature stability would allow the same rationalisation logic to be tested under more application-specific constraints.
- Incorporating richer processing metadata (temperatures, cooling rates, strain history) would strengthen the interpretation of overlap by linking mechanical equivalence more directly to microstructural equivalence.

- Introducing industrial weighting factors (production volumes, market prevalence, supply-chain constraints) would improve the practical relevance of retained grades by balancing technical sufficiency with real-world adoption.
- The exploratory generative modelling work can be developed into cluster-conditioned inverse design, where recyclable compositions are generated within the mechanically feasible envelopes defined by KMRP, and then validated using CALPHAD or experimental trials.

While these implications and future directions illustrate the broader relevance and potential impact of the KMRP framework, it is also important to acknowledge the methodological and data-related limitations that constrain the scope and generalisability of the present research.

9.2 Limitations of the Study

Despite the promising outcomes of the KMRP framework, several limitations were identified during the research process. The limitations outlined below do not affected the core findings; instead, they highlight important aspects of methodological enhancement and further validation.

Limited data availability

The analysis relied on major standards (AISI, ASTM, EN, GBT); however, despite this broad scope, several commercially available steel grades were excluded due to a lack of publicly available or complete data. Some standards provided only partial datasets (for instance, missing elongation or yield strength values), which directly affected the clustering and reduction phases. Moreover, inconsistencies in naming conventions and formatting across sources required manual validation and introduced potential errors in the specification of steel grades. Although these issues were addressed through literature cross-validation and manual review, the absence of a unified, highly reliable global database of steel grades remains a structural limitation that may impact the robustness of any data-driven approach.

Incomplete processing information

In this study, processing routes were represented only via binary flags (e.g., Annealed = 1), without precise process parameters such as temperatures, or cooling rates. The simplified representation was a direct consequence of the limited availability of detailed processing metadata across the selected standards. As a result, the model may have underrepresented processing-property interactions that are critical to material performance. This limitation likely

affected the accuracy of the supervised learning models in predicting mechanical properties, which are known to be processing-sensitive.

Industrial usage validation

While the KMRP framework effectively reduced the variety of steel grades based on mechanical performance and sustainability criteria, the reduction process did not include information on the industrial majority or production volumes of specific grades. Commercial importance, historical usage, and manufacturing frequency are often confidential or poorly documented across standards and markets. Based on this fact, some retained grades may be technically optimal but rarely used in practice. In contrast, others that are widely produced globally may have been excluded due to their marginal performance. Incorporating production-based weightings or real-world demand data in future iterations of KMRP could enhance its practical relevance and industrial adoption.

Focus on mechanical properties only

This study focused on mechanical properties (hardness, UTS, YS, and elongation) as the basis for clustering and grade reduction. Consequently, the generated groups reflect similarity only within this limited performance space. However, other critical properties (such as corrosion resistance, weldability, fatigue behaviour, and thermal stability) were not included due to the lack of standardised or widely available data across steel grades. The absence of these attributes may result in groupings that overlook important application-related properties, potentially limiting the suitability of retained grades in specific applications. Future extensions of this work should integrate multi-property clustering once broader and more reliable datasets become available.

9.3 Recommendations for Future Direction

While this study successfully introduced the multi-step framework (KMRP) for the purpose of rationalising steel grades, several directions exist for extending and enhancing this research.

(1) Extension to other alloy systems and industrial domains

This study focused on carbon and stainless steels. However, the underlying methodology of the KMRP algorithm is not limited to ferrous alloys. An extension of this work would involve applying the approach to other alloy families, including low-alloy steels, tool steels, aluminium alloys, titanium alloys, and nickel-based superalloys. These materials are increasingly used in

critical industries such as aerospace, automotive, and medical devices, where performance optimisation and material simplification are relevant.

Moreover, the principles of clustering based on functional performance and subsequent reduction via domain-specific constraints such as cost, or sustainability are applicable far beyond metallurgy. For instance, in pharmaceutical development, KMRP could potentially be used to reduce the number of candidate molecules that contain similar therapeutic efficacy but vary in chemical structure or production cost.

Future research could explore how KMRP performs in these non-metallic or interdisciplinary contexts, especially when combined with domain-specific features.

(2) Inverse alloy design development via cluster-conditioned GANs

While this study focused on reducing the number of steel grades within the existing standards, an exploratory extension was performed to investigate the potential of generative AI for inverse alloy design, aiming to propose new alloy specifications within each cluster generated by KMRP that are more recyclable. This inverse design direction builds on the KMRP framework by seeking to generate novel steel compositions that achieve target properties within each performance-based cluster.

Preliminary experiments were conducted using three conditional generative architectures (CGAN, CGAN-Tune, and WGAN-GP) conditioned on UTS, elongation, and cluster membership as defined by KMRP. The target was to generate single representative alloy compositions that matched the highest mechanical performance observed within each cluster. Among the tested models, WGAN-GP showed better performance in producing chemically valid compositions that conformed to cluster-specific compositional bounds while meeting maximum UTS and elongation values, particularly within clusters 1 and 3, corresponding to high-strength martensitic and high-ductility austenitic steels, respectively. This work remains exploratory and was limited by the sparsity of certain cluster regions in the original dataset. Broader and more balanced training data will be necessary for robust implementation. Future research can refine this approach by collecting more standard steel grades and integrating broader property targets (such as corrosion resistance, weldability, and fatigue strength) into the conditioning criteria. Moreover, including physical and thermodynamic constraints during generation, along with experimental validation and CALPHAD-based evaluations, will be essential to ensure the feasibility and scientific validity of the generated alloys.

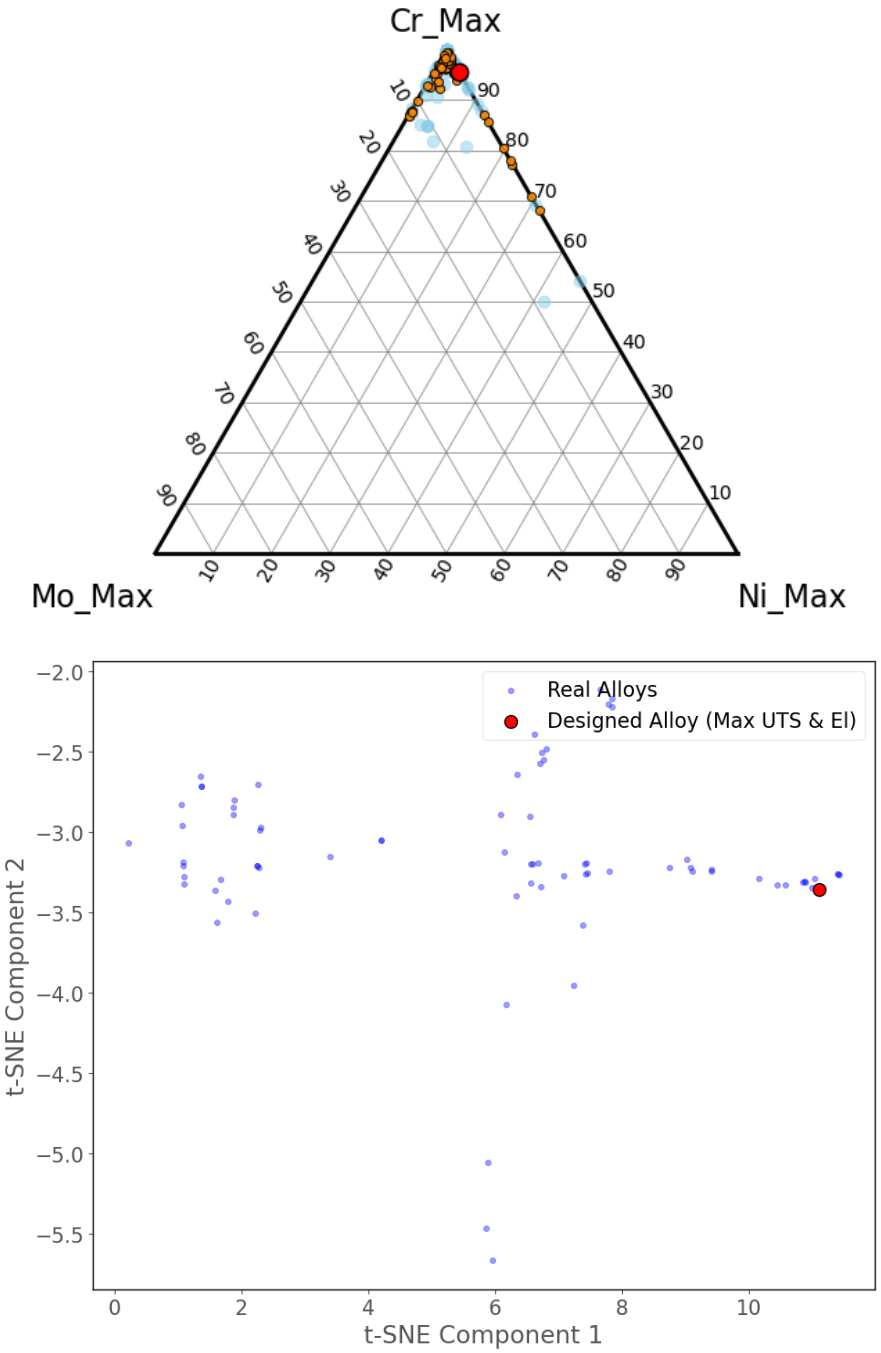


Figure 9.1: (a) Ternary plot showing the designed stainless steel alloy composition (red) relative to real Cluster 0 alloys (blue) in Mn–C–Si space. (b) t-SNE projection illustrating the designed alloy (red) within the distribution of real alloys (blue).

(3) Database expansion and integration of KMRP with LLM-based decision-support platform

The current study was conducted using structured data derived from international steel standards, with a focus on chemical composition, binary-encoded processing conditions, and core mechanical properties. However, real-world alloy often involves multi-objective considerations beyond these parameters, including corrosion resistance, thermal/electrical

conductivity, weldability, cost, and application-specific constraints. Future research could aim to systematically expand the KMRP framework by incorporating such multi-attribute data into the clustering and reduction process, enabling a more holistic rationalisation of steel grades.

In parallel, the emergence of large language models (LLMs) such as GPT-4 and Gemini presents a novel opportunity for integrating KMRP into knowledge-augmented alloy systems. Specifically, retrieval-augmented generation (RAG) techniques could be explored as a way to enhance structured datasets with unstructured domain knowledge extracted from alloy standards, datasheets, manufacturing protocols, or environmental policies. This could bridge the gap between a data-driven approach and the contextual interpretation of alloy specifications that are often documented in textual form.

From a research perspective, such an integration raises important questions around model interpretability, information fidelity, and multi-modal learning. For instance, future work could investigate whether LLMs can accurately parse and extract critical features from technical standards and how this information might be used to dynamically refine KMRP-based clustering or filtering. A knowledge-augmented decision framework would thus allow material scientists to query alloy options based on both numerical constraints (e.g., UTS > 800 MPa) and textual guidance (e.g., “suitable for high-temperature structural applications with low chromium content”), offering a richer, more interactive design experience.

Ultimately, this direction invites interdisciplinary collaboration between materials science, data science, and environmental engineering, enabling the development of AI-assisted platforms that enhance transparency, traceability, and scientific reasoning in sustainable alloy system.

Appendix A:

Data Understanding

The parameters reported in Table A.9.1 correspond to the chemical composition features used in this study. For each alloying element, the suffix _Min and _Max denote the minimum and maximum composition limits specified for a given steel grade in the examined international standards. The column count indicates the number of valid records available for each parameter (n = 538). Summary statistics (mean, standard deviation, minimum, maximum, and quartiles) describe the distribution of these values across the examined dataset.

Table A.9.1: Summary of chemical compositions in the examined dataset.

Chemical Element	count	Mean	Std Dev	Min	25%	50%	75%	Max	Type
Fe_Min	538	0.508	0.292	0.0006	0.26	0.50	0.77	0.9998	Float
Fe_Max	538	0.478	0.297	0.0013	0.21	0.47	0.74	0.9986	Float
Mn_Min	538	0.488	0.291	0.0014	0.23	0.48	0.73	0.9988	Float
Mn_Max	538	0.509	0.286	0.0002	0.26	0.51	0.74	0.9995	Float
C_Min	538	0.516	0.285	0.0014	0.28	0.50	0.77	0.9983	Float
C_Max	538	0.494	0.291	0.0035	0.24	0.48	0.77	0.9983	Float
Si_Min	538	0.481	0.286	0.0006	0.23	0.47	0.77	0.9943	Float
Si_Max	538	0.488	0.298	0.0001	0.24	0.47	0.77	0.9978	Float
P_Min	538	0.488	0.278	0.0017	0.25	0.47	0.77	0.9992	Float
P_Max	538	0.494	0.285	0.0085	0.24	0.47	0.77	0.9993	Float
S_Min	538	0.482	0.286	0.0044	0.22	0.47	0.77	0.9962	Float
S_Max	538	0.482	0.299	0.0000	0.21	0.49	0.77	0.9977	Float
Cr_Min	538	0.505	0.292	0.0019	0.24	0.53	0.77	0.9989	Float
Cr_Max	538	0.494	0.286	0.0034	0.26	0.48	0.77	0.9988	Float
Ni_Min	538	0.483	0.296	0.0014	0.21	0.46	0.77	0.9987	Float
Ni_Max	538	0.514	0.290	0.0002	0.26	0.52	0.77	0.9989	Float
Mo_Min	538	0.492	0.287	0.0004	0.25	0.48	0.77	0.9933	Float
Mo_Max	538	0.513	0.295	0.0002	0.25	0.53	0.77	0.9988	Float
Cu_Min	538	0.510	0.293	0.0009	0.23	0.53	0.77	0.9999	Float
Cu_Max	538	0.488	0.290	0.0011	0.22	0.49	0.77	0.9996	Float
Ti_Min	538	0.472	0.293	0.0030	0.21	0.43	0.77	0.9967	Float
Ti_Max	538	0.512	0.288	0.0025	0.24	0.51	0.77	0.9997	Float
Al_Min	538	0.502	0.295	0.0000	0.23	0.52	0.77	0.9972	Float
Al_Max	538	0.494	0.275	0.0044	0.26	0.47	0.77	0.9958	Float
V_Min	538	0.507	0.277	0.0001	0.27	0.50	0.77	0.9994	Float
V_Max	538	0.497	0.289	0.0002	0.25	0.49	0.77	0.9995	Float
N_Min	538	0.520	0.288	0.0007	0.26	0.52	0.77	0.9977	Float
N_Max	538	0.506	0.277	0.0009	0.27	0.50	0.77	0.9995	Float
Nb_Min	538	0.519	0.286	0.0009	0.29	0.52	0.77	0.9994	Float
Nb_Max	538	0.466	0.295	0.0058	0.19	0.45	0.77	0.9997	Float

Table A.9.2 summarises the processing-condition variables included in the dataset. Each processing condition is represented as a binary indicator, where 1 denotes that the treatment is specified for a given steel grade and 0 denotes that it is not specified. The column Total Count indicates the total number of records evaluated (n = 538), while Applied (1s) and Not Applied (0s) report the frequency of each processing condition across the dataset.

Because international standards and databases typically report processing routes qualitatively and without consistent information on temperatures, times, or cooling rates, processing conditions were harmonised into this binary format to ensure consistency across ASTM, AISI, EN, UNS, and GB sources. The Percentage Applied column provides an overview of the distribution of treatments, indicating that no single processing condition dominates the dataset.

Table A.9.2: Summary of processing conditions in the dataset.

Processing Condition	Total Count	Applied (1s)	Not Applied (0s)	Percentage Applied (%)	Type
Hot_Roll	538	253	288	46.77	Boolean
Cold_Roll	538	273	268	50.46	Boolean
Cold_Drawn	538	265	276	48.98	Boolean
Annealing	538	289	252	53.42	Boolean
Cold_Finished	538	278	263	51.39	Boolean
Hot_Finished	538	260	281	48.06	Boolean
Normalising	538	278	263	51.39	Boolean
Carburising	538	258	283	47.69	Boolean
Quenching	538	275	266	50.83	Boolean
Turned	538	275	266	50.83	Boolean
Temper	538	263	278	48.61	Boolean
Hardened	538	260	281	48.06	Boolean

Hierarchical clustering result

Figure A.9.2(a) presents the dendrogram obtained from agglomerative hierarchical clustering applied to the scaled mechanical property dataset, including hardness, UTS, YS, and elongation. Clustering was performed using Ward's linkage method, where the agglomerative distance represents the increase in within-cluster variance at each merge step. The dendrogram is shown in truncated form to highlight the higher-level cluster structure and relative linkage distances between major branches.

Figure A.9.2(b) shows the corresponding two-dimensional projection of the same dataset in principal component space, with samples coloured according to their hierarchical cluster assignments. PCA was used solely for visualisation purposes, while clustering was performed

in the original scaled mechanical property space. The PCA projection illustrates the spatial distribution of the identified clusters and their relative overlap in the reduced-dimensional representation.

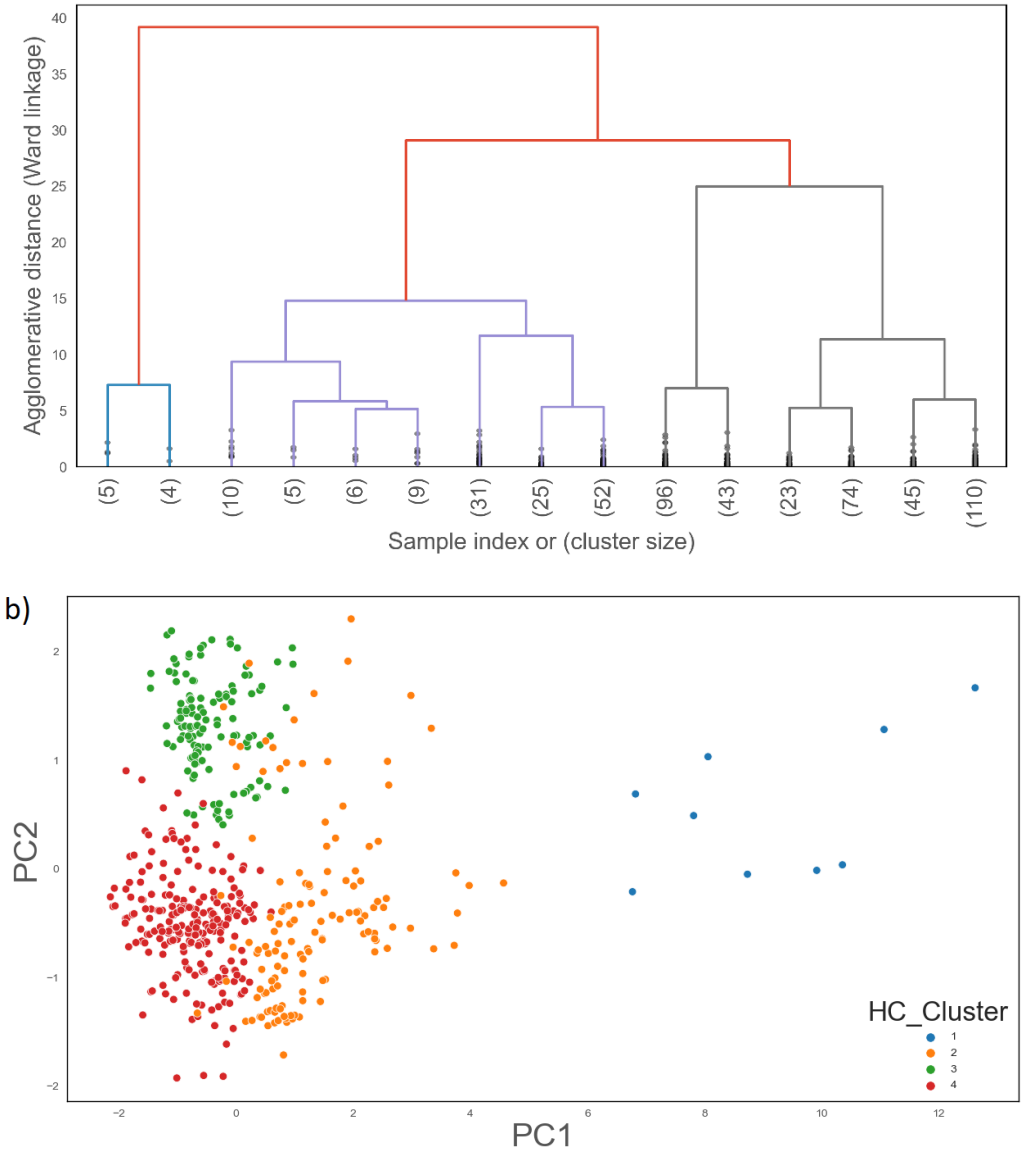


Figure A.9.2: (a) Dendrogram generated and (b) Distribution of steel grades based on PCA-transformed 2D from agglomerative hierarchical clustering of scaled mechanical properties (Hardness, UTS, YS, and Elongation). PC1 primarily represents variation associated with strength-related properties (hardness, UTS, YS), while PC2 reflects variation dominated by elongation.

DBSCAN result

Figure A.9.3(a) shows the k-distance plot computed using the 5-nearest neighbours of each data point, based on the scaled mechanical property dataset (hardness, UTS, YS, and elongation). The k-distance values were sorted in ascending order to support the selection of the DBSCAN neighbourhood radius (ε) by identifying the transition between dense and sparse regions in the feature space.

Figure A.9.3(b) presents a two-dimensional visualisation of the DBSCAN clustering outcome projected onto the first two principal components. Samples are coloured according to their DBSCAN cluster assignments, including points classified as noise (cluster −1). PCA was used for visualisation, while clustering was performed in the original scaled mechanical property space.

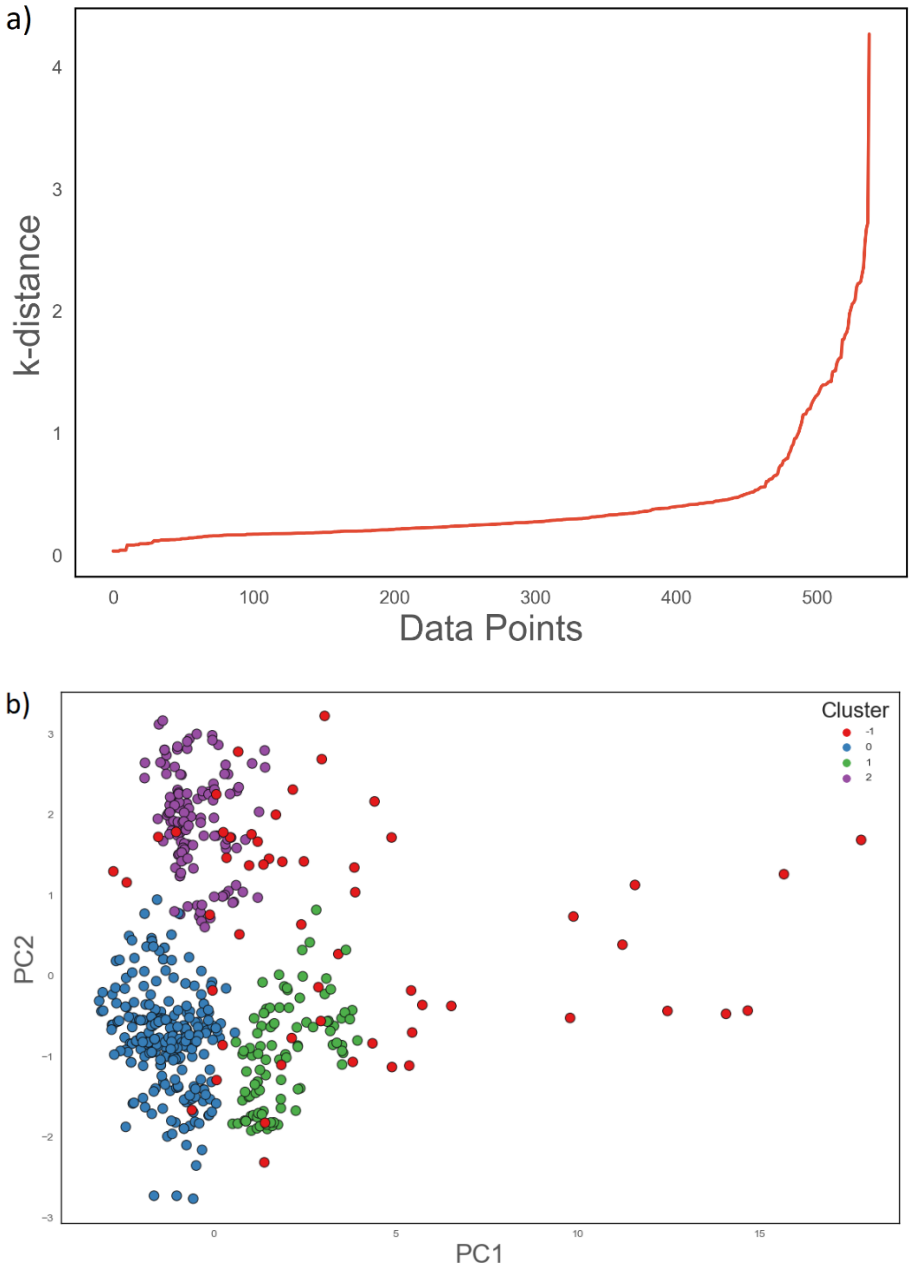


Figure A.9.3: (a) k-distance plot using 5-nearest neighbours to estimate eps value and (b) Distribution of steel grades based on PCA-transformed 2D DBSCAN results using scaled mechanical properties (Hardness, UTS, YS, and Elongation). PC1 primarily represents variation associated with strength-related properties (hardness, UTS, YS), while PC2 reflects variation dominated by elongation.

K-Means result

K-Means clustering was evaluated on the same scaled mechanical property dataset (hardness, UTS, YS, and elongation) using multiple values of K. Table A.9.3 reports the Silhouette Score and Davies–Bouldin Index (DBI) for each configuration, where higher Silhouette and lower DBI indicate improved cohesion and separation. Figure A.9.4 visualises the K-Means assignments in the first two principal components for interpretability; PCA was used only for visualisation, while clustering was performed in the original scaled feature space.

Table 9.3: Evaluation of K-Means clustering performance for different values of K using Silhouette Score and Davies-Bouldin Index.

K	Silhouette score	Davies Bouldin Index
2	0.47	1.00
3	0.45	0.79
4	0.58	0.65
5	0.50	0.77
6	0.42	0.84
7	0.40	0.95
8	0.39	0.97
9	0.37	0.99
10	0.37	0.98

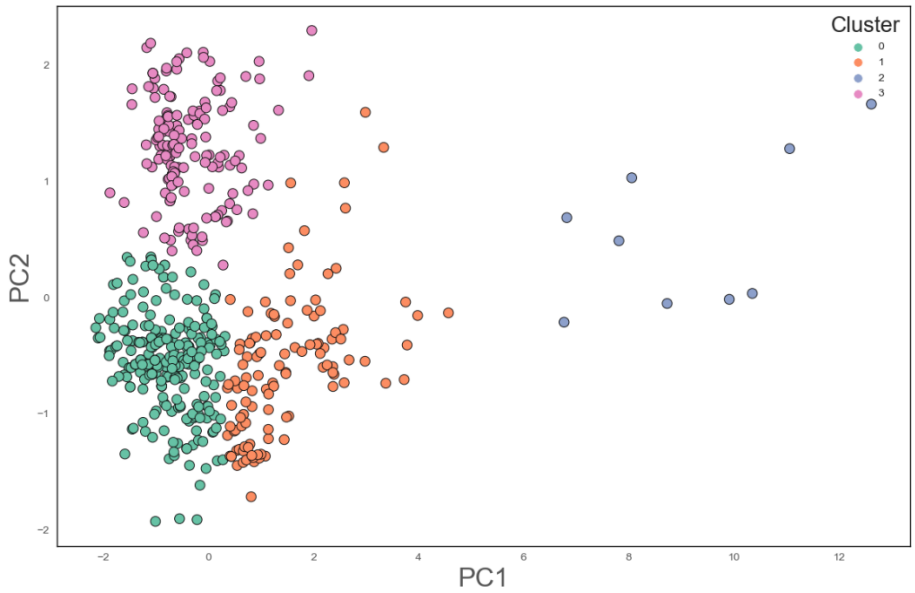


Figure 9.3: Distribution of steel grades based on PCA-transformed 2D K-Means results using scaled mechanical properties (Hardness, UTS, YS, and Elongation). PC1 primarily represents variation associated with strength-related properties (hardness, UTS, YS), while PC2 reflects variation dominated by elongation.

Appendix B – KMRP code

```
from __future__ import annotations

from typing import Dict, Tuple, List, Optional

from copy import deepcopy

import numpy as np

try:

    from sklearn.cluster import KMeans

except ImportError:

    raise ImportError("scikit-learn is required for KMeans. Install via: pip
install scikit-learn")


class KMRP:

    """

    K-Means Reduction Process (KMRP)


    Phase I/II: Apply K-Means to group grades in a compact mechanical-property
space.

    Phase III : Within each (cluster × Category), remove redundant grades while
preserving

                full UTS and Elongation coverage and respecting tramp/CRM
preference.


    Expected grades_data schema (per grade key):

        {

          "Category": "Carbon Steel" | "Stainless Steel" | ...,

          "UTS": (min_uts, max_uts),

          "Elongation": (min_el, max_el),

          # Optional if available:

          "YS": (min_ys, max_ys),

          # For tramp/CRM calculations:

          "Cu_Max": float,

          "Mo_Max": float, "V_Max": float, "Ni_Max": float, "Ti_Max": float

        }

    """
```

```
    # ---------- Construction ----------

    def __init__(self, grades_data: Dict[str, Dict]):
        # Work on a deep copy to avoid side effects
        self.grades_data = deepcopy(grades_data)
        self.labels_: Dict[str, int] = {}         # grade -> cluster label
        self.clusters_: Dict[int, List[str]] = {} # label -> [grades]
        self.kmeans_: Optional[KMeans] = None

    # ---------- Utilities ----------

    @staticmethod
    def _midpoint(r: Tuple[float, float]) -> float:
        return 0.5 * (float(r[0]) + float(r[1]))

    @staticmethod
    def _interval_union_gaps(intervals: List[Tuple[float, float]]) ->
Tuple[List[Tuple[float, float]], List[Tuple[float, float]]]:
        """
        Given a list of (start, end) with start <= end (continuous values),
        returns (merged_intervals, gaps_between_merged).

        This avoids discretising to integers (more robust than range(...)).
        """
        if not intervals:
            return [], []
        intervals = sorted([(min(a, b), max(a, b)) for a, b in intervals],
key=lambda x: x[0])

        merged = []
        cur_start, cur_end = intervals[0]
        for s, e in intervals[1:]:
            if s <= cur_end:  # overlap or touching
                cur_end = max(cur_end, e)
            else:
```

```
                merged.append((cur_start, cur_end))
                cur_start, cur_end = s, e
        merged.append((cur_start, cur_end))

        # compute gaps between merged intervals
        gaps = []
        for i in range(len(merged) - 1):
            left_end = merged[i][1]
            right_start = merged[i + 1][0]
            if right_start > left_end:
                gaps.append((left_end, right_start))
        return merged, gaps

    @staticmethod
    def _covers(interval_a: Tuple[float, float], interval_b: Tuple[float, float])
-> bool:
        """
        True if interval_a fully covers interval_b (continuous).
        """
        return interval_a[0] <= interval_b[0] and interval_a[1] >= interval_b[1]

    @staticmethod
    def _calc_tramp_sum(d: Dict) -> float:
        # You can extend this list if needed (e.g., ["Cu_Max", "Sn_Max"])
        tramp_elements = ["Cu_Max"]
        return float(sum(d.get(k, 0.0) for k in tramp_elements))

    @staticmethod
    def _calc_critical_sum(d: Dict) -> float:
        critical_elements = ["Mo_Max", "V_Max", "Ni_Max", "Ti_Max"]
        return float(sum(d.get(k, 0.0) for k in critical_elements))

    # ---------- Coverage metrics ----------
```

```
    @staticmethod
    def _coverage_and_gaps(block: Dict[str, Dict], key: str) ->
Tuple[Optional[float], Optional[float], List[Tuple[float, float]]]:
        """
        For a given property key ("UTS" or "Elongation"), compute overall min,
max, and any gaps.
        Uses continuous interval merging (no integer discretisation).
        """
        if not block:
            return None, None, []
        intervals = [tuple(map(float, block[g][key])) for g in block]
        merged, gaps = KMRP._interval_union_gaps(intervals)
        overall_min = merged[0][0]
        overall_max = merged[-1][1]
        return overall_min, overall_max, gaps


    # ---------- Overlap detection ----------


    def _find_overlapping_grades(self, pool: Dict[str, Dict], grade: str) ->
List[str]:
        """
        Grades in 'pool' that fully cover the selected grade in UTS OR Elongation.
        Note: 'cover' means the other interval fully contains the selected
interval
        (no 'subset' acceptance).
        """
        assert grade in self.grades_data, f"Grade '{grade}' not found."
        sel_uts = tuple(map(float, self.grades_data[grade]["UTS"]))
        sel_el = tuple(map(float, self.grades_data[grade]["Elongation"]))


        overlaps = []
        for g in pool:
            if g == grade:
                continue
            g_uts = tuple(map(float, self.grades_data[g]["UTS"]))
```

```
            g_el = tuple(map(float, self.grades_data[g]["Elongation"]))

            full_uts_cover = self._covers(g_uts, sel_uts)
            full_el_cover = self._covers(g_el, sel_el)

            if full_uts_cover or full_el_cover:
                overlaps.append(g)
        return overlaps

    # ---------- KMeans (Phase I/II) ----------

    def _build_feature_matrix(self) -> Tuple[np.ndarray, List[str]]:
        """
        Build a compact feature matrix from range midpoints:
            [UTS_mid, El_mid, (optional) YS_mid]
        """
        names = []
        feats = []

        for name, d in self.grades_data.items():
            if "UTS" not in d or "Elongation" not in d:
                # skip malformed entries
                continue
            uts_mid = self._midpoint(d["UTS"])
            el_mid = self._midpoint(d["Elongation"])
            row = [uts_mid, el_mid]
            if "YS" in d and d["YS"] is not None:
                try:
                    row.append(self._midpoint(d["YS"]))
                except Exception:
                    pass
            names.append(name)
            feats.append(row)
```

```
        if not feats:
            raise ValueError("No valid grades with UTS/Elongation found to build features.")
        return np.array(feats, dtype=float), names

    def fit_kmeans(self, n_clusters: int, random_state: int = 42, n_init: int = 10) -> Dict[str, int]:
        """
        Fit K-Means on mid-point features and store cluster labels for each grade.
        Returns dict grade->cluster_label.
        """
        X, names = self._build_feature_matrix()
        km = KMeans(n_clusters=n_clusters, random_state=random_state, n_init=n_init)
        labels = km.fit_predict(X)
        self.kmeans_ = km

        self.labels_ = {name: int(lbl) for name, lbl in zip(names, labels)}
        # Build cluster map
        clusters = {}
        for name, lbl in self.labels_.items():
            clusters.setdefault(lbl, []).append(name)
        self.clusters_ = clusters
        return self.labels_

    # ---------- Elimination (Phase III) ----------

    def _initial_coverage_snapshot(self, grades: Dict[str, Dict]) -> Dict[str, Tuple[float, float, List[Tuple[float, float]]]]:
        """
        Returns baseline coverage (min, max, gaps) for UTS and Elongation.
        """
        uts_min, uts_max, uts_gaps = self._coverage_and_gaps(grades, "UTS")
        el_min, el_max, el_gaps = self._coverage_and_gaps(grades, "Elongation")
        return {
```

```
            "UTS": (uts_min, uts_max, uts_gaps),
            "Elongation": (el_min, el_max, el_gaps),
        }


    @staticmethod
    def _same_coverage(a: Tuple[Optional[float], Optional[float],
List[Tuple[float, float]]],
                       b: Tuple[Optional[float], Optional[float],
List[Tuple[float, float]]],
                       tol: float = 1e-9) -> bool:
        """
        Compare two coverage triplets (min, max, gaps). Uses a small tolerance for
floats.
        Gaps must match in count and approximately in endpoints.
        """
        (a_min, a_max, a_gaps), (b_min, b_max, b_gaps) = a, b
        if a_min is None and b_min is None:
            return True
        if a_min is None or b_min is None:
            return False
        if abs(a_min - b_min) > tol or abs(a_max - b_max) > tol:
            return False
        if len(a_gaps) != len(b_gaps):
            return False
        for (al, ar), (bl, br) in zip(a_gaps, b_gaps):
            if abs(al - bl) > tol or abs(ar - br) > tol:
                return False
        return True


    def _eliminate_within_block(self, block_grades: List[str], tramp_priority:
str) -> Tuple[List[str], List[str]]:
        """
        Eliminate redundant grades inside a block (i.e., within same cluster AND
same Category).
        """
        # Create a local view (dict) for this block
```

```
        block = {g: self.grades_data[g] for g in block_grades}

        # Record initial coverage
        baseline = {
            "UTS": self._coverage_and_gaps(block, "UTS"),
            "Elongation": self._coverage_and_gaps(block, "Elongation"),
        }

        remaining = list(block.keys())
        eliminated: List[str] = []

        i = 0
        while i < len(remaining):
            g = remaining[i]

            # Build pool without g
            pool = {h: self.grades_data[h] for h in remaining if h != g}

            # If removing g breaks coverage (min/max/gaps) -> keep g
            pool_cov_uts = self._coverage_and_gaps(pool, "UTS")
            pool_cov_el = self._coverage_and_gaps(pool, "Elongation")
            uts_ok = self._same_coverage(pool_cov_uts, baseline["UTS"])
            el_ok = self._same_coverage(pool_cov_el, baseline["Elongation"])

            if uts_ok and el_ok:
                # Candidates that fully cover g in UTS or Elongation
                overlaps = self._find_overlapping_grades(pool, g)

                if overlaps:
                    sel_tramp = self._calc_tramp_sum(self.grades_data[g])
                    sel_crit = self._calc_critical_sum(self.grades_data[g])

                    if tramp_priority == "min":
                        # Prefer lower tramp; on ties prefer <= critical
```

```
                        better_tramp_exists =
any(self._calc_tramp_sum(self.grades_data[o]) < sel_tramp for o in overlaps)
                        if better_tramp_exists:
                            eliminated.append(g)
                            remaining.remove(g)
                            i = 0
                            continue
                        # tie on tramp -> prefer lower/equal critical
                        tie_on_tramp =
all(abs(self._calc_tramp_sum(self.grades_data[o]) - sel_tramp) < 1e-12 for o in
overlaps)
                        if tie_on_tramp:
                            better_crit_exists =
any(self._calc_critical_sum(self.grades_data[o]) <= sel_crit for o in overlaps)
                            if better_crit_exists:
                                eliminated.append(g)
                                remaining.remove(g)
                                i = 0
                                continue


                    elif tramp_priority == "max":
                        # Prefer higher tramp; on ties prefer <= critical (your
original secondary rule)
                        worse_tramp_exists =
any(self._calc_tramp_sum(self.grades_data[o]) > sel_tramp for o in overlaps)
                        if worse_tramp_exists:
                            eliminated.append(g)
                            remaining.remove(g)
                            i = 0
                            continue
                        # tie on tramp -> prefer lower/equal critical
                        tie_on_tramp =
all(abs(self._calc_tramp_sum(self.grades_data[o]) - sel_tramp) < 1e-12 for o in
overlaps)
                        if tie_on_tramp:
                            better_crit_exists =
any(self._calc_critical_sum(self.grades_data[o]) <= sel_crit for o in overlaps)
```

```
                            if better_crit_exists:
                                eliminated.append(g)
                                remaining.remove(g)
                                i = 0
                                continue
                    else:
                        raise ValueError("Invalid tramp_priority. Use 'min' or
'max'.")
            i += 1

        return remaining, eliminated


    def eliminate(self, tramp_priority: str = "min") -> Tuple[List[str],
List[str]]:
        """
        Run Phase III elimination across all clusters and categories.
        Returns (remaining_grades, eliminated_grades).
        """
        if not self.clusters_:
            # If user forgot to call fit_kmeans, treat all grades as one cluster
            self.clusters_ = {0: list(self.grades_data.keys())}
            self.labels_ = {g: 0 for g in self.grades_data}

        remaining_global: List[str] = []
        eliminated_global: List[str] = []

        for lbl, grades in self.clusters_.items():
            # Partition by Category to avoid cross-type substitution
            by_cat: Dict[str, List[str]] = {}
            for g in grades:
                cat = self.grades_data[g].get("Category", "Unspecified")
                by_cat.setdefault(cat, []).append(g)

            for cat, block in by_cat.items():
```

```
                kept, removed = self._eliminate_within_block(block,
tramp_priority=tramp_priority)
                remaining_global.extend(kept)
                eliminated_global.extend(removed)

        # Deduplicate in case of any overlaps
        remaining_global = sorted(set(remaining_global))
        eliminated_global = sorted(set(eliminated_global))

        return remaining_global, eliminated_global



    def run(self, n_clusters: int, tramp_priority: str = "min", random_state: int
= 42, n_init: int = 10) -> Tuple[Dict[str, int], List[str], List[str]]:
        """
        Full pipeline:
          1) Fit K-Means (Phase I/II)
          2) Eliminate within (cluster × Category) (Phase III)
        Returns (labels_dict, remaining, eliminated).
        """
        labels = self.fit_kmeans(n_clusters=n_clusters, random_state=random_state,
n_init=n_init)
        remaining, eliminated = self.eliminate(tramp_priority=tramp_priority)
        return labels, remaining, eliminated


if __name__ == "__main__":
    kmrp = KMRP(grades_dict)
    labels, remaining, eliminated = kmrp.run(n_clusters=2, tramp_priority="min")
    print("Cluster labels:", labels)
    print("Remaining Grades:", remaining)
```

References

Alabi RO, Elmusrati M, Leivo I, Almangush A, Mäkitie AA. Machine learning explainability in nasopharyngeal cancer survival using LIME and SHAP. Sci Rep 2023;13:8984. https://doi.org/10.1038/s41598-023-35795-0.

Ali A, Abd Razak S, Othman SH, Eisa TAE, Al-Dhaqm A, Nasser M, et al. Financial Fraud Detection Based on Machine Learning: A Systematic Literature Review. Applied Sciences 2022;12. https://doi.org/10.3390/app12199637.

Almeida JS. Predictive non-linear modeling of complex data by artificial neural networks. Curr Opin Biotechnol 2002;13:72–6. https://doi.org/https://doi.org/10.1016/S0958-1669(02)00288-4.

Al-Shaykh OK, Mersereau RM. Lossy compression of noisy images. IEEE Transactions on Image Processing 1998;7:1641–52. https://doi.org/10.1109/83.730376.

ASTM International. Steel Standards. ASTM International 2025. https://store.astm.org/products-services/standards-and-publications/standards/steel-standards.html?utm_source=chatgpt.com (accessed September 4, 2025).

Azimi SM, Britz D, Engstler M, Fritz M, Mücklich F. Advanced Steel Microstructural Classification by Deep Learning Methods. Sci Rep 2018;8:2128. https://doi.org/10.1038/s41598-018-20037-5.

Bailera M, Lisbona P, Peña B, Romeo LM. A review on CO2 mitigation in the Iron and Steel industry through Power to X processes. Journal of CO2 Utilization 2021;46:101456. https://doi.org/https://doi.org/10.1016/j.jcou.2021.101456.

Bailey N. 1 - Factors influencing weldability. In: Bailey N, editor. Weldability of Ferritic Steels, Woodhead Publishing; 1994, p. 1–44. https://doi.org/https://doi.org/10.1533/9781845698935.1.

Bandyopadhyay S, Thakur SS, Mandal JK. Product recommendation for e-commerce business by applying principal component analysis (PCA) and K-means clustering: benefit for the society. Innov Syst Softw Eng 2021;17:45–52. https://doi.org/10.1007/s11334-020-00372-5.

Banerjee A, Dave RN. Validating clusters using the Hopkins statistic. 2004 IEEE International Conference on Fuzzy Systems (IEEE Cat. No.04CH37542), vol. 1, 2004, p. 149–53 vol.1. https://doi.org/10.1109/FUZZY.2004.1375706.

Baştanlar Y, Özuysal M. Introduction to Machine Learning. In: Yousef M, Allmer J, editors. miRNomics: MicroRNA Biology and Computational Analysis, Totowa, NJ: Humana Press; 2014, p. 105–28. https://doi.org/10.1007/978-1-62703-748-8_7.

Battaglia GJ. Mean square error. AMP Journal of Technology 1996;5:31–6.

Bebis G, Georgiopoulos M. Feed-forward neural networks. IEEE Potentials 1994;13:27–31. https://doi.org/10.1109/45.329294.

Bell S, Davis B, Javaid A, Essadiqi E. Final Report on Effect of Impurities in Steel. 2006. https://doi.org/10.13140/RG.2.2.33946.85440.

Bengfort B, Bilbro R, Danielsen N, Gray L, McIntyre K, Roman P, et al. Yellowbrick 2018. https://doi.org/10.5281/zenodo.1206264.

Bhadeshia HKDH. Driving force for martensitic transformation in steels. Metal Science 1981;15:175–7. https://doi.org/10.1179/030634581790426714.

Biau G. Analysis of a Random Forests Model. Journal of Machine Learning Research 2010;13.

Bifarin O. Interpretable machine learning with tree-based shapley additive explanations: Application to metabolomics datasets for binary classification. PLoS One 2023;18:e0284315. https://doi.org/10.1371/journal.pone.0284315.

Black JT. DeGarmo's Materials and Processes in Manufacturing. Tenth edition. Hoboken, NJ : Wiley, [2008] ©2008; 2008.

Black JT, Kohser RA. DeGarmo's materials and processes in manufacturing. John Wiley & Sons; 2017.

Black SC, Chiles V, Lissaman AJ, Martin SJ. 2 - Primary Forming Processes. In: Black SC, Chiles V, Lissaman AJ, Martin SJ, editors. Principles of Engineering Manufacture (Third Edition), Oxford: Butterworth-Heinemann; 1996, p. 13–67. https://doi.org/https://doi.org/10.1016/B978-034063195-9/50033-X.

Blengini GA, Latunussa CEL, Eynard U, De Matos CT, Wittmer D, Georgitzikis K, et al. Study on the EU's list of critical raw materials (2020) final report. Publications Office of the European Union: Luxembourg 2020.

Bodyakova A, Belyakov A. Microstructure and Mechanical Properties of Structural Steels and Alloys. Materials 2023;16. https://doi.org/10.3390/ma16145188.

Bordekar H, Cersullo N, Brysch M, Philipp J, Hühne C. eXplainable artificial intelligence for automatic defect detection in additively manufactured parts using CT scan analysis. J Intell Manuf 2025;36:957–74. https://doi.org/10.1007/s10845-023-02272-4.

Bowyer J, Bratkovich S, Fernholz K, Frank M, Groot H, Howe J, et al. Understanding Steel Recovery and Recycling Rates and Limits to Recycling. Dovetail Partners Outlook 2015.

Breiman L. Random Forests. Mach Learn 2001a;45:5–32. https://doi.org/10.1023/A:1010933404324.

Breiman L. Random Forests. Mach Learn 2001b;45:5–32. https://doi.org/10.1023/A:1010933404324.

Bringas JE. Handbook of Comparative World Steel Standards, 2008.

Bringas JE. Hand Book of Comparative World Steel Standard s. 2004.

Britannica TE of E. stainless steel. Https://WwwBritannicaCom/Technology/Stainless-Steel 2021.

Broadbent C. Steel's recyclability: demonstrating the benefits of recycling steel to achieve a circular economy. International Journal of Life Cycle Assessment 2016;21:1658–65. https://doi.org/10.1007/s11367-016-1081-1.

Bushra A, Kim D, Kan Y, Yi G. AutoSCAN: automatic detection of DBSCAN parameters and efficient clustering of data in overlapping density regions. PeerJ Comput Sci 2024;10:e1921. https://doi.org/10.7717/peerj-cs.1921.

Capdevila C, Caballero FG, Andrés CG de. Determination of Ms Temperature in Steels: A Bayesian Neural Network Model. ISIJ International 2002;42:894–902. https://doi.org/10.2355/isijinternational.42.894.

Carvalho TP, Soares FAAMN, Vita R, Francisco R da P, Basto JP, Alcalá SGS. A systematic literature review of machine learning methods applied to predictive maintenance. Comput Ind Eng 2019;137:106024. https://doi.org/https://doi.org/10.1016/j.cie.2019.106024.

Castelli IE, Olsen T, Datta S, Landis DD, Dahl S, Thygesen KS, et al. Computational screening of perovskite metal oxides for optimal solar light capture. Energy Environ Sci 2012;5:5814–9. https://doi.org/10.1039/C1EE02717D.

Charles RG, Douglas P, Dowling M, Liversage G, Davies ML. Towards Increased Recovery of Critical Raw Materials from WEEE– evaluation of CRMs at a component level and pre-processing methods for interface optimisation with recovery processes. Resour Conserv Recycl 2020;161:104923. https://doi.org/https://doi.org/10.1016/j.resconrec.2020.104923.

Chen Q, Zobel J, Verspoor K. Evaluation of a Machine Learning Duplicate Detection Method for Bioinformatics Databases. Proceedings of the ACM Ninth International Workshop on Data and Text Mining in Biomedical Informatics, New York, NY, USA: Association for Computing Machinery; 2015, p. 4–12. https://doi.org/10.1145/2811163.2811175.

Chen S, Kaufmann T. Development of Data-Driven Machine Learning Models for the Prediction of Casting Surface Defects. Metals (Basel) 2022;12. https://doi.org/10.3390/met12010001.

Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. 2016a. https://doi.org/10.1145/2939672.2939785.

Chen T, Guestrin C. Xgboost: A scalable tree boosting system. Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, 2016b, p. 785–94.

Cheng T, Chen G. Prediction of mechanical properties of hot-rolled strip steel based on PCA-GBDT method. J Phys Conf Ser 2021;1774:012002. https://doi.org/10.1088/1742-6596/1774/1/012002.

Chi C-M, Vossler P, Fan Y, Lv J. Asymptotic properties of high-dimensional random forests. The Annals of Statistics 2022;50. https://doi.org/10.1214/22-AOS2234.

Chiang MM-T, Mirkin B. Intelligent Choice of the Number of Clusters in K-Means Clustering: An Experimental Study with Different Cluster Spreads. J Classif 2010;27:3–40. https://doi.org/10.1007/s00357-010-9049-5.

Chou P, Tsai J, Chou J. Modeling and Optimizing Tensile Strength and Yield Point on a Steel Bar Using an Artificial Neural Network With Taguchi Particle Swarm Optimizer. IEEE Access 2016;4:585–93. https://doi.org/10.1109/ACCESS.2016.2521162.

Circular economy overview. EllenmacarthurfoundationOrg 2021.

Colpaert H. Metallography of Steels: Interpretation of Structure and the Effects of Processing. ASM International; 2018. https://doi.org/10.31399/asm.tb.msisep.9781627082594.

Commission B. World Commission on Environment and Development (WCED)(1987). Our Common Future 1987.

Conejo AN, Birat J-P, Dutta A. A review of the current environmental challenges of the steel industry and its value chain. J Environ Manage 2020;259:109782. https://doi.org/https://doi.org/10.1016/j.jenvman.2019.109782.

Costa APO, Seabra MRR, Tavares JMRS, de Sá JMAC, Santos AD. Analysis of influence of chemical components in the mechanical properties in the steel databases using machine learning. Congress on NumericalMethods in Engineering CMN 2022 (2022. Las Palmas de Gran Canaria), International Center for Numerical Methods in Engineering (CIMNE); 2022, p. 101–17.

Cybenko G V. Approximation by superpositions of a sigmoidal function. Mathematics of Control, Signals and Systems 1989;2:303–14.

Dabiri H, Farhangi V, Moradi MJ, Zadehmohamad M, Karakouzian M. Applications of Decision Tree and Random Forest as Tree-Based Machine Learning Techniques for Analyzing the Ultimate Strain of Spliced and Non-Spliced Reinforcement Bars. Applied Sciences 2022a;12. https://doi.org/10.3390/app12104851.

Dabiri H, Kheyroddin A, Faramarzi A. Predicting tensile strength of spliced and non-spliced steel bars using machine learning- and regression-based methods. Constr Build Mater 2022b;325:126835. https://doi.org/10.1016/j.conbuildmat.2022.126835.

Daehn KE, Cabrera Serrenho A, Allwood JM. How Will Copper Contamination Constrain Future Global Steel Recycling? Environ Sci Technol 2017;51:6599–606. https://doi.org/10.1021/acs.est.7b00997.

Daigo I, Fujimura L, Hayashi H, Yamasue E, Ohta S, Huy TD, et al. Quantifying the total amounts of tramp elements associated with carbon steel production in Japan. ISIJ International 2017;57:388–93. https://doi.org/10.2355/isijinternational.ISIJINT-2016-500.

Daigo I, Goto Y. Comparison of Tramp Element Contents of Steel Bars from Japan and China. ISIJ International 2015;55:2027–32. https://doi.org/10.2355/isijinternational.ISIJINT-2015-166.

Dantas PV, Sabino da Silva W, Cordeiro LC, Carvalho CB. A comprehensive review of model compression techniques in machine learning. Applied Intelligence 2024;54:11804–44. https://doi.org/10.1007/s10489-024-05747-w.

Darma R, Wulan W. From Nearby Stars to Solar Kinematics: New Insight from Gaia DR2 Catalogue 2021;4.

Davies DL, Bouldin DW. A Cluster Separation Measure. IEEE Trans Pattern Anal Mach Intell 1979;PAMI-1:224–7. https://doi.org/10.1109/TPAMI.1979.4766909.

Davis JR. Carbon and Alloy Steels, 1996.

Demir-Kavuk O, Kamada M, Akutsu T, Knapp E-W. Prediction using step-wise L1, L2 regularization and feature selection for small data sets with large number of features. BMC Bioinformatics 2011;12:412. https://doi.org/10.1186/1471-2105-12-412.

Devi R, Thigarasu V. A Semantic Deduplication of Temporal Dynamic Records from Multiple Web Databases. Indian J Sci Technol 2015;8. https://doi.org/10.17485/ijst/2015/v8i34/75103.

Dieter GE, Bacon D. Mechanical metallurgy. vol. 3. McGraw-hill New York; 1976.

Ding C, He X. K-Means Clustering via Principal Component Analysis. Proceedings of the Twenty-First International Conference on Machine Learning, New York, NY, USA: Association for Computing Machinery; 2004, p. 29. https://doi.org/10.1145/1015330.1015408.

Dou B, Zhu Z, Merkurjev E, Ke L, Chen L, Jiang J, et al. Machine Learning Methods for Small Data Challenges in Molecular Science. Chem Rev 2023;123:8736–80. https://doi.org/10.1021/acs.chemrev.3c00189.

Dubey SR, Singh SK, Chaudhuri BB. Activation functions in deep learning: A comprehensive survey and benchmark. Neurocomputing 2022;503:92–108. https://doi.org/https://doi.org/10.1016/j.neucom.2022.06.111.

Dundu M. Evolution of stress–strain models of stainless steel in structural engineering applications. Constr Build Mater 2018;165:413–23. https://doi.org/https://doi.org/10.1016/j.conbuildmat.2018.01.008.

Durmaz A, Potu S, Romich D, Möller J, Nützel R. Microstructure quality control of steels using deep learning. Front Mater 2023;10. https://doi.org/10.3389/fmats.2023.1222456.

Dutta S. Different types and new applications of stainless steel. Stainless Steel 2018;62:86–91.

Dworak S, Rechberger H, Fellner J. How will tramp elements affect future steel recycling in Europe? – A dynamic material flow model for steel in the EU-28 for the period 1910 to 2050. Resour Conserv Recycl 2022;179. https://doi.org/10.1016/j.resconrec.2021.106072.

Ester M, Kriegel H-P, Sander J, Xu X. A density-based algorithm for discovering clusters in large spatial databases with noise. kdd, vol. 96, 1996a, p. 226–31.

Ester M, Kriegel H-P, Sander J, Xu X. A density-based algorithm for discovering clusters in large spatial databases with noise. Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, AAAI Press; 1996b, p. 226–31.

Fahim A. A varied density-based clustering algorithm. J Comput Sci 2023;66:101925. https://doi.org/https://doi.org/10.1016/j.jocs.2022.101925.

Fahim A. K and starting means for k-means algorithm. J Comput Sci 2021;55:101445. https://doi.org/https://doi.org/10.1016/j.jocs.2021.101445.

Fanca A, Puscasiu A, Gota D-I, Valean H. Recommendation Systems with Machine Learning. 2020 21th International Carpathian Control Conference (ICCC), 2020, p. 1–6. https://doi.org/10.1109/ICCC49264.2020.9257290.

Feng Z, Gs Z. The Role of Molybdenum in Steel Alloys: Enhancing Performance and Applications. Advanced Materials Science Research Adv Mate Sci Res 2024;7:233–4. https://doi.org/10.37532/aaasmr.2024.7(6).233-234.

Fischer CC, Tibbetts KJ, Morgan D, Ceder G. Predicting crystal structure by merging data mining with quantum mechanics. Nat Mater 2006;5:641–6. https://doi.org/10.1038/nmat1691.

Fitriya LA, Purboyo TW, Prasasti AL. A review of data compression techniques. International Journal of Applied Engineering Research 2017;12:8956–63.

Fränti P, Sieranoja S. How much can k-means be improved by using better initialization and repeats? Pattern Recognit 2019;93:95–112. https://doi.org/https://doi.org/10.1016/j.patcog.2019.04.014.

Freund Y, Schapire RE. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. J Comput Syst Sci 1997;55:119–39. https://doi.org/https://doi.org/10.1006/jcss.1997.1504.

Frie C, Riza Durmaz A, Eberl C. Exploration of materials fatigue influence factors using interpretable machine learning. Fatigue Fract Eng Mater Struct 2024;47:2752–73. https://doi.org/10.1111/ffe.14315.

Friedman J. Greedy Function Approximation: A Gradient Boosting Machine. The Annals of Statistics 2000;29. https://doi.org/10.1214/aos/1013203451.

Gagolewski M, Bartoszuk M, Cena A. Genie: A new, fast, and outlier-resistant hierarchical clustering algorithm. Inf Sci (N Y) 2016;363:8–23. https://doi.org/https://doi.org/10.1016/j.ins.2016.05.003.

Gao H, Liu J, Daigo I. Methodology development for estimating the impact of restriction factors to promote national steel recycling. Resour Conserv Recycl 2025;215:108052. https://doi.org/https://doi.org/10.1016/j.resconrec.2024.108052.

García Gutiérrez I, Pina C, Tobajas R, Elduque D. Incorporating composition into life cycle assessment of steel grades. J Clean Prod 2024;472:143538. https://doi.org/https://doi.org/10.1016/j.jclepro.2024.143538.

Glorot X, Bordes A, Bengio Y. Deep Sparse Rectifier Neural Networks. vol. 15. 2010.

Gola J, Britz D, Staudt T, Winter M, Schneider AS, Ludovici M, et al. Advanced microstructure classification by data mining methods. Comput Mater Sci 2018;148:324–35. https://doi.org/https://doi.org/10.1016/j.commatsci.2018.03.004.

Gonsalves T, Upadhyay J. Chapter Eight - Integrated deep learning for self-driving robotic cars. In: Shaw RN, Ghosh A, Balas VE, Bianchini M, editors. Artificial Intelligence for Future Generation Robotics, Elsevier; 2021, p. 93–118. https://doi.org/https://doi.org/10.1016/B978-0-323-85498-6.00010-1.

Goyal V, Yadav A, Mukherjee R. Performance Evaluation of Machine Learning and Deep Learning Models for Temperature Prediction in Poultry Farming. 2022 10th International Conference on Emerging Trends in Engineering and Technology - Signal and Information Processing (ICETET-SIP-22), 2022, p. 1–6. https://doi.org/10.1109/ICETET-SIP-2254415.2022.9791771.

Grigoryan H. A Stock Market Prediction Method Based on Support Vector Machines (SVM) and Independent Component Analysis (ICA). Database Systems Journal 2016;7.

Groza V, Udrescu M, Bozdog A, Udrescu L. Drug Repurposing Using Modularity Clustering in Drug-Drug Similarity Networks Based on Drug–Gene Interactions. Pharmaceutics 2021;13. https://doi.org/10.3390/pharmaceutics13122117.

Guide SD. Structural stainless steel. American Institute of Steel Construction: Chicago, IL, USA 2013.

Guo S, Yu J, Liu X, Wang C, Jiang Q. A predicting model for properties of steel using the industrial big data based on machine learning. Comput Mater Sci 2019;160:95–104. https://doi.org/10.1016/j.commatsci.2018.12.056.

Hagedorn W, Jäger S, Wieczorek L, Kronenberg P, Greiff K, Weber S, et al. More than recycling – The potential of the circular economy shown by a case study of the metal working industry. J Clean Prod 2022;377:134439. https://doi.org/https://doi.org/10.1016/j.jclepro.2022.134439.

Hagelstein K. Globally sustainable manganese metal production and use. J Environ Manage 2009;90:3736–40. https://doi.org/https://doi.org/10.1016/j.jenvman.2008.05.025.

Halkidi M. Hierarchial Clustering. In: LIU L, ÖZSU MT, editors. Encyclopedia of Database Systems, Boston, MA: Springer US; 2009, p. 1291–4. https://doi.org/10.1007/978-0-387-39940-9_604.

Hall JN, Fekete JR. 2 - Steels for auto bodies: A general overview. In: Rana R, Singh SB, editors. Automotive Steels, Woodhead Publishing; 2017, p. 19–45. https://doi.org/https://doi.org/10.1016/B978-0-08-100638-2.00002-X.

Hamad M, Salih S. Removing Duplicate Records from Data Warehouse by Q-gram and Neural Network. International Journal of Engineering and Technology 2016;8:2374–82. https://doi.org/10.21817/ijet/2016/v8i5/160805121.

Hamming RW. Coding and information theory. Prentice-Hall, Inc.; 1986.

Han J, Kamber M, Pei J. 2 - Getting to Know Your Data. In: Han J, Kamber M, Pei J, editors. Data Mining: Concepts and Techniques (Third Edition), Boston: Morgan Kaufmann; 2012, p. 39–82. https://doi.org/https://doi.org/10.1016/B978-0-12-381479-1.00002-2.

Happonen A, Usmani U, Watada J. A Review of Unsupervised Machine Learning Frameworks for Anomaly Detection in Industrial Applications, 2022, p. 158–89. https://doi.org/10.1007/978-3-031-10464-0_11.

Hapuwatte B, Last N, Karsli S, Aher G, Morris K, Ferrero V. From Principles to Practice: Product Design Foundations for a Circular Economy. 2025. https://doi.org/10.1115/IMECE2024-146537.

Harris J. Engineering metallurgy: Part 1 applied physical metallurgy 1994.

Harwarth M, Brauer A, Huang Q, Pourabdoli M, Mola J. Influence of Carbon on the Microstructure Evolution and Hardness of Fe–13Cr–xC (x = 0–0.7 wt.%) Stainless Steel. Materials 2021;14:5063. https://doi.org/10.3390/ma14175063.

Hasan MN, Binte Malek M, Begum AA, Rahman M, Nurul M, Mollah H. medicina Assessment of Drugs Toxicity and Associated Biomarker Genes Using Hierarchical Clustering. Medicina (B Aires) 2019;55:451. https://doi.org/10.3390/medicina55080451.

Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition (Springer Series in Statistics). 2009.

Hayatu Hassan I, Abdullahi M, Ahmad B, Ali Y. K-MEANS CLUSTERING ALGORITHM BASED CLASSIFICATION OF SOIL FERTILITY IN NORTH WEST NIGERIA. FUDMA Journal of Sciences 2020;Vol. 4:780–7. https://doi.org/10.33003/fjs-2020-0402-363.

Hilal W, Gadsden SA, Yawney J. Financial Fraud: A Review of Anomaly Detection Techniques and Recent Advances. Expert Syst Appl 2022;193:116429. https://doi.org/https://doi.org/10.1016/j.eswa.2021.116429.

Hoerl AE, Kennard RW. Ridge Regression: Biased Estimation for Nonorthogonal Problems. Technometrics 1970;12:55–67. https://doi.org/10.2307/1267351.

Hooke R. Lectures “de Potentia Restitutiva”, Or of Spring, Explaining the Power of Springing Bodies, to which are Added Some Collections ... by Robert Hooke ... J. Martyn; 1678.

Hovhannisyan H, Qi W, Lu K, Yang R, Wang J. Whispers in the cloud storage: A novel cross-user deduplication-based covert channel design. Peer Peer Netw Appl 2018;11:277–86. https://doi.org/10.1007/s12083-016-0483-y.

Hozumi Y, Wang R, Yin C, Wei G-W. UMAP-assisted K-means clustering of large-scale SARS-CoV-2 mutation datasets. Comput Biol Med 2021;131:104264. https://doi.org/https://doi.org/10.1016/j.compbiomed.2021.104264.

Huang A, Wang K, Zhao Y, Wang W, Xicheng W, Peng J. Effect of Copper Addition on the Formability of 304L Austenitic Stainless Steel. J Mater Eng Perform 2022;32. https://doi.org/10.1007/s11665-022-07367-2.

Huffman DA. A Method for the Construction of Minimum-Redundancy Codes. Proceedings of the IRE 1952;40:1098–101. https://doi.org/10.1109/JRPROC.1952.273898.

Hwang R-C, Chen Y-J, Huang H-C. Artificial Intelligent Analyzer for Mechanical Properties of Rolled Steel Bar by Using Neural Networks. Expert Syst Appl 2010;37:3136–9. https://doi.org/10.1016/j.eswa.2009.09.069.

Ikotun AM, Ezugwu AE, Abualigah L, Abuhaija B, Heming J. K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. Inf Sci (N Y) 2023;622:178–210. https://doi.org/https://doi.org/10.1016/j.ins.2022.11.139.

Imam M, Ueji R, Fujii H. Effect of online rapid cooling on microstructure and mechanical properties of friction stir welded medium carbon steel. J Mater Process Technol 2016;230:62–71.

Islam T, Rashed HMMA. Classification and Application of Plain Carbon Steels. Reference Module in Materials Science and Materials Engineering, Elsevier; 2019. https://doi.org/https://doi.org/10.1016/B978-0-12-803581-8.10268-1.

Jagadish H, Madar J, Ng R. Semantic Compression and Pattern Extraction with Fascicles. 1999.

Jalalian S. Rationalised Steel Grade Dataset Using the K-Means Reduction Process (KMRP) 2026. https://doi.org/10.5281/zenodo.18259932.

Jalalian S. Carbon and Stainless Steel Grades 2025. https://doi.org/10.5281/zenodo.17704410.

Jammulamadaka H, Pisupati S V. A Critical Review of Extraction Methods for Vanadium from Petcoke Ash. Fuels 2023;4:58–74. https://doi.org/10.3390/fuels4010005.

Javaid A, Essadiqi E. Final report on scrap management, sorting and classification of steel 2003.

Jayaprakash C, Damodaran BB, V S, Soman KP. Dimensionality Reduction of Hyperspectral Images for Classification using Randomized Independent Component Analysis. 2018 5th International Conference on Signal Processing and Integrated Networks (SPIN), 2018, p. 492–6. https://doi.org/10.1109/SPIN.2018.8474266.

Jeffus L. Welding: Principles and Applications. Cengage Learning; 2016.

Jeon J, Seo N, Jung J-G, Kim H-S, Son SB, Lee S-J. Prediction and mechanism explain of austenite-grain growth during reheating of alloy steel using XAI. Journal of Materials Research and Technology 2022;21:1408–18. https://doi.org/https://doi.org/10.1016/j.jmrt.2022.09.119.

Jeon J, Seo N, Jung J-G, Son SB, Lee S-J. Analysis of Prediction Mechanisms and Feature Importance of Martensite Start Temperature of Alloy Steel via Explainable Artificial Intelligence. Mater Trans 2023;64:2196–201. https://doi.org/10.2320/matertrans.MT-MI2022004.

Jeon J, Seo N, Son SB, Lee S-J, Jung M. Application of Machine Learning Algorithms and SHAP for Prediction and Feature Analysis of Tempered Martensite Hardness in Low-Alloy Steels. Metals (Basel) 2021;11. https://doi.org/10.3390/met11081159.

Jia T, Liu Z, Hu H, Wang G. The Optimal Design for the Production of Hot Rolled Strip with “Tight Oxide Scale” by Using Multi-objective Optimization. ISIJ International 2011;51:1468–73. https://doi.org/10.2355/isijinternational.51.1468.

Jiang L, Wang A, Tian N, Zhang W, Fan Q. BP Neural Network of Continuous Casting Technological Parameters and Secondary Dendrite Arm Spacing of Spring Steel. Journal of Iron and Steel Research International 2011;18:25–9. https://doi.org/10.1016/S1006-706X(11)60099-X.

Jimoh RG, Yusuf RM, Olatunde YO, Kayode SY. Application of dimensionality reduction on classification of colon cancer using ICA and K-NN algorithm. Anale Ser Informatică 2018;6:55–9.

Jin F, Wu H, Liu Y, Zhao J, Wang W. Varying-scale HCA-DBSCAN-based anomaly detection method for multi-dimensional energy data in steel industry. Inf Sci (N Y) 2023;647:119479. https://doi.org/https://doi.org/10.1016/j.ins.2023.119479.

Jin X, Han J. K-Means Clustering. In: Sammut C, Webb GI, editors. Encyclopedia of Machine Learning, Boston, MA: Springer US; 2010, p. 563–4. https://doi.org/10.1007/978-0-387-30164-8_425.

John JM, Shobayo O, Ogunleye B. An Exploration of Clustering Algorithms for Customer Segmentation in the UK Retail Market. Analytics 2023;2:809–23. https://doi.org/10.3390/analytics2040042.

Jolliffe I, Cadima J. Principal component analysis: A review and recent developments. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 2016;374:20150202. https://doi.org/10.1098/rsta.2015.0202.

José P, Nieto G, Manuel González Suárez V, Carlos Álvarez Antón J, Mayo Bayón R, Ángel J, et al. A New Predictive Model of Centerline Segregation in Continuous Cast Steel Slabs by Using Multivariate Adaptive Regression Splines Approach. Materials 2015;8:3562–83. https://doi.org/10.3390/ma8063562.

Kalpakjian S, Schmid S, Sekar V. Manufacturing Engineering and Technology. 2013.

Kamel M, Kyla T, Scott A, Stephen F, Rachel F, Jui-Hua H, et al. Unlocking the Potential of Clustering and Classification Approaches: Navigating Supervised and Unsupervised Chemical Similarity. Environ Health Perspect 2025;132:085002. https://doi.org/10.1289/EHP14001.

Kangas P, Chai G. Use of Advanced Austenitic and Duplex Stainless Steels for Applications in Oil & Gas and Process Industry. Adv Mat Res 2013;794:645–69.

Kansal T, Bahuguna S, Singh V, Choudhury T. Customer Segmentation using K-means Clustering. 2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS), 2018, p. 135–9. https://doi.org/10.1109/CTEMS.2018.8769171.

Kaur R, Chana I, Bhattacharya J. Data deduplication techniques for efficient cloud storage management: a systematic review. J Supercomput 2018;74:2035–85. https://doi.org/10.1007/s11227-017-2210-8.

Keehan E, Karlsson L, Andrén H-O. Influence of carbon, manganese and nickel on microstructure and properties of strong steel weld metals: Part 1 – Effect of nickel content. Science and Technology of Welding & Joining 2006;11:1–8. https://doi.org/10.1179/174329306X77830.

Khan B, Rauf A, Javed H, Khusro S. Removing Fully and Partially Duplicated Records through K-Means Clustering. International Journal of Engineering and Technology 2012;4:750–4. https://doi.org/10.7763/IJET.2012.V4.477.

Kiangala SK, Wang Z. An effective adaptive customization framework for small manufacturing plants using extreme gradient boosting-XGBoost and random forest ensemble learning algorithms in an Industry 4.0 environment. Machine Learning with Applications 2021;4:100024. https://doi.org/https://doi.org/10.1016/j.mlwa.2021.100024.

Kim H, Inoue J, Kasuya T. Unsupervised microstructure segmentation by mimicking metallurgists' approach to pattern recognition. Sci Rep 2020;10:17835. https://doi.org/10.1038/s41598-020-74935-8.

Kim J, Sovacool BK, Bazilian M, Griffiths S, Lee Junghwan, Yang M, et al. Decarbonizing the iron and steel industry: A systematic review of sociotechnical systems, technological innovations, and policy options. Energy Res Soc Sci 2022;89:102565. https://doi.org/https://doi.org/10.1016/j.erss.2022.102565.

Kim K-S, Kang J-H, Kim S-J. Effects of carbon and nitrogen on precipitation and tensile behavior in 15Cr-15Mn-4Ni austenitic stainless steels. Materials Science and Engineering: A 2018;712:114–21. https://doi.org/https://doi.org/10.1016/j.msea.2017.11.099.

Kingma D, Ba J. Adam: A Method for Stochastic Optimization. International Conference on Learning Representations 2014.

Kishor G, Mugada KK, Mahto RP, Badheka V. Effect of Process Conditions on Mechanical and Metallurgical Properties of Wire Arc Additively Manufactured 316L Stainless Steel. J Mater Eng Perform 2025;34:8431–51. https://doi.org/10.1007/s11665-024-10033-4.

Kishore K, Sheikh MN, Hadi MNS. A critical analysis of electric arc furnace (EAF) slag for sustainable geopolymer concrete production. Materials Today Sustainability 2025;29:101064. https://doi.org/https://doi.org/10.1016/j.mtsust.2024.101064.

Kitahara AR, Holm EA. Microstructure Cluster Analysis with Transfer Learning and Unsupervised Learning. Integr Mater Manuf Innov 2018;7:148–56. https://doi.org/10.1007/s40192-018-0116-9.

Klar E, Samal PK. Powder Metallurgy Stainless Steels: Processing, Microstructures, and Properties, 2007.

Kolajo T, Daramola O. Human-centric and semantics-based explainable event detection: a survey. Artif Intell Rev 2023;56:1–40. https://doi.org/10.1007/s10462-023-10525-0.

Kolotyrkin YM, Reformatskaya II, Pan`shin EA. The mechanism for increasing pitting resistance of stainless steels by the addition of molybdenum 1994;30.

Korkmaz S, Zararsiz G, Goksuluk D. Drug/nondrug classification using Support Vector Machines with various feature selection strategies. Comput Methods Programs Biomed 2014;117:51–60. https://doi.org/https://doi.org/10.1016/j.cmpb.2014.08.009.

Koumarelas I, Jiang L, Naumann F. Data Preparation for Duplicate Detection. J Data and Information Quality 2020;12. https://doi.org/10.1145/3377878.

Krauss G. Steels heat treatment and processing principles ASM International. Materials Park 1993.

Kussul E, Baidyk T, Kasatkina L, Lukovich V. Rosenblatt Perceptrons for handwritten digit recognition. vol. 2. 2001. https://doi.org/10.1109/IJCNN.2001.939589.

Lalam S, Tiwari PK, Sahoo S, Dalal AK. Online prediction and monitoring of mechanical properties of industrial galvanised steel coils using neural networks. Ironmaking & Steelmaking 2019;46:89–96. https://doi.org/10.1080/03019233.2017.1342424.

Lecun Y, Bottou L, Orr G, Müller K-R. Efficient BackProp 2000.

Lee J-W, Park C, Do Lee B, Park J, Goo NH, Sohn K-S. A machine-learning-based alloy design platform that enables both forward and inverse predictions for thermo-mechanically controlled processed (TMCP) steel alloys. Sci Rep 2021;11:11012. https://doi.org/10.1038/s41598-021-90237-z.

Li F, Wu J, Dong F, Lin J, Sun G, Chen H, et al. Ensemble Machine Learning Systems for the Estimation of Steel Quality Control. 2018 IEEE International Conference on Big Data (Big Data), 2018, p. 2245–52. https://doi.org/10.1109/BigData.2018.8622583.

Liang X, Fu H, Cui M, Liu G. materials Effect of Intercritical Tempering Temperature on Microstructure Evolution and Mechanical Properties of High Strength and Toughness Medium Manganese Steel 2022. https://doi.org/10.3390/ma15062162.

Likas A, Vlassis N, J. Verbeek J. The global k-means clustering algorithm. Pattern Recognit 2003;36:451–61. https://doi.org/10.1016/S0031-3203(02)00060-2.

Lin L-C, Berger AH, Martin RL, Kim J, Swisher JA, Jariwala K, et al. In silico screening of carbon-capture materials. Nat Mater 2012;11:633–41.

Liu G, Singha M, Pu L, Neupane P, Feinstein J, Wu H-C, et al. GraphDTI: A robust deep learning predictor of drug-target interactions from multiple heterogeneous data. J Cheminform 2021;13:58. https://doi.org/10.1186/s13321-021-00540-0.

Liu X, Bao Y, Zhao L, Gu C. Establishment and Application of Steel Composition Prediction Model Based on t-Distributed Stochastic Neighbor Embedding (t-SNE) Dimensionality Reduction Algorithm. Journal of Sustainable Metallurgy 2024;10:509–24. https://doi.org/10.1007/s40831-024-00798-2.

López-Rubio E, Palomo EJ, Ortega-Zamorano F. Unsupervised learning by cluster quality optimization. Inf Sci (N Y) 2018;436–437:31–55. https://doi.org/https://doi.org/10.1016/j.ins.2018.01.007.

Lu J, Chen L, Yin J, Huang T, Bi Y, Kong X, et al. Journal of Biomolecular Structure and Dynamics Identification of new candidate drugs for lung cancer using chemical-chemical interactions, chemical-protein interactions and a K-means clustering algorithm Identification of new candidate drugs for lung cancer using chemical-chemical interactions, chemical-protein interactions and a K-means clustering algorithm 2016. https://doi.org/10.1080/07391102.2015.1060161.

Lu Q, Lai Q, Chai Z, Wei X, Xiong X, Yi H, et al. Revolutionizing car body manufacturing using a unified steel metallurgy concept. Sci Adv 2025;7:eabk0176. https://doi.org/10.1126/sciadv.abk0176.

Lüdeke-Freund F, Gold S, Bocken NMP. A Review and Typology of Circular Economy Business Model Patterns. J Ind Ecol 2019;23:36–61. https://doi.org/https://doi.org/10.1111/jiec.12763.

Lundberg S, Lee S-I. A Unified Approach to Interpreting Model Predictions. 2017.

Lv Y, Ma T, Tang M, Cao J, Tian Y, Al-Dhelaan A, et al. An efficient and scalable density-based clustering algorithm for datasets with complex structures. Neurocomputing 2016;171:9–22. https://doi.org/https://doi.org/10.1016/j.neucom.2015.05.109.

Lyman T. Metals Handbook, American Society for Metals. 1948 1948.

M. WH. The Metallurgy of Iron and Steel. Nature 1922;110:507. https://doi.org/10.1038/110507a0.

Ma B, Yang C, Li A, Chi Y, Chen L. A Faster DBSCAN Algorithm Based on Self-Adaptive Determination of Parameters. Procedia Comput Sci 2023;221:113–20. https://doi.org/10.1016/j.procs.2023.07.017.

Van Der Maaten L, Hinton G. Visualizing data using t-SNE. Journal of Machine Learning Research 2008;9:2579–625.

Madhavan S, Sridharan NV, Mahanta TK, Vaithiyanathan S. Diagnosis of Surface Defects in Hot-Rolled Steel from Deep Learning Features Using Machine Learning Algorithms. Arab J Sci Eng 2024. https://doi.org/10.1007/s13369-024-09744-6.

Madhulatha TS. An overview on clustering methods. ArXiv Preprint ArXiv:12051117 2012.

Madugula SS, John L, Nagamani S, Gaur AS, Poroikov V V, Sastry GN. Molecular descriptor analysis of approved drugs using unsupervised learning for drug repurposing. Comput Biol Med 2021;138:104856. https://doi.org/https://doi.org/10.1016/j.compbiomed.2021.104856.

MakeItFrom.com. MakeItFrom 2023. https://www.makeitfrom.com/ (accessed June 3, 2023).

Malhat MG, Mousa HM, El-Sisi AB. Clustering of chemical data sets for drug discovery. 2014 9th International Conference on Informatics and Systems 2014:DEKM-11-DEKM-18.

Mandagere N, Zhou P, Smith MA, Uttamchandani S. Demystifying data deduplication. Proceedings of the ACM/IFIP/USENIX Middleware '08 Conference Companion, New York, NY, USA: Association for Computing Machinery; 2008, p. 12–7. https://doi.org/10.1145/1462735.1462739.

Martin J. Materials for engineering. Woodhead Publishing; 2006.

MatWeb. MatWeb 2022, Online Materials Information Resource. Https://WwwMatwebCom/ 2023. https://www.matweb.com/ (accessed July 20, 2023).

McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. Bull Math Biol 1990;52:99–115.

Mehta B, Höglund L, Mu W, Hedström P. Toward Scrap-Tolerant Steels: Investigating the Role of Cu and Sn Micro-Segregations on Solidification Microstructure and Cracking. Journal of Sustainable Metallurgy 2025;11. https://doi.org/10.1007/s40831-025-01093-4.

Meister D, Kaiser J, Brinkmann A, Cortes T, Kuhn M, Kunkel J. A study on data deduplication in HPC storage systems. Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis, Washington, DC, USA: IEEE Computer Society Press; 2012.

Meng C, Zeleznik OA, Thallinger GG, Kuster B, Gholami AM, Culhane AC. Dimension reduction techniques for the integrative analysis of multi-omics data. Brief Bioinform 2016;17:628–41. https://doi.org/10.1093/bib/bbv108.

Michler T. Austenitic Stainless Steels. Reference Module in Materials Science and Materials Engineering, Elsevier; 2016. https://doi.org/https://doi.org/10.1016/B978-0-12-803581-8.02509-1.

Mililitsky M, De Wispelaere N, Petrov R, Ramos JE, Reguly A, Hänninen H. Characterization of the mechanical properties of low-nickel austenitic stainless steels. Materials Science and Engineering: A 2008;498:289–95. https://doi.org/https://doi.org/10.1016/j.msea.2008.08.012.

Milligan GW, Cooper MC. A study of standardization of variables in cluster analysis. J Classif 1988;5:181–204. https://doi.org/10.1007/BF01897163.

Millner G, Mücke M, Romaner L, Scheiber D. Machine learning mechanical properties of steel sheets from an industrial production route. Materialia (Oxf) 2023;30. https://doi.org/10.1016/j.mtla.2023.101810.

Mohseni S, Zarei N, Ragan E. A Multidisciplinary Survey and Framework for Design and Evaluation of Explainable AI Systems. ACM Trans Interact Intell Syst 2021;11:1–45. https://doi.org/10.1145/3387166.

Montesinos López OA, Montesinos López A, Crossa Jose. Fundamentals of Artificial Neural Networks and Deep Learning. In: Montesinos López OA, Montesinos López A, Crossa José, editors. Multivariate Statistical Machine Learning Methods for Genomic Prediction, Cham: Springer International Publishing; 2022, p. 379–425. https://doi.org/10.1007/978-3-030-89010-0_10.

Müllner D. Modern hierarchical, agglomerative clustering algorithms 2011.

Murray JW, Jin X, Cleaver CJ, Azevedo JMC, Liao Z, Zhou W, et al. A review of principles and options for the re-use of machining chips by solid, semi-solid or melt-based processing. J Mater Process Technol 2024;331:118514. https://doi.org/https://doi.org/10.1016/j.jmatprotec.2024.118514.

Murtagh F, Contreras P. Algorithms for hierarchical clustering: An overview. Wiley Interdisc Rew: Data Mining and Knowledge Discovery 2012;2:86–97. https://doi.org/10.1002/widm.53.

Murtagh F, Legendre P. Ward's Hierarchical Clustering Method: Clustering Criterion and Agglomerative Algorithm 2011.

Nair V, Hinton GE. Rectified linear units improve restricted boltzmann machines. Proceedings of the 27th international conference on machine learning (ICML-10), 2010, p. 807–14.

Nakamura S, Kondo Y, Kagawa S, Matsubae K, Nakajima K, Nagasaka T. MaTrace: Tracing the fate of materials over time and across products in open-loop recycling. Environ Sci Technol 2014;48:7207–14. https://doi.org/10.1021/es500820h.

Nguyen H, Cao H, Nguyen V, Pham D. Evaluation of Explainable Artificial Intelligence: SHAP, LIME, and CAM. 2021.

Nielsen C V, Martins PAF. Chapter 3 - Finite element simulation: A user's perspective**Chris V. Nielsen (Technical University of Denmark, Lyngby, Denmark) and Paulo A.F. Martins (IDMEC, Instituto Superior Técnico, University of Lisbon, Portugal). In: Nielsen C V, Martins PAF, editors. Metal Forming, Academic Press; 2021, p. 109–80. https://doi.org/https://doi.org/10.1016/B978-0-323-85255-5.00011-X.

Nisbet R, Miner G, Yale K. Chapter 9 - Classification. In: Nisbet R, Miner G, Yale K, editors. Handbook of Statistical Analysis and Data Mining Applications (Second Edition), Boston: Academic Press; 2018, p. 169–86. https://doi.org/https://doi.org/10.1016/B978-0-12-416632-5.00009-8.

Nixon MS, Aguado AS. 12 - Distance, classification and learning. In: Nixon MS, Aguado AS, editors. Feature Extraction and Image Processing for Computer Vision (Fourth Edition), Academic Press; 2020, p. 571–604. https://doi.org/https://doi.org/10.1016/B978-0-12-814976-8.00012-9.

Nogara J, Zarrouk SJ. Corrosion in geothermal environment Part 2: Metals and alloys. Renewable and Sustainable Energy Reviews 2018;82:1347–63. https://doi.org/https://doi.org/10.1016/j.rser.2017.06.091.

Nouri A, Wen C. 3 - Stainless steels in orthopedics. In: Wen C, editor. Structural Biomaterials, Woodhead Publishing; 2021, p. 67–101. https://doi.org/https://doi.org/10.1016/B978-0-12-818831-6.00008-2.

Nsengiyumva W, Zhong S, Lin J, Zhang Q, Zhong J, Huang Y. Advances, limitations and prospects of nondestructive testing and evaluation of thick composites and sandwich structures: A state-of-the-art review. Compos Struct 2021;256:112951. https://doi.org/https://doi.org/10.1016/j.compstruct.2020.112951.

Ohno H, Matsubae K, Nakajima K, Kondo Y, Nakamura S, Fukushima Y, et al. Optimal Recycling of Steel Scrap and Alloying Elements: Input-Output based Linear Programming Method with Its Application to End-of-Life Vehicles in Japan. Environ Sci Technol 2017;51:13086–94. https://doi.org/10.1021/acs.est.7b04477.

Ohno H, Matsubae K, Nakajima K, Kondo Y, Nakamura S, Nagasaka T. Toward the efficient recycling of alloying elements from end of life vehicle steel scrap. Resour Conserv Recycl 2015a;100:11–20. https://doi.org/https://doi.org/10.1016/j.resconrec.2015.04.001.

Ohno H, Matsubae K, Nakajima K, Kondo Y, Nakamura S, Nagasaka T. Toward the efficient recycling of alloying elements from end of life vehicle steel scrap. Resour Conserv Recycl 2015b;100:11–20. https://doi.org/https://doi.org/10.1016/j.resconrec.2015.04.001.

Omar ZA, Bakar MAA, Zamzuri ZH, Ariff NM. Duplicate Detection Using Unsupervised Random Forests: A Preliminary Analysis. 2022 3rd International Conference on Artificial Intelligence and Data Sciences (AiDAS), 2022, p. 66–71. https://doi.org/10.1109/AiDAS56890.2022.9918724.

Organization WH. Vanadium-Environmental Health Criteria 81 1984.

Ormesher I. Duplicate Detection with GenAI. 2024. https://doi.org/10.48550/arXiv.2406.15483.

Owen C V, Schmidt FA, Carlson ON. Effect of copper, tin, nickel, and chromium on mechanical properties of low-carbon steel. Metals Technology 1976;3:441–5. https://doi.org/10.1179/030716976803392222.

Ozerdem MS, Kolukisa S. Artificial neural network approach to predict the mechanical properties of Cu–Sn–Pb–Zn–Ni cast alloys. Mater Des 2009;30:764–9. https://doi.org/https://doi.org/10.1016/j.matdes.2008.05.019.

P. Švec, L. Frischerová, J. David. Usage of clustering methods for sequence plan optimization in steel production. Metalurgija 2016;55:485–8.

Padmanaban M, Radha R, Devkunvar S, Bhatt Vaishnav N. PSO Algorithm to Select Subsets of Q-Gram Features for Record Duplicate Detection. vol. 82. 2013.

Paluszek M, Thomas S. An Overview of Machine Learning, 2017, p. 3–15. https://doi.org/10.1007/978-1-4842-2250-8_1.

Papavinasam S. Chapter 3 - Materials. In: Papavinasam S, editor. Corrosion Control in the Oil and Gas Industry, Boston: Gulf Professional Publishing; 2014, p. 133–77. https://doi.org/https://doi.org/10.1016/B978-0-12-397022-0.00003-0.

Park S, Kim C, Kang N. Artificial Neural Network-Based Modelling for Yield Strength Prediction of Austenitic Stainless-Steel Welds. Applied Sciences 2024;14. https://doi.org/10.3390/app14104224.

Park Seobin, Kayani SH, Euh K, Seo E, Kim H, Park Sangeun, et al. High strength aluminum alloys design via explainable artificial intelligence. J Alloys Compd 2022;903. https://doi.org/10.1016/j.jallcom.2022.163828.

Patel S V, Jokhakar VN. A random forest based machine learning approach for mild steel defect diagnosis. 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC) 2016:1–8.

Péalat C, Bouleux G, Cheutet V. Improved time series clustering based on new geometric frameworks. Pattern Recognit 2022;124:108423. https://doi.org/https://doi.org/10.1016/j.patcog.2021.108423.

Pearl J. The limitations of opaque learning machines. Possible Minds 2019;25:13–9.

Pedregosa F, Varoquaux G, Gramfort A and M V., Thirion B and GO and BM and PP and WR and DV and VJ and PA and CD and BM and PM and DE. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research 2011;12:2825–30.

Pedrycz W, Cios K, Swiniarski R, Kurgan L. Unsupervised Learning: Association Rules, 2007, p. 289–306. https://doi.org/10.1007/978-0-387-36795-8_10.

Peña JM, Lozano JA, Larrañaga P. An empirical comparison of four initialization methods for the K-Means algorithm. Pattern Recognit Lett 1999;20:1027–40.

Perafan-Lopez JC, Ferrer-Gregory VL, Nieto-Londoño C, Sierra-Pérez J. Performance Analysis and Architecture of a Clustering Hybrid Algorithm Called FA+GA-DBSCAN Using Artificial Datasets. Entropy 2022;24. https://doi.org/10.3390/e24070875.

Pink CM. Chapter 11 - Forensic Ancestry Assessment Using Cranial Nonmetric Traits Traditionally Applied to Biological Distance Studies. In: Pilloud MA, Hefner JT, editors. Biological Distance Analysis, San Diego: Academic Press; 2016, p. 213–30. https://doi.org/https://doi.org/10.1016/B978-0-12-801966-5.00011-1.

Pise NN, Kulkarni P. A Survey of Semi-Supervised Learning Methods. 2008 International Conference on Computational Intelligence and Security, vol. 2, 2008, p. 30–4. https://doi.org/10.1109/CIS.2008.204.

Prakash A. Pre-processing techniques for preparing clean and high-quality data for diabetes prediction. International Journal of Research Publication and Reviews 2024;5:458–65. https://doi.org/10.55248/gengpi.5.0224.0412.

Pu IM. Chapter 9 - Audio compression. In: Pu IM, editor. Fundamental Data Compression, Oxford: Butterworth-Heinemann; 2006, p. 171–88. https://doi.org/https://doi.org/10.1016/B978-075066310-6/50012-X.

Raabe D. The Materials Science behind Sustainable Metals and Alloys. Chem Rev 2023a;123:2436–608.

Raabe D. The Materials Science behind Sustainable Metals and Alloys. Chem Rev 2023b;123:2436–608. https://doi.org/10.1021/acs.chemrev.2c00799.

De Ras K, Van de Vijver R, Galvita V V, Marin GB, Van Geem KM. Carbon capture and utilization in the steel industry: challenges and opportunities for chemical engineering. Curr Opin Chem Eng 2019;26:81–7.

Reck B, Graedel T. Challenges in Metal Recycling. Science 2012;337:690–5. https://doi.org/10.1126/science.1217501.

Refaeilzadeh P, Tang L, Liu H. Cross-Validation. In: LIU L, ÖZSU MT, editors. Encyclopedia of Database Systems, Boston, MA: Springer US; 2009, p. 532–8. https://doi.org/10.1007/978-0-387-39940-9_565.

Reijnders L. Conserving functionality of relatively rare metals associated with steel life cycles: a review. J Clean Prod 2016;131:76–96. https://doi.org/https://doi.org/10.1016/j.jclepro.2016.05.073.

Ren Y, Wang N, Li M, Xu Z. Deep density-based image clustering. Knowl Based Syst 2020;197:105841. https://doi.org/https://doi.org/10.1016/j.knosys.2020.105841.

Ribeiro M, Singh S, Guestrin C. "Why Should I Trust You?": Explaining the Predictions of Any Classifier. 2016. https://doi.org/10.18653/v1/N16-3020.

Rigatti SJ. Random Forest. J Insur Med 2017;47:31–9. https://doi.org/10.17849/insm-47-01-31-39.1.

Rocha A, Marques P. The Use of Machine Learning to Predict Steel Properties – A Review on the Latest Works, 2024. https://doi.org/10.5772/intechopen.1004639.

Rousseeuw PJ. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math 1987;20:53–65. https://doi.org/https://doi.org/10.1016/0377-0427(87)90125-7.

Rousseeuw PJ, Kaufman L. Finding groups in data: An introduction to cluster analysis. John Wiley & Sons Hoboken, New Jersey; 2009.

Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. Nature 1986;323:533–6. https://doi.org/10.1038/323533a0.

Sadegh-Zadeh S-A, Sadeghzadeh N, Soleimani O, Shiry Ghidary S, Movahedi S, Mousavi S-Y. Dimensionality reduction techniques for EEG-based emotional state classification. American Journal of Neurodegenerative Diseases 2024;13:23–33. https://doi.org/10.62347/ZWRY8401.

Salvetr P, Gokhman A, Svoboda M, Donik Č, Podstranská I, Kotous J, et al. Effect of Cu Alloying on Mechanical Properties of Medium-C Steel after Long-Time Tempering at 500 °C. Materials 2023;16. https://doi.org/10.3390/ma16062390.

Samuel AL. Some Studies in Machine Learning Using the Game of Checkers. IBM J Res Dev 1959;3:210–29. https://doi.org/10.1147/rd.33.0210.

Sandeep MS, Tiprak K, Kaewunruen S, Pheinsusom P, Pansuk W. Shear strength prediction of reinforced concrete beams using machine learning. Structures 2023;47:1196–211. https://doi.org/https://doi.org/10.1016/j.istruc.2022.11.140.

Santos I, Nieves J, Penya YK, Bringas PG. Machine-learning-based mechanical properties prediction in foundry production. 2009 ICCAS-SICE 2009:4536–41.

Satyam Kumar. 7 Ways to Handle Missing Values in Machine Learning. Towards Data Science 2020. https://towardsdatascience.com/ (accessed September 6, 2022).

Sayood K. Chapter 2 - Mathematical Preliminaries for Lossless Compression. In: Sayood K, editor. Introduction to Data Compression (Fifth Edition), Morgan Kaufmann; 2018, p. 11–39. https://doi.org/https://doi.org/10.1016/B978-0-12-809474-7.00002-1.

Sayood K. Dictionary Techniques. In: Sayood K, editor. Introduction to Data Compression (Fourth Edition), Boston: Morgan Kaufmann; 2012, p. 135–61. https://doi.org/https://doi.org/10.1016/B978-0-12-415796-5.00005-3.

Schmidhuber J. Deep learning in neural networks: An overview. Neural Networks 2015;61:85–117. https://doi.org/https://doi.org/10.1016/j.neunet.2014.09.003.

Schubert E, Sander J, Ester M, Kriegel HP, Xu X. DBSCAN Revisited, Revisited: Why and How You Should (Still) Use DBSCAN. ACM Trans Database Syst 2017;42. https://doi.org/10.1145/3068335.

Schulte R, Lechner M, Merklein M. Fundamental analysis For the application of hybrid semi-finished products in sheet-bulk metal forming. J Mater Process Technol 2020;283:116709. https://doi.org/https://doi.org/10.1016/j.jmatprotec.2020.116709.

Serajzadeh S, Karimi Taheri A. An investigation on the effect of carbon and silicon on flow behavior of steel. Mater Des 2002;23:271–6. https://doi.org/https://doi.org/10.1016/S0261-3069(01)00080-2.

S&G. Steel-grades.org 2024. https://www.steel-grades.com/ (accessed November 13, 2024).

Shahri HH, Shahri SH. Eliminating Duplicates in Information Integration: An Adaptive, Extensible Framework. IEEE Intell Syst 2006;21:63–71. https://doi.org/10.1109/MIS.2006.90.

Shapley LS. A value for n-person games 1953.

Sharma RC. Principles of heat treatment of steels. New Age International; 2003.

Singh M. Application of Steel in Automotive Industry, 2016.

Singh S, Bhadeshia HKDH, MacKay D, Carey HC, Martín ÍLS. Neural network analysis of steel plate processing. Ironmaking & Steelmaking 1998;25:355–65.

Sotoodeh K. Chapter Four - Corrosion study and material selection for cryogenic valves in an LNG plant. In: Sotoodeh K, editor. Cryogenic Valves for Liquefied Natural Gas Plants, Gulf Professional Publishing; 2022, p. 175–211. https://doi.org/https://doi.org/10.1016/B978-0-323-99584-9.00001-X.

Srikanth M, Asmatulu R. Chapter 8 - Nanotechnology Safety in the Construction and Infrastructure Industries. In: Asmatulu R, editor. Nanotechnology Safety, Amsterdam: Elsevier; 2013, p. 99–113. https://doi.org/https://doi.org/10.1016/B978-0-444-59438-9.00008-4.

Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. J Mach Learn Res 2014;15:1929–58.

Steorts R, Hall R, Fienberg S. SMERED: A Bayesian Approach to Graphical Record Linkage and De-duplication. J Am Stat Assoc 2014;111. https://doi.org/10.1080/01621459.2015.1105807.

Sterjovski Z, Nolan D, Carpenter K, Dunne D, Norrish J. Artificial Neural Network for Modeling the Mechanical Properties of Steels in Various Applications. J Mater Process Technol 2005;170:536–44. https://doi.org/10.1016/j.jmatprotec.2005.05.040.

Štrumbelj E, Kononenko I. Explaining prediction models and individual predictions with feature contributions. Knowl Inf Syst 2014;41:647–65.

Sugimoto K, Sawada Y. The Role of Alloyed Molybdenum in Austenitic Stainless Steels in the Inhibition of Pitting in Neutral Halide Solutions. Corrosion 1976;32:347–52. https://doi.org/10.5006/0010-9312-32.9.347.

Sutton RS, Barto AG. Reinforcement learning: An introduction. vol. 1. MIT press Cambridge; 1998.

Swartout WR. Explaining and justifying expert consulting programs. Proceedings of the 7th International Joint Conference on Artificial Intelligence - Volume 2, San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.; 1981, p. 815–23.

Swateelagna S, Singh M, Rahul MR. Explainable Machine Learning based approach for the design of new refractory high entropy alloys. Intermetallics (Barking) 2024;167. https://doi.org/10.1016/j.intermet.2024.108198.

Szandała T. Review and Comparison of Commonly Used Activation Functions for Deep Neural Networks. In: Bhoi AK, Mallick PK, Liu C-M, Balas VE, editors. Bio-inspired Neurocomputing, Singapore: Springer Singapore; 2021, p. 203–24. https://doi.org/10.1007/978-981-15-5495-7_11.

Teo P Ter, Zakaria SK, Salleh SZ, Taib MAA, Mohd Sharif N, Abu Seman A, et al. Assessment of Electric Arc Furnace (EAF) Steel Slag Waste's Recycling Options into Value Added Green Products: A Review. Metals (Basel) 2020;10. https://doi.org/10.3390/met10101347.

Thierry D, Persson D, Le Bozec N. Atmospheric Corrosion of Zinc and Zinc Alloyed Coated Steel. In: Wandelt K, editor. Encyclopedia of Interfacial Chemistry, Oxford: Elsevier; 2018, p. 55–78. https://doi.org/https://doi.org/10.1016/B978-0-12-409547-2.13431-6.

Tibshirani R, Walther G, Hastie T. Estimating the number of clusters in a data set via the gap statistic. J R Stat Soc Series B Stat Methodol 2001;63:411–23.

Tiwari T, Jalalian S, Mendis C, Eskin D. Classification of T6 Tempered 6XXX Series Aluminum Alloys Based on Machine Learning Principles. JOM 2023;75:4526–37. https://doi.org/10.1007/s11837-023-06025-9.

Tomala J, Urbaniec M. TOWARDS SUSTAINABLE DEVELOPMENT IN THE EUROPEAN UNION: A CRITICAL RAW MATERIALS PERSPECTIVE. Economics and Environment 2024;88. https://doi.org/10.34659/eis.2024.88.1.654.

Torgo L, Gama J. Regression by classification. In: Borges DL, Kaestner CAA, editors. Advances in Artificial Intelligence, Berlin, Heidelberg: Springer Berlin Heidelberg; 1996, p. 51–60.

Trautman B, Tóth L, Fabian R. Comparison of Tool Steels for Tube End Flanging. Acta Materialia Transylvanica 2023;5:93–8. https://doi.org/10.33924/amt-2022-02-09.

Tsafou K, Jensen LJ. Integrative Systems Biology. In: Bradshaw RA, Stahl PD, editors. Encyclopedia of Cell Biology, Waltham: Academic Press; 2016, p. 245–51. https://doi.org/https://doi.org/10.1016/B978-0-12-394447-4.40042-8.

Tu X, Fu C, Huang A, Chen H, Ding X. DBSCAN Spatial Clustering Analysis of Urban “Production–Living–Ecological” Space Based on POI Data: A Case Study of Central Urban Wuhan, China. Int J Environ Res Public Health 2022;19. https://doi.org/10.3390/ijerph19095153.

Tullis T, Albert B. Chapter 9 - Special Topics. In: Tullis T, Albert B, editors. Measuring the User Experience (Second Edition), Boston: Morgan Kaufmann; 2013, p. 209–36. https://doi.org/https://doi.org/10.1016/B978-0-12-415781-1.00009-1.

Tune XGBoost “min_child_weight” Parameter. XGBoost 2024. https://xgboosting.com/configure-xgboost-min_child_weight-parameter/ (accessed August 21, 2025).

Turing AM. Computing machinery and intelligence. Parsing the turing test, Springer; 2009, p. 23–65.

Tzortzis G, Likas A. The MinMax k-Means clustering algorithm. Pattern Recognit 2014;47:2505–16. https://doi.org/https://doi.org/10.1016/j.patcog.2014.01.015.

Uranga P, Shang C-J, Senuma T, Yang J-R, Guo A-M, Mohrbacher H. Molybdenum alloying in high-performance flat-rolled steel grades. Adv Manuf 2020;8:15–34. https://doi.org/10.1007/s40436-019-00285-y.

Valenzuela L, Pichara K. Unsupervised classification of variable stars. Mon Not R Astron Soc 2018;474:3259–72.

Valkenborg D, Rousseau A-J, Geubbelmans M, Burzykowski T. Unsupervised learning. American Journal of Orthodontics and Dentofacial Orthopedics 2023;163:877–82. https://doi.org/10.1016/j.ajodo.2023.04.001.

Vandeginste BGM, Massart DL, Buydens LMC, De Jong S, Lewi PJ, Smeyers-Verbeke J. Chapter 30 - Cluster analysis. In: Vandeginste BGM, Massart DL, Buydens LMC, De Jong S, Lewi PJ, Smeyers-Verbeke J, editors. Data Handling in Science and Technology, vol. 20, Elsevier; 1998, p. 57–86. https://doi.org/https://doi.org/10.1016/S0922-3487(98)80040-3.

Vapnik V. The Nature of Statistical Learning Theory. Statistics for Engineering and Information Science, vol. 8, 2000, p. 1–15. https://doi.org/10.1007/978-1-4757-3264-1_1.

Veasey A. The physical separation and recovery of metals from waste, volume one. CRC Press; 1993.

Veasey T. An overview of metals recycling by physical separation methods. Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering 1997;211:61–4.

Vichi M, Cavicchia C, Groenen PJF. Hierarchical Means Clustering. J Classif 2022;39:553–77. https://doi.org/10.1007/s00357-022-09419-7.

Villalba G, Segarra M, Fernández AI, Chimenos JM, Espiell F. A proposal for quantifying the recyclability of materials. Resour Conserv Recycl 2002;37:39–53. https://doi.org/https://doi.org/10.1016/S0921-3449(02)00056-3.

Vander Voort GF. Metallography and microstructures. ASM international; 2004.

Walczak S, Cerpa N. Artificial Neural Networks. In: Meyers RA, editor. Encyclopedia of Physical Science and Technology (Third Edition), New York: Academic Press; 2003, p. 631–45. https://doi.org/https://doi.org/10.1016/B0-12-227410-5/00837-1.

Wall M, Rechtsteiner A, Rocha L. Singular Value Decomposition and Principal Component Analysis. In A Practical Approach to Microarray Data Analysis 2002;5. https://doi.org/10.1007/0-306-47815-3_5.

Wang H, Lv F, Zhan Z, Zhao H, Li J, Yang K. Predicting the Tensile Properties of Automotive Steels at Intermediate Strain Rates via Interpretable Ensemble Machine Learning. World Electric Vehicle Journal 2025;16. https://doi.org/10.3390/wevj16030123.

Wang H, Pan L, Chen Y, Cai Z, Zhao Y, Liu G. Research on the Cold Rolling Process, Microstructure and Properties of 305 Austenitic Stainless Steel Thin Strips. Materials 2024;17:1250. https://doi.org/10.3390/ma17061250.

Wang X, Zhang W, Lin Z, Li H, Zhang Y, Quan W, et al. Current Status of Image Recognition Technology in the Field of Corrosion Protection Applications. Coatings 2024;14. https://doi.org/10.3390/coatings14081051.

Wang Y, Huang H, Rudin C, Shaposhnik Y. Understanding how dimension reduction tools work: an empirical approach to deciphering t-SNE, UMAP, TriMAP, and PaCMAP for data visualization. Journal of Machine Learning Research 2021;22:1–73.

Ward L, Agrawal A, Choudhary A, Wolverton C. A general-purpose machine learning framework for predicting properties of inorganic materials. NPJ Comput Mater 2016;2:16028. https://doi.org/10.1038/npjcompumats.2016.28.

Wiens M, Verone-Boyle A, Henscheid N, Podichetty J, Burton J. A Tutorial and Use Case Example of the eXtreme Gradient Boosting (XGBoost) Artificial Intelligence Algorithm for Drug Development Applications. Clin Transl Sci 2025;18. https://doi.org/10.1111/cts.70172.

Wiklund C-M, Helle M, Kohl T, Järvinen M, Saxén H. Feasibility study of woody-biomass use in a steel plant through process integration. J Clean Prod 2017;142:4127–41. https://doi.org/https://doi.org/10.1016/j.jclepro.2016.09.210.

World Steel Association. World steel production in 2021. Https://WorldsteelOrg/Wp-Content/Uploads/2021-World-Steel-in-FiguresPdf 2021.

Worrell E, Reuter M. Handbook of Recycling: State-of-the-art for Practitioners, Analysts, and Scientists. 2014.

Wozniak P, Akerlof C, Amrose S, Brumby S, Casperson D, Gisler G, et al. Classification of ROTSE Variable Stars using Machine Learning 2001;33:1495.

Wu Y, Zhang L, Bhatti UA, Huang M. Interpretable Machine Learning for Personalized Medical Recommendations: A LIME-Based Approach. Diagnostics 2023;13. https://doi.org/10.3390/diagnostics13162681.

Xie Q, Suvarna M, Li J, Zhu X, Cai J, Wang X. Online prediction of mechanical properties of hot rolled steel plate using machine learning. Mater Des 2021;197:109201. https://doi.org/https://doi.org/10.1016/j.matdes.2020.109201.

Xie S, Lawniczak AT, Gan C. Optimal number of clusters in explainable data analysis of agent-based simulation experiments. J Comput Sci 2022;62:101685. https://doi.org/https://doi.org/10.1016/j.jocs.2022.101685.

Xin G, Fan P. A lossless compression method for multi-component medical images based on big data mining. Sci Rep 2021;11:12372. https://doi.org/10.1038/s41598-021-91920-x.

Xiong J, Zhang T, Shi S. Machine learning of mechanical properties of steels. Sci China Technol Sci 2020;63:1247–55. https://doi.org/10.1007/s11431-020-1599-5.

Xue J, Lee C, Wakeham SG, Armstrong RA. Using principal components analysis (PCA) with cluster analysis to study the organic geochemistry of sinking particles in the ocean. Org Geochem 2011;42:356–67. https://doi.org/https://doi.org/10.1016/j.orggeochem.2011.01.012.

Xue Y, Tong Y, Neri F. An ensemble of differential evolution and Adam for training feed-forward neural networks. Inf Sci (N Y) 2022;608:453–71. https://doi.org/https://doi.org/10.1016/j.ins.2022.06.036.

Yang B, He J, Zhang G, Guo J, editors. Chapter 11 - Applications of vanadium in the steel industry. Vanadium, Elsevier; 2021, p. 267–332. https://doi.org/https://doi.org/10.1016/B978-0-12-818898-9.00011-5.

Yang J, Wang Y, Gao X, Zuo R, Song L, Jin C, et al. Vanadium: A Review of Different Extraction Methods to Evaluate Bioavailability and Speciation. Minerals 2022;12. https://doi.org/10.3390/min12050642.

Yassine A, Mohamed L, Al Achhab M. Intelligent recommender system based on unsupervised machine learning and demographic attributes. Simul Model Pract Theory 2021;107:102198. https://doi.org/https://doi.org/10.1016/j.simpat.2020.102198.

Yilmaz Eroglu D, Guleryuz E. Enhanced Three-Stage Cluster-Then-Classify Method (ETSCCM). Metals (Basel) 2025;15. https://doi.org/10.3390/met15030318.

Yu B, An R, Zhao G. Spatial and temporal disparity of the in-use steel stock for China. Resour Conserv Recycl 2020;155:104667. https://doi.org/https://doi.org/10.1016/j.resconrec.2019.104667.

Yu B, Zheng Z, Dai J. K-DGHC: A hierarchical clustering method based on K-dominance granularity. Inf Sci (N Y) 2023;632:232–51. https://doi.org/https://doi.org/10.1016/j.ins.2023.03.012.

Yu L, Zunger A. Identification of Potential Photovoltaic Absorbers Based on First-Principles Spectroscopic Screening of Materials. Phys Rev Lett 2012;108:68701. https://doi.org/10.1103/PhysRevLett.108.068701.

Yuan C, Yang H. Research on K-Value Selection Method of K-Means Clustering Algorithm. J (Basel) 2019;2:226–35. https://doi.org/10.3390/j2020016.

Zhan Z, Shi Z, Wang Z, Lu W, Chen Z, Zhang D, et al. Effect of Manganese on the Strength–Toughness Relationship of Low-Carbon Copper and Nickel-Containing Hull Steel. Materials 2024;17. https://doi.org/10.3390/ma17051012.

Zhang F, Yang Y, Shan Q, Li Z, Bi J, Zhou R. Microstructure Evolution and Mechanical Properties of 0.4C-Si-Mn-Cr Steel during High Temperature Deformation. Materials 2020;13. https://doi.org/10.3390/ma13010172.

Zhang X, Saniie J. Unsupervised Learning for 3D Ultrasonic Data Compression. 2021. https://doi.org/10.1109/IUS52206.2021.9593654.

Zhang Y, Xu X. Machine Learning Steel Ms Temperature. Simulation 2021;97:383–425. https://doi.org/10.1177/0037549721995574.

Zhang Y, Zhao Y. Automated Clustering Algorithms for Classification of Astronomical Objects. Astron Astrophys 2004;422. https://doi.org/10.1051/0004-6361:20040141.

Zhao X, Yang W. Prediction of mechanical property and costs of high-strength steel using machine learning. Canadian Metallurgical Quarterly 2025;0:1–13. https://doi.org/10.1080/00084433.2025.2471616.

Zheng H, Ye X, Li J, Jiang L, Liu Z, Wang G, et al. Effect of carbon content on microstructure and mechanical properties of hot-rolled low carbon 12Cr-Ni stainless steel. Materials Science and Engineering A-Structural Materials Properties Microstructure and Processing - MATER SCI ENG A-STRUCT MATER 2010;527:7407–12. https://doi.org/10.1016/j.msea.2010.08.023.

Zhitenev A, Salynova M, Shamshurin A, Ryaboshuk S, Kolnyshenko V. Database Clustering after Automatic Feature Analysis of Nonmetallic Inclusions in Steel. Metals (Basel) 2021;11. https://doi.org/10.3390/met11101650.

Zivic F, Malisic AK, Grujovic N, Stojanovic B, Ivanovic M. Materials informatics: A review of AI and machine learning tools, platforms, data repositories, and applications to architectured porous materials. Mater Today Commun 2025;48:113525. https://doi.org/https://doi.org/10.1016/j.mtcomm.2025.113525.

Zou J, Han Y, So S-S. Overview of Artificial Neural Networks. In: Livingstone DJ, editor. Artificial Neural Networks: Methods and Applications, Totowa, NJ: Humana Press; 2009, p. 14–22. https://doi.org/10.1007/978-1-60327-101-1_2.