

**An intelligent design
framework for multi-principal
element alloys using
optimized deep learning**

**A Thesis Submitted for the
Degree of Doctor of
Philosophy**

By

Syed Hassan Fatimi

**Department of Computer
Science, Brunel University
London**

2026

I would like to dedicate this thesis to my family.

Declaration

I, Syed Hassan Fatimi, hereby declare that this thesis and the works presented in it are entirely my own. Some of the works have been previously published in journals, and this has been mentioned in the thesis. Where I have consulted the work of others, this is always clearly stated.

Syed Hassan Fatimi

December 2025

Publications

The following paper reports the research work resulting from this thesis.

- S. H. Fatimi, Z. Wang, I. T. Chang, W. Liu and X. Liu, A novel hyperparameter optimization approach for supervised classification: phase prediction of multi-principal element alloys, *Cognitive Computation*, vol. 17, pp. 50, 2025.

Abstract

Multi-principal element alloys (MPEAs) are a new type of alloys which have shown promising applications but they lack a comprehensive theoretical framework due to insufficient experimental analysis. To circumvent these issues, machine learning has proven to be effective in assisting design of MPEAs by generating a prediction mechanism of properties of alloys.

The core object of this research is to develop a prediction framework for MPEAs via dense neural networks which uses atomic composition to predict alloy phase and its physical properties, making this a supervised learning approach. This framework will then be supplemented with hyperparameter optimization to improve its prediction power and allow us to get a better theoretical understanding of interactions between atoms in an MPEA. For the purposes of optimization, genetic algorithm (GA) is utilized with novel hyperparameters which serve to include data pre-processing within the framework of hyperparameter optimization. Ablation studies were conducted which indicate improvement in prediction by including novel hyperparameters, thus justifying our approach. This model is then tested on recently discovered MPEAs and is able to predict their alloy phases with high accuracy, making this a viable approach for experimental design.

Building on our prediction model, we further explored hyperparameter optimization by using teachnig-learning-based optimization algorithm (TLBO), which is a population based heuristic algorithm. TLBO has less control parameters than GA making it more feasible, but it suffers from premature convergence. To this affect, we introduce modifications in conventional TLBO to create balance between local and global search. Each approach is tested and compared with conventional TLBO, and the results depict notable improvement. Combinations of the three modifications are compared with each other and the best combination is termed as enhanced TLBO (eTLBO), which when compared with GA illustrates improvement in prediction accuracy.

The outcomes of this reseach provide a deeper understanding of complex mechanisms surrounding MPEAs while signifying the importance of optimization in supervised learning models. This thesis leverages the findings to discuss implications regarding high-throughput alloy design, and determine feasible machine learning systems to support it.

Acknowledgements

I would like to take this opportunity to express my sincerest gratitude to my principle supervisor Prof. Zidong Wang and my second supervisor Prof. Isaac Chang for their guidance and support during my PhD study.

It is an honor of mine to have Prof. Wang as my supervisor. During these four years of research, he has been nothing but compassionate and understanding. His efforts and ideas led me to appreciate research and mature as a person. His insight and professional guidance, as well as his constant encouragement was a source of motivation for me in my PhD and for that I will be forever grateful. I would also like to take this opportunity to express my sincere gratitude to Prof. Chang for his encouragement and assistance during the past four years. His ability to explain complex concepts and ideas of research are truly exceptional and I have always appreciated his guidance in my PhD journey.

I would like to thank Dr. Weibo Liu for helpful discussions, suggestions and support of my research and for his aid in publishing my first research paper. Also, I would like to extend my thanks to the members of Intelligent Data Analysis Centre, specifically Prof. Xiaohui Liu, for his motivation and encouragement during my yearly reviews.

I am sincerely thankful to my fellow research colleagues Jingzhong Fang and Linwei Chen for helpful discussions, suggestions and support of my research. Their high productivity and research impact kept me motivated to improve as a researcher. Giving countless presentations in front of them gave me much needed confidence and for that I am grateful.

I would like to extend my gratitude towards my flatmates and close friends Adel Safi and Abdul Rahman Safi, for making these four years an enjoyable experience. Their company provided invaluable entertainment as well as support, and I will always be grateful for their friendship.

In the end, I would like to express my deepest thanks and appreciation towards

my family. My parents provided unconditional support and encouragement, which served as motivation for me to keep moving forward. I would also like to extend this gratitude towards my grandmother and siblings. Our endless conversations filled me with much-needed comfort and solace. Last but not least I would like to extend a heartfelt thanks to my wife, who was a massive source of motivation during critical stages of my research and thesis writing. Thank you for all your adoration and care.

Contents

1	Introduction	15
1.1	Background and Context	15
1.1.1	Overview of the Topic	15
1.1.2	Historical Background	16
1.2	Research Problem	18
1.2.1	Problem Statement	18
1.2.2	Gaps in Existing Knowledge	19
1.3	Objectives of the Study	21
1.3.1	Research Objectives	21
1.4	Scope and Limitations	21
1.4.1	Scope	21
1.4.2	Limitations	22
1.4.3	Practical Implications	22
1.5	Methodology Overview	22
1.5.1	Research Design	22
1.5.2	Challenges	23
1.5.3	Contributions	23
1.6	Structure of the Thesis	24
2	Background	26
2.1	Introduction	26
2.2	Multi-Principal Element Alloys	26
2.2.1	Alloy Phases	27
2.2.2	Phase Space and MPEAs	28
2.2.3	Hume-Rothery Rules	29
2.3	Machine Learning Approach	31
2.4	Genetic Algorithm	34

2.4.1	Advantages of Genetic Algorithm	35
2.4.2	Disadvantages and Limitations of Genetic Algorithm	36
2.5	Teaching-Learning Based Optimization	36
2.5.1	Advantages of TLBO	38
2.5.2	Disadvantages and Limitations of TLBO	38
3	A Novel Hyperparameter Optimization Approach for Supervised Classification: Phase Prediction of Multi-Principal Element Alloys	41
3.1	Motivation	41
3.2	Background	45
3.2.1	Multi-Principal Element Alloys	45
3.2.2	Description of Data Set	45
3.3	Methodology	48
3.3.1	ANN based Classification	48
3.3.2	Genetic Algorithm	49
3.3.3	Hyperparameter Optimization	53
3.4	Experimental Results	55
3.4.1	Data Statistics	55
3.4.2	Data Pre-Processing	57
3.4.3	Hyperparameter Optimization	57
3.4.4	Analysis of Results	58
3.4.5	Validation of Prediction Model	59
3.4.6	Ablation Studies	60
3.5	Conclusion	61
4	Teaching Learning Based Optimization	67
4.1	Introduction	67
4.2	Background	69
4.2.1	Multi-Principal Element Alloys	69
4.2.2	Data Set Description	69
4.3	TLBO Algorithms	70
4.3.1	Basic TLBO Algorithm	70
4.3.2	Popular TLBO Variants	71
4.4	Novel TLBO Algorithm	72
4.4.1	Incorporation of Previous Knowledge in TLBO	73

4.4.2	Noisy TLBO with Weighted Teaching	74
4.4.3	TLBO with Gaussian Sampling and Increased Knowledge Shar- ing	75
4.5	Experimental Results	76
4.5.1	Benchmark Results	76
4.5.2	Hyperparameter Optimization	78
4.5.3	Analysis of Results	80
4.6	Conclusion	80
5	Periodic Table Representation	91
5.1	Introduction	91
5.1.1	Limitations of Hume-Rothery Rules in MPEAs	91
5.1.2	Qualitative Description of Periodic Table Representation . . .	91
5.2	Methodology	92
5.2.1	Framework of Periodic Table Representation	92
5.2.2	Convolutional Neural Network	94
5.3	Experiments and Results	96
5.3.1	Data Collection and Preprocessing	96
5.3.2	Architecture of CNN	96
5.3.3	PTR Results	97
5.3.4	Validation of PTR Prediction Model	99
5.4	Conclusion	100
6	Conclusion	101
6.1	Introduction	101
6.2	Summary of the Research	101
6.3	Future Research Directions	103

List of Figures

2.1	Phase diagram of tertiary alloy system, depicting high entropy region [41].	29
2.2	Flowchart of TLBO, depicting teaching and learner phase.	40
3.1	Distribution of classes in alloy dataset depicting imbalance	48
3.2	Depiction of dropout in neural networks by comparing networks with and without applying dropout [118].	49
3.3	Flowchart of Genetic Algorithm. p_{mut} is mutation probability and r is a uniformly distributed random number between 0 and 1.	63
3.4	Depiction of chromosomes where a) shows normal chromosome with fixed length genes; b) shows chromosome having a single variable length gene which consists of values $b_{71}, b_{72} \dots b_{7n}$, where n is an integer which can vary between a lower and upper bound.	64
3.5	Depiction of boxplot with interquartile distance.	64
3.6	Variation of number of outliers with k_{iqr}	64
3.7	Distribution of classes within training and testing sets when using stratified sampling.	65
3.8	Training process of ANN, depicting evolution of loss function with number of epochs.	65
3.9	Fitness plot of GA with respect to number of generations. Here, Population size = 20, Number of Generations = 5 and Elitism = 2 . .	65
3.10	Fitness plot of GA with respect to number of generations, with a zoomed in view enclosed within the plot depicting change in fitness close to saturation. Here, Population size = 40, Number of Generations = 20 and Elitism = 3.	66
3.11	Confusion matrix of testing data	66

4.1	Distribution of marks obtained for a group of learners, depicting the formulation of TLBO [73]	70
4.2	Flowchart of TLBO Variant 1, depicting teaching and learner phase. .	82
4.3	Flowchart of TLBO Variant 2, depicting teaching and learner phase. .	83
4.4	Flowchart of TLBO Variant 3, depicting teaching and learner phase. .	84
4.5	Comparison between optimization results of GA and TLBO for $F_1(x)$	85
4.6	Comparison between optimization results of GA and TLBO for $F_2(x)$	85
4.7	Comparison between optimization results of GA and TLBO for $F_3(x)$	86
4.8	Comparison between optimization results of GA and TLBO for $F_4(x)$	86
4.9	Comparison between optimization results of GA and TLBO for $F_5(x)$	87
4.10	Comparison between optimization results of GA and TLBO for $F_6(x)$	87
4.11	Comparison between optimization results of GA and TLBO for $F_7(x)$	88
4.12	Comparison between optimization results of GA and TLBO for $F_8(x)$	88
4.13	Comparison between optimization results of GA and TLBO for $F_9(x)$	89
4.14	Fitness plot of enhanced TLBO	89
4.15	Confusion matrix of testing data for enhanced TLBO	90
5.1	Illustration of periodic table representation with $Ce_{68}Al_{10}Cu_{20}Fe_2$ as an example [154]	93
5.2	Depiction of a typical CNN architecture containing convolution, pool- ing and fully-connected layers.	95
5.3	Fitness plot of GA with respect to number of generations for PTR. Here, Population size = 20, Number of Generations = 10, Elitism = 2	98
5.4	Comparison plot of fitness between GA, original TLBO and enhanced TLBO developed in Chapter 3	98

List of Tables

3.1	List of hyperparameters and their respective ranges.	55
3.2	Descriptive Statistics values of features in our dataset.	56
3.3	List of hyperparameters with the best testing accuracy	59
3.4	Prediction results of new alloys for validation of model	59
3.5	Comparison among various classification algorithms	60
3.6	Comparison results of ablation studies	61
4.1	List of benchmark functions considered in this study	77
4.2	Mathematical description of benchmark functions	77
4.3	List of comparison parameters for benchmark functions	78
4.4	Statistical results of testing accuracy comparison between GA and TLBO when applied to MPEA data.	78
4.5	Statistical results of testing accuracy comparison between original TLBO and its modified variants when applied to hyperparameter op- timization of phase prediction for MPEA data.	79
4.6	Statistical results of testing accuracy comparison between original TLBO and its modified versions when applied to hyperparameter op- timization of phase prediction for MPEA data.	79
5.1	List of hyperparameters and their respective ranges.	97
5.2	List of hyperparameters with the best testing accuracy using PTR with enhanced TLBO	99
5.3	Comparison of prediction accuracies between PTR-based optimiza- tion and use of Hume-Rothery features.	99

List of Abbreviations

MPEA	multi-principal element alloy
GA	genetic algorithm
TLBO	teaching learning based optimization
CCA	complex concentrated alloy
HEA	high entropy alloy
ML	machine learning
SS	solid solution
IM	intermetallic
AM	amorphous
FCC	face-centered cubic
BCC	body-centered cubic
HCP	hexagonal closed pack
MP	multi-phase
DFT	density functional theory
PTR	periodic table representation

Chapter 1

Introduction

1.1 Background and Context

1.1.1 Overview of the Topic

The field of alloy design is defined as the study of combining elements in atomic concentrations, such that it results in an alloy with particular chemical and physical properties. In alloy design and metallurgy, the distribution of composition within elements exists in the form of a principal element with large concentrations and minor elements added as impurities and traces. A new form of alloys based upon the studies of Cantor [1] and Yeh

citeyeh2004nanostructured have emerged recently, termed as multi-principal element alloys (MPEA). These contain elements in equiatomic concentrations, making all individual metals as principal components. MPEAs have shown promising properties surpassing conventional alloys [2] as well as a new ability to resist change, which can be exploited for applications in extreme environments [3].

Since the introduction of MPEAs in the early 2000s, they have attracted research interest due to their unique thermodynamic foundations, unconventional microstructure behaviour and promising combination of mechanical and functional properties [4]. Hence, their study is significant from both a fundamental scientific perspective and for their potential usage in harsh and demanding environments. This leads to the question of how we can design MPEAs to be utilized in useful applications. The field of MPEA design can be interpreted as the study of how combining metallic elements in different concentrations leads to particular atomic structures in an MPEA, and how these atomic structures can be linked to physical and chemical properties

of the respective MPEA.

Recent improvements in data-driven techniques have led to an increased usage in various scientific domains, specifically alloy design [5]. Machine learning techniques have proven to be useful in high-throughput MPEA design by creating viable prediction models of critical alloy properties, including atomic phases [6], [7], [8] and mechanical properties [9], [10], [11]. These models create suggestive guidelines for alloys with targeted properties, hence proving to be invaluable for experimental alloy design as well as providing a deeper theoretical understanding of complex alloy systems like MPEAs.

The problem of machine learning-based prediction framework of MPEAs can be termed as supervised learning approach, where a model is trained on labelled data, with the aim to predict the label given a set of inputs, also termed as features. Supervised learning can be further subdivided into two categories: classification and regression. Classification deals with categorical labels and regression is used for continuous labels. To train robust prediction models that accurately span the search space of a labelled data set, choosing suitable features is vital for success [12]. The choice of feature can be guided by expert knowledge of problem domain, or via feature extraction using intelligent statistical methods. Both these feature selection approaches are explored in this study, with the aim to compare and contrast model performance.

In machine learning algorithms, there exist many tunable parameters which must be adjusted according to the problem since the performance of the algorithm hinges on these hyperparameters [13]. This necessitates the use of appropriate optimization algorithms which can iteratively change hyperparameters with the aim to improve performance of the model. Among myriad optimization algorithms, population-based heuristic techniques are widely used since they are gradient-free and can be applied to a variety of continuous, discrete or multi-objective problems. Two such techniques are Genetic Algorithm (GA) [14] and Teaching Learning Based Optimization (TLBO) [15] are explored in this study, with the objective to create accurate prediction models for MPEA discovery and design.

1.1.2 Historical Background

The study of alloys and their properties has existed for centuries and has been utilized in the industry since the industrial age with the advent of stronger and

better performing machinery [16]. These studies have been restricted because the design of conventional alloys is similar to a dilute solution, where one or two principal elements serve as a solvent and other metals are added in small concentrations which act as solutes. If we observe the phase space of these alloys, we can see that the conventional alloys are only exploring the corners and edges of phase space [17]. The central region of phase space signifies that every element is a principal element with none acting as either solvent or solute. Phase space and phase diagrams are explained in Chapter 2, Section 2.2.

From the above description of conventional and special alloys, historically over five millennia the alloy design, alloy production, and alloy selection were all based on one principal-element or one-compound concept. This alloy concept has generated numerous practical alloys contributing to civilization and daily life. But it still limits the degree of freedom in the composition of the alloy and thus restricts the development of special microstructures, properties, and applications. Consequently, materials science and engineering of alloys is not fully explored since those alloys outside this conventional scheme have not been included.

Early foundations for MPEA concepts can be traced to classical alloy theory and phase stability studies in the mid-20th century. Work on substitutional solid solutions, Hume–Rothery electron concentration rules and transition metal alloy systems demonstrated that multicomponent alloys could form stable solid solutions under appropriate electronic and size-factor conditions [18], [19]. However, most alloy development during this period remained focused on binary and ternary systems due to experimental and thermodynamic modeling limitations. Further research led to an increased interest in complex alloy systems in connection with superalloys, metallic glasses, and intermetallic compounds, where multiple alloying elements were used to tailor performance. However, these systems still typically maintained a dominant principal element in large concentrations.

The modern history of MPEAs is marked by two independent and simultaneous landmark publications. Yeh et. al. [20] introduced the term high-entropy alloys to describe alloys composed of multiple principal elements in near-equiatomic ratios, emphasizing the potential stabilizing role of high configurational entropy in promoting simple solid solution phases. Their work demonstrated that equiatomic multicomponent systems such as AlCoCrCuFeNi could form simple crystal structures despite high compositional complexity. In the same year, Cantor et al [1] reported the formation of single-phase solid solutions in equiatomic multicomponent transition-

metal alloys, most notably the now well-known CoCrFeMnNi system, often referred to as the Cantor alloy. These studies provided the first systematic experimental validation that multi-element alloys could exhibit stable, simple crystal structures and useful mechanical properties.

As the field of MPEA matured, it became evident that not all compositionally complex alloys possess high configurational entropy or equiatomic compositions. This recognition motivated the use of broader terminology such as compositionally complex alloys (CCAs), to include non-equiatomic and entropy-moderate systems derived from similar design. These expanded definitions reflect a shift from entropy-centered classification toward a more general composition-space perspective [4], [21].

The historical development of MPEAs has also been closely tied to advances in computational thermodynamics and high-throughput methods. The growth of CALPHAD databases for multicomponent systems, first-principles electronic structure calculations, and materials informatics has enabled vast exploration of high-dimensional alloy spaces. These tools have accelerated alloy discovery and helped interpret early MPEA concepts within a viable thermodynamic framework [4], [22].

Now, MPEAs consist of a broad class of materials such as structural, refractory, lightweight and functional alloy systems. Their historical trajectory signifies the importance of optimum alloy design strategies. These strategies continue to influence modern alloy development, therefore aiding the integration between experimental metallurgy and data-driven alloy design.

1.2 Research Problem

1.2.1 Problem Statement

The research questions discussed in this section aim to advance intelligent design of MPEAs through supervised prediction models utilizing hyperparameter optimization and suitable optimization algorithms. The first step is collecting data and pre-processing it. Second step is creating algorithms to achieve the goals of this study. The last step is implementing the algorithms on alloy data to aid high-throughput design. The following research questions were considered for this thesis.

- **Problem Statement 1:** How can we work with small and low quality datasets of experimental MPEA synthesis which contain outliers and imbalanced distribution of labels, and clean it effectively such that it retains crucial informa-

tion without adversely affecting the training process of prediction model? To address this question, a novel outlier detection approach is utilized which is incorporated within hyperparameter optimization. This way, hyperparameter optimization takes care of outlier detection and removal while maintaining an effective prediction quality.

- **Problem Statement 2:** How can we create an accurate alloy phase prediction model for MPEAs? To address this question, a supervised deep neural network based classification model is created and implemented on MPEA data, with testing accuracy as the performance metric. The model is made robust with the use of hyperparameter optimization, where population-based heuristic algorithms GA and TLBO are utilized.
- **Problem Statement 3:** How can we improve TLBO such that it can be employed for high-dimensional MPEA data? To address this query, certain modifications are made in the conventional TLBO setup which serve to slow down convergence, hence claiming its superiority over GA due to the presence of less ad-hoc control parameters, making it more feasible for application.

These research questions are designed to systemically identify factors influencing alloy design. By examining both existing and novel elements, this thesis aims to address limitations of current approaches and contribute towards more durable and generalizable models.

1.2.2 Gaps in Existing Knowledge

MPEAs have matured rapidly as a field, however several substantive gaps in knowledge still remain. The most consequential gap when it comes to the application of machine learning techniques to MPEAs is the lack of comprehensive databases. Sporadic databases exist in literature but they are confined to specific groups of MPEAs or certain key chemical or physical properties [4]. A critical aim of this study is to collect and compile MPEA data from different sources in literature, such that not only the current study but future studies can benefit from it. This is imperative for data science as indicated in the name, requires clear and concise data to train models and make statistical predictions and qualitative analysis.

Another key gap existing in the framework of MPEAs is the use of Hume-Rothery rules-based identifying features. Hume-Rothery rules define how metallic elements

combine to form alloys of different structures [18]. These rules have been used to create features of MPEAs which have been utilized in machine learning-based prediction models [23]. These alloy features have been used in literature to help identify phase formation, specifically for solid solution phases [24]. They provide a guideline for whether a combination of metallic atom types and concentrations can form a solid solution or not. Although they hold historic significance, they have limited applicability to modern alloy systems which are multicomponent and complex. This is due to the fact that Hume-Rothery rules were developed for binary alloy systems, and any extension towards multicomponents suffers from over-simplification [25]. These rules are qualitative and empirical, as opposed to being derived from first-principle thermodynamics [26]. Their specificity for classifying only solid solutions is another drawback, since they provide no guidance for predicting ordered compounds and complex intermetallics, as well as amorphous phases, which lack long-range order [27]. Hume-Rothery based features are used extensively in machine learning studies applied to MPEAs because of lack of alternatives. These rules are finicky but still provide a decent ground work for complex alloy systems. Our study aims to supplement Hume-Rothery based features with additional statistical features which will aid in the process of predictive model building. Another approach that this study explores is to remove these features all together and use first-principle chemical knowledge from position of metals on the periodic table.

Teaching learning based optimization (TLBO) is a recent addition to the long list of population based heuristic optimization algorithms. It has a similar structure compared to genetic algorithm (GA) and particle swarm optimization (PSO), but it has less tunable parameters, making it more feasible for applications. This comes with some drawbacks, mainly that TLBO relies heavily on mean-based population updates in its two phases which can rapidly reduce population diversity, leading to premature convergence [28]. The current study aims to fill in the gaps of TLBO performance by introducing additional mechanisms to slow down convergence and improve its capability to find optimum solutions.

The use of Hume-Rothery rules as a guideline to define features used for MPEA phase prediction is widely adopted in research [25], [26], [27], however it is based only on interactions among binary alloys [4]. Its extension towards multi-component systems is inaccurate and unreliable, although its wide usage is due to the fact that it is historically significant and there is no reliable theoretical framework to describe the complex interactions between atoms of each principal element in a large unit cell

of an MPEA. A critical aim of this study is to supplement or replace Hume-Rothery rules with a more feasible and accurate representation of chemical behaviour in MPEAs at the atomic scale which guide phase formation.

1.3 Objectives of the Study

1.3.1 Research Objectives

This research aims to create a high-throuput design framework for MPEA with predictive machine learning methods, test it on existing experimental alloy data and advance the theoretical knowledge of MPEA engineering. Following are the core objectives of this research.

- **Objective 1:** Conduct a comprehensive literature review on MPEA phase prediction using machine learning, identifying current methodologies and their effectiveness, as well as evaluating challenges and gaps related to model performance and generalizability.
- **Objective 2:** Systemically collect experimental data of MPEA synthesis and identify key characteristics and issues with the dataset.
- **Objective 3:** Create a novel supervised prediction model for MPEA phase prediction with hyperparameter optimization.
- **Objective 4:** Improve upon existing optimization techniques by exploring new approaches to balance local and global search.

Objectives 1 and 2 are related to problem statement 1. Objective 3 is designed to solve problem statement 2. For the solution of problem statement 3, objective 4 is created. Hence, each objective in this study serves to answer a particular problem statement defined in Section 1.2.1.

1.4 Scope and Limitations

1.4.1 Scope

Phase prediction of MPEAs is the main scope of this research. This includes data collection, algorithm selection, algorithm modifications and implementation. The

framework developed in this research can be extrapolated towards other alloy types such as metallic glasses and complex concentrated alloys. The prediction model is flexible and geared towards addition of new synthesized MPEAs which will continuously improve the ability of generalization of the model for all MPEAs and potentially all alloys.

It should be noted here that the MPEA dataset used in this study is entirely real, and no synthetic data is present. Synthetic data has been used in literature to supplement the dataset [29], [30], but they are not equivalent to experimentally verified data in terms of reliability.

1.4.2 Limitations

This research is mainly aimed at MPEAs which have been synthesized and reported in literature. Any extrapolation towards simulated data created by density functional theory (DFT) , molecular dynamics (MD) or any ab-initio quantum mechanical methodology is not considered. We understand that these simulated alloys can supplement the original experimental dataset but they should be treated with caution since they each have their own limitations in predicting alloy phases.

1.4.3 Practical Implications

The concept of circular metals can be termed as the study of recyclable and reusable alloys in an approach to manage precious resources and minimize waste and negative environmental impact. Our research will aid in advancing the study of MPEAs as circular metals, led by the Interdisciplinary Centre for Circular Metals, a research centre which aims to make UK the first nation to fully reuse and recycle metals. This would mean mining and extraction of metals will be significantly reduced making a cleaner and more sustainable environment.

1.5 Methodology Overview

1.5.1 Research Design

To answer the research questions indicated above, this thesis uses a quantitative approach of intelligent data analysis to create a prediction framework for MPEAs. By adopting this design approach, we can effectively identify potential alloy phases

and link them to mechanical and chemical properties, hence aiding in the search for applicable MPEAs with desirable characteristics.

Following this orientation, deep neural networks are employed with hyperparameter optimization to undertake a robust exploration of data. For the purposes of hyperparameter optimization, two gradient-free optimization techniques, GA and TLBO, are utilized with novel hyperparameters to support preprocessing and feature selection.

1.5.2 Challenges

The application of intelligent data-driven techniques is reliant upon good quality data. This is a major challenge in this study since the data of MPEAs used for training is obtained from experimental formulation and analysis of alloys, and contains many core issues listed below with the relevant problem statement in parenthesis..

1. Small database due to lack of experimental verification (Problem statement 1).
2. Skewed distribution of phases, since some phases are more desirable and hence published more often (Problem statement 1).
3. Presence of outliers and noise which are an inherent by-product of experimental design which can contain systematic as well as random errors (Problem statement 1).
4. Use of non-standard algorithms and lack of standardized testing due to the fact that the focus of prior research is on alloy data rather than improvements and standardization of algorithms and lacks proper justification of using a particular technique (Problem statement 2).

1.5.3 Contributions

Contributions of the work presented in this thesis have been separated into major and minor categories described below.

Major Contributions

1. A novel method of pre-processing data and optimizing hyperparameters simultaneously by introducing outlier detection and feature subset selection within

the framework of hyperparameter optimization.

2. Introducing a novel encoding scheme of GA which can simultaneously handle continuous and discrete data as well as dynamically changing chromosome length.
3. Introducing novel modifications within original TLBO with the purpose of balancing exploration and exploitation and avoiding premature convergence.
4. Using an enhanced feature engineering scheme by using periodic table representation which outperforms Huume-Rothery based features for both GA and enhanced TLBO.

Minor Contributions

1. A comprehensive database is designed with the explicit purpose of being utilized in machine learning techniques, containing atomic concentrations of each alloy with their respective features and phase.
2. Use of GA for hyperparameter optimization to create a prediction mechanism for MPEA phases by using empirical features as inputs and alloy phase as output.
3. Creating a robust and accurate phase prediction model to be readily utilized in experimental design as an aid for alloy discovery with tailored properties.

1.6 Structure of the Thesis

The thesis is structured as follows:

- **Chapter 2: Background** - In this chapter we provide necessary background information of the knowledge related to this thesis. Firstly, we introduce the concept of MPEAs and their various definitions supported by literature. Then, we briefly introduce machine learning and present a thorough review of its applications regarding MPEAs.
- **Chapter 3: A Novel Hyperparameter Optimization Approach for Supervised Classification : Phase Prediction of Multi-Principal Element Alloys** - Here, we have proposed a novel hyperparameter optimization

method for ANNs with the addition of two novel hyperparameters: outlier detection and feature subset selection. These hyperparameters serve to simultaneously pre-process and train it to provide accurate alloy phase predictions. GA is utilized as the optimization algorithm with a novel dynamic length chromosome operator which can handle genetic operations for variable length individuals in the population. Results obtained from ablations studies have indicated significant improvement of accuracy with the inclusion of new hyperparameters, hence justifying our approach. We tested this algorithm on recently published MPEAs, and the predicted alloy phases were in agreement with actual phases, making this a viable approach for aiding experimental analysis of MPEAs.

- **Chapter 4: Teaching Learning Based Optimization** - Here, a novel modifications in TLBO are proposed which serve to slow down convergence and create a better balance between exploration and exploitation. These TLBO variants are tested on MPEA phase data and show promising results. A novel permutation scheme is utilized to find the optimum configuration of modifications, and the resulting variant is termed as enhanced TLBO. The approach is validated by comparing it with GA, where enhanced TLBO exhibits better prediction accuracy.
- **Chapter 5: Periodic Table Representation** - In this chapter, periodic table representation is utilized in conjunction with CNNs, which are a better representation of chemical properties of multicomponent alloys as compared to Hume-Rothery rules, which are accurate only for binary and tertiary systems. Hyperparameter optimization is applied to CNN architecture with novel hyperparameters and modified TLBO and GA serving as optimization algorithms. Prediction accuracy results show significant improvement compared with using Hume-Rothery based features.
- **Chapter 6: Conclusion** - In the last chapter, a comprehensive summary of the work done in this thesis is presented, alongside some research directions for the future.

Chapter 2

Background

2.1 Introduction

This chapter explains concepts relevant to the study of MPEAs and the use of machine learning techniques to quantify MPEAs for experimental design and theoretical understanding. The literature presented in this chapter is identified through a systematic literature review using Google Scholar with the following keywords: "machine learning", "high entropy alloys", "multi-principal element alloys", "phase prediction" for Section 2.2 and Section 2.3, "genetic algorithm" for Section 2.4, "teaching learning based optimization" for Section 2.5. It should also be noted that no assistance of artificial intelligence or large language model is used anywhere in the thesis.

2.2 Multi-Principal Element Alloys

The study of MPEAs started with the publishing of two papers by Cantor[1] and Yeh[20] in 2004. Their findings pushed forward the path to new discoveries in functional alloys and metallurgy. These alloys were found to be resilient against changes in composition of a single element which was not possible for the case of conventional alloys where small changes in concentrations leads to noticeable changes in physical properties. This allows MPEAs to be better suited for recycling as they will retain their physical characteristics even in the presence of variations in their composition.

The first results on crystalline multi-principal element alloys (MPEAs) [1] and high entropy alloys (HEAs) [31], [32], [33], [34] were published in the same year,

but background work began much earlier [35]. The terms, ‘high entropy alloys’ (HEA) and ‘multi-principal element alloys’ (MPEAs) were both introduced in these publications. Henceforth, these two terms were used interchangeably in literature.

MPEAs are often categorized as having high configurational entropy. Entropy of a system is the degree of disorder, more specifically it is the number of possible microstates that can occur in a statistical ensemble. In regards to atomic configuration, entropy is the number of unique possible ways atoms can arrange themselves in a lattice. For MPEAs the number of possibilities is much larger than conventional alloys due to the presence of multiple principal elements with near equiatomic concentrations. This leads to high configurational entropy, hence MPEAs are also widely known as high entropy alloys. The standard definition of MPEAs is an alloy having at least five principal elements, each having an atomic percentage of 5-35%. The unique feature of MPEAs is even with high configurational entropy, they exhibit stable phases for some concentrations. The search for such stable phases is ongoing and many different experimental and computational analysis has been utilized in literature.

2.2.1 Alloy Phases

Alloy phases are defined as distinct regions within an alloy with uniform structure. These structures determine mechanical, thermal and chemical properties of alloys [36]. They can exist in various ordered, semi-ordered or unordered forms. A specific subgroup of alloy phases is called solid solution, where atoms of multiple principal elements are completely soluble in the solid state. Examples of solid solution phases are body-centered cubic (BCC), face-centered cubic (FCC) and hexagonal close-packed (HCP). Sometimes, instead of forming a solution, principal elements chemically combine with each other to form compounds. These are termed as intermetallic phases and are characterized by their brittleness and hardness, making them useful in applications where strength and durability are critical [37]. It is not necessary for alloys to exist in ordered structure, as evident by amorphous phase which does not have long-range order. They are used in applications requiring high elasticity and wear resistance [38]. Another key aspect of study of phases is the fact that alloys can have multiple phases coexisting simultaneously in different regions of the alloy sample. Such alloys are termed as multiphase alloys. This enables a balance of properties that single-phase alloys cannot achieve, making them essential

in advanced engineering and industrial applications [39].

2.2.2 Phase Space and MPEAs

Phase space is the set of thermodynamic conditions under which an alloy can exist. For an alloy, the relevant variables include chemical composition, temperature, pressure and processing conditions. Mathematically, the phase space can be represented as

$$(T, P, x_1, x_2, \dots, x_n) \quad (2.1)$$

where T is temperature, P denotes pressure and x_i represents the composition of each element i in an alloy with n total elements.

A phase diagram is a map showing which phases are thermodynamically stable under specific conditions. Most engineering phase diagrams plot the temperature on vertical axis and composition on horizontal axis, and each region represents a different equilibrium phase or combination of phases. However, this is only the case for binary systems. For ternary alloy systems, the phase diagram is represented using a triangle with its vertices depicting each of the three components, hence every point inside the triangle corresponds to one unique alloy composition. Alloys along the edges illustrate binary systems, whereas compositions inside the triangle depict all the possible ternary alloys [40].

As opposed to conventional alloys which only explore corners and edges of phase diagram, MPEAs contain configurations in the center of the phase diagram, leading to a more complete theoretical picture. Equatomic MPEAs lie at the center, and near-equatomic MPEAs explore areas close to the center. This region is termed as "high entropy" region, because MPEAs have a significantly higher configurational entropy than conventional alloys existing in the corners [4]. This phenomenon is depicted in Fig. 2.1, where phase diagram of a tertiary alloy system is illustrated, with high entropy region situated at the center. Here, configurational entropy is represented as a contour color map with central red region depicting MPEAs with high entropy.

MPEAs are defined in terms of either entropy or composition. One of the earliest papers defines HEAs as, "those composed of five or more principal elements in equimolar ratios" [20]. This requirement is restrictive and the following sentence in the same paper expands this definition to include, "principal elements with

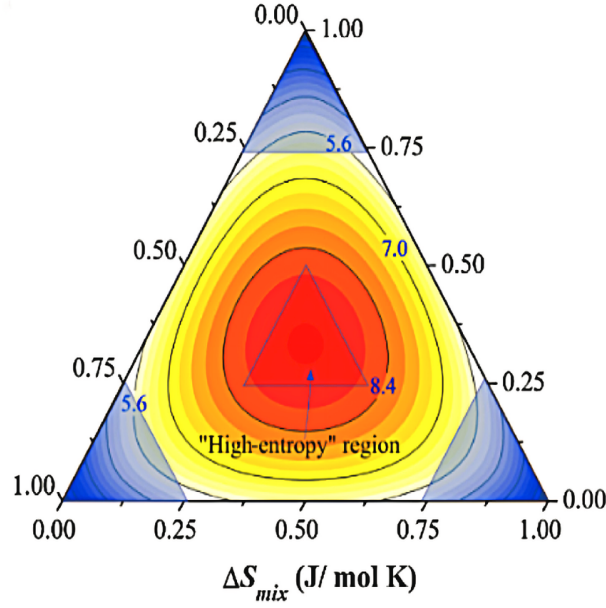


Figure 2.1: Phase diagram of tertiary alloy system, depicting high entropy region [41].

the concentration of each element being between 35 and 5 at%.” Thus, HEAs do not necessarily have to be equimolar, increasing the number of HEAs significantly. HEAs may also contain minor elements to change the properties of the base alloy, further expanding the number of HEAs [42]. This composition-based definition is based only on elemental compositions and places no restrictions on the magnitude of configurational entropy.

Whether initially intended or not, HEAs have become linked with finding single-phase solid solutions by manipulating configurational entropy. There are several terms that are used to describe the sheer scale of composition space. These terms include multi-principal element alloys (MPEAs), complex concentrated alloys (CCAs) and baseless alloys.

2.2.3 Hume-Rothery Rules

The Hume-Rothery rules were originally designed to explain the formation of solid solutions in binary alloy systems [27]. These empirical rules describe the conditions under which a metallic element can be dissolved into another to form stable solid solution phases. These principles have been extended to tertiary and quinary systems, as well as multi-component systems like MPEAs [43]. The underlying physical quantities associated with these rules can be used to engineer features that can be used as inputs to a machine learning algorithm for predicting alloy phases. They

incorporate domain knowledge within the learning process since they are derived from fundamental atomic characteristics, hence enabling efficient exploration of the large compositional search space. The methodology presented in Chapter 3 and Chapter 4 will use Hume-Rothery rules to generate input features. These features are termed as Hume-Rothery features.

Limitations of Hume-Rothery Rules for MPEAs

Despite their vast use, Hume-Rothery rules contain several key limitations that restrict their application to complex alloy systems like MPEAs. One major drawback of these rules lies in their empirical and qualitative nature. The criteria used most notably is the 15% atomic size difference threshold. This is not derived from first-principles but instead represent generalizations based on a limited set of experimental values. As a result, they lack predictive meticulousness and often fail to capture the nuance between enthalpy and entropy that mainly governs phase stability in alloy systems [18][26].

A second limitation involves the oversimplified treatment of electronic effects. Valence electron concentration is an important indicator of phase formation in Hume-Rothery rules but it does not adequately account for directional bonding, orbital hybridization, or band structure effects that are critical in transition metal alloys and complex intermetallic compounds. Also, alloys that satisfy the traditional Hume-Rothery criteria may exhibit limited solubility or unexpected phase formation due to electronic structures not considered by the rules [44].

Another assumption that Hume-Rothery rules make is the idealized substitutional behaviour. Hence, they are poorly suited to describing systems with significant interstitial solubility, strong chemical ordering, or compound formation [45]. In many functional materials such as MPEAs and CCAs, the atomic interactions are non-ideal, and lattice distortions, local chemical environments, and short-range order play a crucial role in determining phase stability [4].

A fourth critical drawback is the limited compositional scope of these rules. They were originally created for binary alloy systems and became unreliable when extended to ternary or multicomponent systems. Modern material design frequently includes high-dimensional compositional spaces where complex interactions cannot be reduced to pairwise comparisons described in Hume-Rothery rules [46]. In such systems these rules act only as a coarse screening tool at best, rather than a reliable

predictive model.

Furthermore, these rules do not incorporate thermodynamic and kinetic factors such as temperature-dependent entropy contributions and diffusion rates. When creating these alloys in a realistic lab setting, their phase formation is often governed by non-equilibrium paths which may lead to microstructures that are substantially different from those predicted by equilibrium-based heuristics such as the Hume-Rothery rules [26].

Taking all these limitations under consideration, In Chapter 5 we have proposed a novel representation scheme for MPEAs which does not involve the use of Hume-Rothery rules and instead engineers features directly from the periodic table.

2.3 Machine Learning Approach

Although machine learning is being extensively applied to problems in materials research, there are several challenges in efficiently exploring such a vast domain of knowledge. Specifically, there are numerous machine learning terminologies that a non-specialist is unfamiliar with. Understanding these terminologies will lead to improved applicability of data science techniques as well as quantitative analysis of results.

According to a study by Meredig, there are five key research areas in machine learning for materials science [47]: (1) Training models to predict materials properties and the validation by experiments; (2) ML methods development for descriptors, multi-property optimization [48], extrapolation detection [49], and uncertainty quantification [50]; (3) High-throughput data collection; (4) ML to accelerate material characterization which advances the current methodologies and techniques; and (5) Integration of physics and physics-based simulations within machine learning [51].

Compared to conventional alloys, MPEAs have a very large compositional space due large number of possible alloys that can be formed by any combination of metallic elements. This makes the search for an optimum alloy increasingly difficult as this space has not been explored in much detail and there are only certain configurations with stable phases which can be used in functional applications. Hence, the trial-and-error methodology of experimental design which was used for conventional alloys is not feasible for MPEAs. This makes the use of data science techniques extremely beneficial in creating and exploring the search space. This has been made possible by using machine learning techniques which operate on the principal that

machines can be instructed to learn by example. This is based on the same learning patterns as humans but the advantage of using computers is they can store and analyze large datasets without any human bias. These techniques revolutionized the IT industry as well as all branches of science. This is evident from the fact that nobel prize for physics and chemistry in 2024 was awarded to researchers based on their work on data science techniques. Machine learning techniques are especially useful when working with high-dimensional, non-linear or noisy data.

Machine learning can be classified into three categories: regression, classification, and clustering. For MPEA studies, regression models are applied for continuous target outputs, such as yield strength whereas classification models are used for discrete outputs, like alloy structures phases. Based on the training input, machine learning can also be classified into three types: supervised learning, semi-supervised learning, and unsupervised learning.

To generate a search space viable for MPEAs, we require a framework of physical and chemical properties. The best way to accomplish this feat is to use phase classification, where physical properties are treated as features and alloy phase is treated as class. Hence we will describe phase classification below.

Generally, there are four categories in alloy phases: solid solution (SS), inter-metallic (IM), amorphous (AM) and mixed phases. Most of the research on MPEAs is primarily aimed at discovering single-phase SS that has promising mechanical properties, such as increasing hardness and strength [52], [53].

Troparevsky et al. proposed a model as the criteria to predict the formation of single-phase MPEAs, which used enthalpies of formation of binary compounds based on density functional theory (DFT) [54]. Lederer et al.'s Ab-initio approach, named as Lederer-ToherVecchio-Curtarolo (LTVC), has the capability to scan through many thousands of systems available in the AFLOW consortium repository [55]. Lederer et al's LTVC is corroborated by monte carlo (MC) simulations for estimating the transition temperature of a SS. Ab-initio energies were incorporated within a mean-field statistical mechanical model where an order parameter followed the evolution of disorder.

For the predictions of specific phases, such as the BCC and FCC, Pei et al. utilized a support vector machine model to predict the stable HEAs from the composition space of 16 metallic elements, which was one order larger than the number of available experimental data [56]. There are 322 as-cast samples considered in this study with a cross validation accuracy over 90 % after training and testing. Pei

et al. anticipated there have 369 FCC and 267 BCC equiatomic MPEAs. Eleven RHEAs agreed with recent experiments, and the 20 quinary ones with the highest melting temperatures are validated through first-principles calculations [56]. Pei et al.'s model was complementary to the calculation of phase diagrams and Ab-initio methods. Li et al. built a SVM model with 322 as-cast MPEAs and a cross validation accuracy over 90 % to forecast the FCC and BCC phases formation [56]. Among them, several RHEAs were screened out based on a high ratio of melting temperature to density. 11 of them complied with recent experiments and the 20 quinary MPEAs were validated through first-principles calculations. Kostichenko developed a novel machine learning based-potential involving relaxation effects, called as low-rank potentials [57]. In combination with MC simulations, the potentials are proved to reach high accuracy as cluster expansion techniques in studying the phase stability in the prototypical NbMoTaW HEA. Meanwhile, Song et al. proposed a method that combines electronegativity difference, CALPHAD and ML to screen the high yield strength region in Co-Cr-Fe-Ni-Mo multi-component. SVMs model with trained data of CALPHAD calculations was established to predict the phase decomposition temperature for obtaining single-phase HEAs. Because the yield strength is positively related to the electronegativity difference, the compositions of HEAs with the phase decomposition temperature of 900 K and the maximum electronegativity are screened from the large amounts of high-throughput machine learning calculations. Moreover, Song et al.'s results are validated by the experiments [58].

Since the IM was found to enhance functional applications via precipitation hardening, comprehensive studies on the IM phases have been conducted [59]. Islam et al.'s machine learning for phase selection in the HEAs utilized an artificial neural network to identify patterns in an experimental dataset [60]. Islam et al. considered five features in their study and analyzed the correlations between them that led to the phase selection in a dataset of 118 alloys of the HEAs and trained a ANN model to classify the resulting phases based on the input features. They found a very high accuracy (> 99 %) in learning of the full dataset as well the most important factor is the VEC and the least important one is the mixing entropy.

For machine learning guided appraisal and exploration of MPEA phase design, Zhou et al. used one-dimensional convolutional neural network (CNN) and ANNs to create a phase prediction framework [61]. Later, further studies were conducted regarding complex phase predictions. For example, Dixit et al. [62] used ANNs with the novel inclusion of processing routes, e.g., arc melting, injection-

casting, and sputtering as an input feature apart from the conventional thermodynamic features. Given the total input features, the proposed model could predict all the coexisting phases, i.e., FCC, BCC, FCC1 + FCC2, BCC1 + BCC2, B2, Laves, C14 and Laves + Sigma, with 87.08 % accuracy. Lee et al.'s later work [63] refined the ANN into a deep learning-based phase predictor. To optimize hyperparameters in the architecture, Bayesian optimization was used, and a conditional generative adversarial network (CGAN) was employed which generates synthetic data to overcome the shortage of experimental data. After adding synthetic data to enhance the training of the model, the performance of the model was largely improved, reaching 93.17 % in accuracy. In addition, the deep learning-based ANN allows interpretability of design features. This study created a comprehensive guidance in the MPEA design and also demonstrated the capability of generating similar and novel compositions via generative deep learning models.

As the framework of phase prediction using machine learning is gradually improved, research tends to uncover the relationships between microstructures and thermodynamic properties to reach higher predictability and accuracy. Some studies created new thermodynamic parameters for training while others focused on feature engineering with rational selection of materials descriptors. There are several materials descriptors correlated with MPEAs. For example, Yang et al. proposed the parameters to predict the phase formation for the HEAs [64]. Their study indicated that the HEAs were stabilized as solid solution phase. Xie et al.'s MD simulation in the AlCoCrCuFeNi MPEA thin film growth results [65] were in excellent agreement with Yang's solid-solution formation rules. Valence electron concentration (VEC) is another key feature to predict the phase stability of the MPEAs. Using this information, Chen et al. employed the VEC to design the HEAs with strength-ductility balance [66].

2.4 Genetic Algorithm

Genetic algorithm (GA) is a type of optimization algorithm inspired by the process of natural selection and evolution. It is used to find solutions to optimization and search problems by mimicking the process of natural selection to evolve a population of solutions over successive generations. GA is particularly useful for problems with large and complex search space, making it a viable candidate to be employed in this study.

Before implementation of GA, the data needs to be encoded such that each individual solution is depicted as a chromosome with each gene within a chromosome signifying decision variables.

Improvement of solutions via GA hinges on two key operators, crossover and mutation. Crossover creates new solutions by mixing solutions from previous generation, and mutation randomly changes a decision variable of a particular solution. The process of GA can be depicted as follows:

1. Initialization of population to represent search space;
2. Evaluation of fitness through an objective function;
3. Selection of suitable individuals for reproduction;
4. Crossover through combination of individuals;
5. Mutation by randomly altering an offspring's genetic code.

This is an iterative process where each iteration is termed as a generation. In every generation, evaluation, selection, crossover and mutation occur and generations proceed until a termination criterion is satisfied, which could be number of generations or minimum tolerance in fitness value. The performance of GA can be sensitive to choice of parameters such as population size, crossover rate, mutation rate and selection criterion.

2.4.1 Advantages of Genetic Algorithm

One of the primary advantages of GA is their capability to search globally. Unlike gradient-based optimization techniques, GA does not depend upon derivative information and is therefore less susceptible to being trapped in a local optima. This property makes them particularly well suited for highly non-linear and multimodal search spaces, making it a viable candidate for phase prediction of MPEAs, as well as other multicomponent alloy systems [67].

Since GA is population-based in nature, it enables simultaneous exploration of multiple regions of the search space. This parallel behaviour enhances robustness and increases the probability of identifying high-quality solutions in problems with a complex fitness landscape. Population diversity also facilitates implicit information sharing via genetic operators like mutation and crossover, which promotes improvement in solutions [14].

Another significant benefit of using GA is its flexibility. It can be readily adapted and employed for continuous, discrete and multi-objective optimization problems [68]. Their applicability is extended by techniques such as custom encoding [69] and constraint-handling [70]. As a result, GA has been successfully applied in various domains from structural optimization to material design and machine learning.

2.4.2 Disadvantages and Limitations of Genetic Algorithm

Despite their versatility, GA suffers from several drawbacks. A major limitation of GA is its computational cost. Evaluation of fitness for a large population over multiple generations can be computationally time-intensive specially for multi-dimensional problems, such as the modelling of multicomponent alloys [71].

Another disadvantage is the lack of guaranteed convergence to global optimum. GA is effective for global search but its heuristic nature implies that convergence is not a guarantee, rather it is probabilistic. This can be due to bad parameter setting, lack of population diversity or premature convergence, all of them leading to suboptimal solutions [72].

The performance of GA is sensitive to the choice of encoding and genetic operators. Encoding schemes, crossover strategies, mutation rates and selection criteria must be carefully designed to match the problem. Inappropriate choices can significantly reduce performance and identifying optimum parameters usually requires extensive tuning or domain knowledge. The reliance of GA on tunable parameter selection hinders its generality and increases complexity [72].

2.5 Teaching-Learning Based Optimization

TLBO is a meta-heuristic optimization algorithm inspired by the environment of a classroom, where students learn from a teacher and disseminate information amongst each other to improve grades of the class.

Optimization is defined in terms of a cost function $f(X)$ which is dependent on a set of independent variables X , where the aim is to find the set of values of X which minimize $f(X)$. We can formulate an optimization problem as: Minimize $f(X)$, subject to $X_i \in [x_i^{\min}, x_i^{\max}]$, where D_n is the dimensions, or rather the number of independent variables under consideration.

TLBO consists of two phases: teaching phase and learner phase. In teaching

phase, students learn from the teacher and improve their fitness. Teacher is defined as the best student in the class.

$$T_i^k = X_{\min f(x)}^k \quad (2.2)$$

where T_i^k denotes the teacher and $X_{\min f(x)}^k$ is the student with the best fitness in generation number k . In teaching phase the population learns from the teacher and updates itself according to the following relation:

$$X_i^{k+1} = X_i^k + r_i(T_i^k - T_f M_i^k) \quad (2.3)$$

where M_i^k is the mean of learners in generation k and r_i is a uniformly distributed random number between 0 and 1. T_f is the teaching factor which can have value of either 1 or 2 and is defined as:

$$T_f = \text{round}[1 + \text{rand}(0, 1)] \quad (2.4)$$

In learner phase, students learn from each other through mutual interaction. We randomly select two students from the population and update their values according to the following criterion:

$$X_i^{k+1} = \begin{cases} X_i^k + r_i(X_i^k - X_j^k), & \text{if } f(X_i^k) < f(X_j^k) \\ X_i^k + r_i(X_j^k - X_i^k), & \text{if } f(X_i^k) > f(X_j^k) \end{cases} \quad (2.5)$$

here, X_i^k and X_j^k represent two distinct learners, such that $i \neq j$, and r_i is a uniformly distributed random number within the range $[0, 1]$. Learner phase takes into consideration the relative difference in cost function values of students such that the update is in the direction of the better student. To ensure this, Eq 2.5 is defined piecewise, where if X_i^k is better than X_j^k , the update is $X_i^k - X_j^k$ in the direction of X_i^k , and vice versa. The entire process of TLBO is schematically illustrated as a flowchart depicted in Fig. 2.2.

The parameter-free nature of TLBO is particularly appealing when compared with other population based intelligent algorithms, where the tuning of several parameters becomes a time-consuming process (e.g., crossover rate, mutation rate, inertia weights). Researchers have extended the canonical TLBO over time, creating various derivative forms which address different challenges posed by complex

search spaces.

2.5.1 Advantages of TLBO

A prominent advantage of TLBO is its minimal reliance on algorithm-specific control parameters. Unlike GA and other evolutionary algorithms, TLBO does not require explicit tuning of crossover rates, mutation rates or inertia weights. This characteristic simplifies the implementation of TLBO and reduces risk of suboptimal performance due to bad selection of tunable parameters [15].

TLBO also demonstrates global search capability due to its population based structure. The teacher phase aids in convergence towards better regions in the search space by guiding the population mean towards the best learner [73]. The learner phase promotes information exchange among students, leading to enhanced exploration. This two-phase learning mechanism provides a balance between exploration and exploitation, hence enabling TLBO to perform effectively on nonlinear and high-dimensional problems.

Another advantage is the simplicity and generality of TLBO. The core update equations of teaching and learner phases are straightforward and problem-independent. Hence, the algorithm can be applied readily to continuous, discrete and multi-objective optimization problems. TLBO has been successfully implemented in various domains such as mechanical design [74], power systems [75], scheduling [76] and machine learning parameter tuning [77]. This demonstrates the versatility of TLBO as well as its ease of implementation.

2.5.2 Disadvantages and Limitations of TLBO

Despite its strengths, TLBO is subject to several limitations. A drawback frequently encountered in research is its relatively slow convergence in the later stages of optimization as well as origin bias [28], [78], [79]. The teacher phase is responsible for accelerating early convergence, however lack of local exploitation search can hinder the algorithm when it is near the optimum. This is specially the case for high-dimensional or ill-conditioned problems.

Another limitation of TLBO is concerned with lack of population diversity. The deterministic selection of a single teacher and update based on population mean can lead to homogenization of the population [80]. This increases the risk of premature convergence since the population becomes more biased towards being stuck in a

local minima. This becomes a critical issue for complex multimodal systems where maintaining a diverse population is crucial for successful convergence.

Although TLBO is described as parameter-free, this claim is not absolute. Practical implications when applying TLBO still require the selection of population size and stopping criteria, similar to other population-based optimization techniques. Most of the modifications added to TLBO to improve its performance introduce additional tunable parameters, like adaptive learning rates [81], [82], [83], probabilistic teaching factors [84] or stochastic perturbation terms [85]. This diminishes the original simplicity of the algorithm.

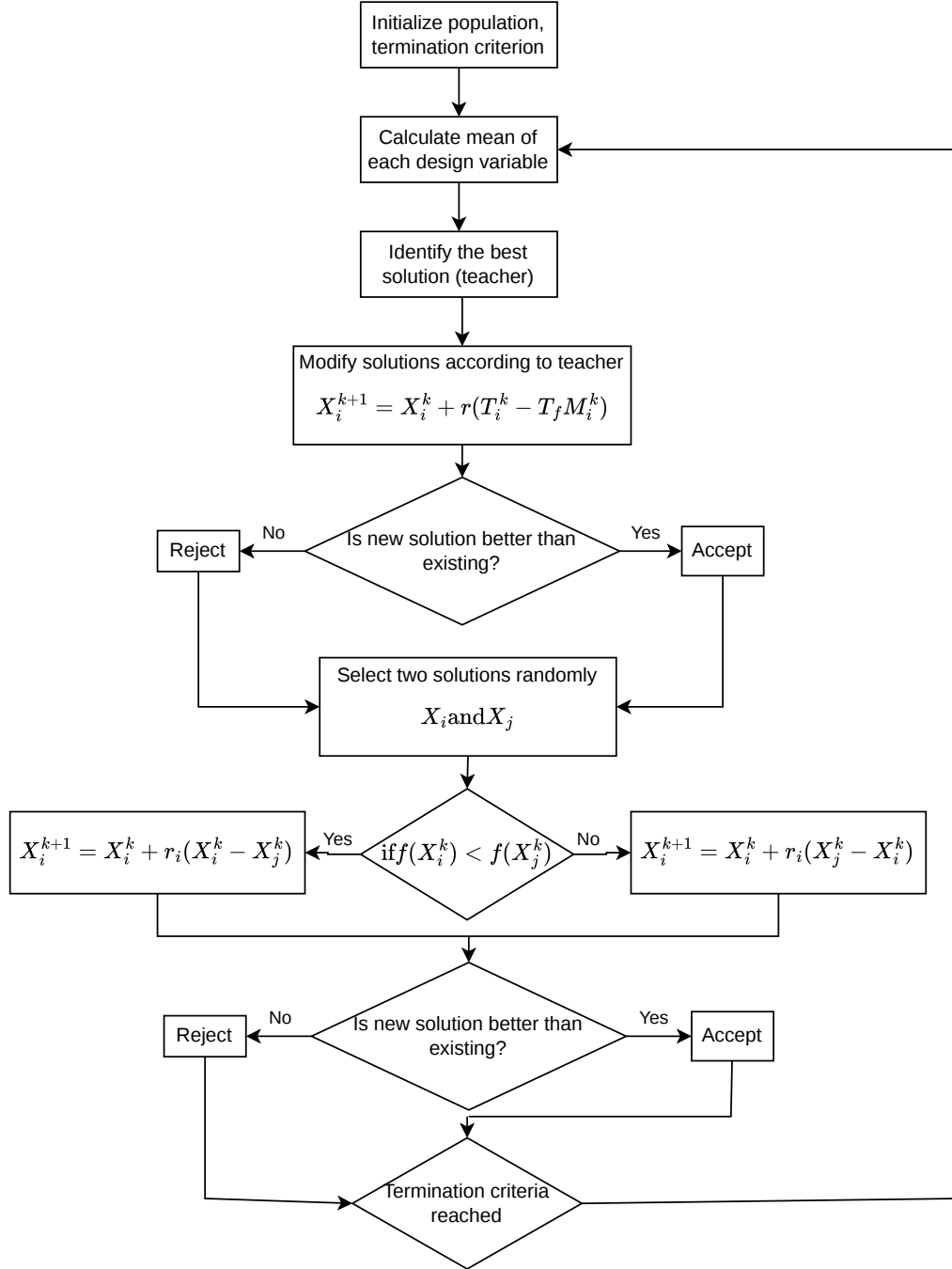


Figure 2.2: Flowchart of TLBO, depicting teaching and learner phase.

Chapter 3

A Novel Hyperparameter Optimization Approach for Supervised Classification: Phase Prediction of Multi-Principal Element Alloys

3.1 Motivation

In contrast with conventional alloys, MPEAs contain multiple principal elements with (near) equiatomic concentrations. They have been shown to exhibit excellent mechanical and functional properties [86] compared with conventional alloys and are expected to be applied to a wider range of fields than conventional alloys. However, the existence of multiple phases as well as a large search space due to the number of possible combinations, makes it tough to select elements and their compositions to make MPEAs with desirable properties. Experimental study of such alloys has been predicated on trial-and-error based approaches which prove to be insufficient due to the presence of a large search space making trial-and-error alloy discovery infeasible, compelling the use of data science approaches for alloy discovery.

With availability of large datasets and advancements in algorithms, there has been a recent surge of interest in the field of machine learning. Machine learning techniques have been used in various applications of material science such as property analysis, discovery of new materials and quantum chemistry [87]. This

includes various approaches of supervised, unsupervised and reinforcement learning techniques which have changed the way material scientists approach problems. This interest has made its way into the study of alloys and has lead to studies on machine learning accelerated atomistic simulations, which includes inter-atomic potentials, molecular dynamics and Monte Carlo simulations, as well as the study of machine learning assisted physical property prediction.

Among the myriad of aforementioned applications, phase prediction of MPEAs has received considerable attention due to the fact that alloy phases tend to guide physical properties [88]. Certain phases are more desirable than others when considering mechanical performance of alloys, such as ductility and yield strength. Single-phase face-centred cubic (FCC) alloys have relatively high ductility but low strength, whereas single-phase body-centred cubic (BCC) alloys have high strength but low ductility which makes them brittle and susceptible to damage. Dual-phase MPEAs can possess a balance between strength and ductility [89]. These guiding factors are important when considering alloy applications and there is a need for a universal indicator of phase prediction of MPEAs. Unfortunately, there does not exist a solid theoretical framework for phase prediction due to the sheer astronomical amount of possible materials. An estimate of possible alloys which could be formed is $\sim 10^{78}$ [46], which is an enormous number. Only a tiny fraction have been synthesized and studied, most of them being conventional alloys. This renders the current trial-and-error method for exploring new alloy configurations obsolete, prompting the use of enhanced search techniques, which may be either model-driven or data-driven.

Phase prediction of MPEAs when translated to machine learning, becomes a supervised multi-class classification problem, where each class signifies a particular alloy phase, and these phases are treated as mutually exclusive. Many studies have been performed tackling this problem using various features and strategies [60], [90], [91], [92], [93], [56], [63]. These studies include the use of artificial neural networks, support vector machines, random forests and deep learning with generative adversarial networks.

It is worth mentioning here that Artificial Neural Networks (ANNs) have been widely used to create various predictive models. ANNs are computational models inspired by the structure and functioning of biological neurons, such as the human brain [94] [95]. They consist of interconnected nodes, called neurons, organized in layers. Each neuron receives input signals, processes them using an activation function, and produces an output signal. ANNs are well suited for high dimensional

data since they are able to learn complex non-linear relations between input and output data. They are flexible in handling numerical and categorical inputs, and are easily scalable to the size of data.

A major challenge in the use of ANNs for predictive modeling is determination of hyperparameters. There exist numerous adjustable parameters which need to be tuned prior to classification, such as number of neurons per layer, learning rate and regularization strength. Finding the optimal set of hyperparameters is a challenging task but necessary task which will eventually influence the performance of the model. Usually, these parameters are based on heuristic rules and are manually modified, which can be time-consuming. Hyperparameter optimization is a process which identifies the best set of hyperparameters that maximizes performance of a model [96], [97]. This problem can be approached as optimizing a loss function over a constrained configuration space [98]. Multiple approaches for optimizing hyperparameters have been used in the past, including grid search, random search, Bayesian optimization and evolutionary algorithms.

Among various optimization techniques, Genetic algorithm (GA), a subset of evolutionary algorithms, is a powerful and flexible tool. GA is a nature-inspired algorithm based on natural selection and Darwinian evolution [99]. They are gradient-free and well suited for high-dimensional and complex systems. GA has been utilized for hyperparameter optimization and has shown significant improvements compared to grid search and Bayesian optimization [100]. Their population-based approach allows parallelism as well as diversity in exploration. Due to the technical merits of GA, it has been widely applied to hyperparameter optimization of ANNs [101], convolutional neural networks [102] and long short term memory networks [103].

The approach of ANN predictive modelling for identifying MPEAs has been utilized in literature before, but there has been no instance of careful hyperparameter selection. This might be due to the fact that application of machine learning to MPEAs is a relatively recent interest of the scientific community, and there exists ample room for improvement. Hyperparameter optimization has the ability to significantly improve predictive capabilities ANN, especially when dealing with complex and noisy data containing outliers [104]. Due to the nature of experimental errors [105], MPEA data obtained from literature is fraught with bad quality data in the form of imbalanced class distributions, sparsity and noise in the form of outliers. Any machine learning approach applied to MPEA prediction must have the capability to deal with these issues. In our study, one of our key objectives is

to incorporate outlier detection in the process of training instead of including it in pre-processing by integrating it within hyperparameter optimization. This will serve to clean the data within the training process and improve the predictive modelling of MPEAs.

Motivated by the above discussions, the objective of this chapter is to develop a framework of supervised classification via hyperparameter optimization with inclusion of novel hyperparameters, whose presence will be helpful in dealing with outliers and feature selection. Specifically, a boxplot based outlier detection and removal method as well as optimal feature subset choice. This strategy is employed to enhance the classification ability of ANNs when dealing with bad quality data. The proposed methodology is applied to phase classification of MPEA.

The main contributions of this chapter can be summarized in the following aspects:

1. a novel hyperparameter optimization algorithm is proposed with two new parameters: outlier detection and feature subset selection; and
2. the aforementioned algorithm is applied to phase classification of MPEAs. Experimental results demonstrate the effectiveness of the proposed hyperparameter optimization method.

The aforementioned contributions are related to first two major and first two minor contributions from Section 1.5.3. Creation of novel hyperparameter optimization technique and application of this algorithm to phase identification of MPEAs are major contributions, and compilation of database as well as the use of GA as a meta-heuristic optimizer for MPEA phase classification are minor contributions described in this chapter.

The remaining parts of this chapter are organized as follows. The background of MPEA phase classification is discussed in Section 3.2, as well as description of utilized data sets. In Section 3.3, the proposed hyperparameter optimization strategy via static and dynamic length chromosome GA is introduced. Experimental results of phase classification of MPEAs are presented Section 3.4. Finally, conclusion and discussion about suitable future work are suggested in Section 3.5.

3.2 Background

3.2.1 Multi-Principal Element Alloys

As described in Section 2.2, MPEAs are a new form of alloys which are attracting increased attention from researchers due to exhibition of excellent mechanical and physical properties. Determining the stable phases present in MPEAs has been an important aspect of their study since the resulting phases influence mechanical properties of alloys, hence affecting their potential applications. One fundamental question regarding MPEAs is what kind of atomic phase will form when large number of different elements are mixed together. Surprisingly, the resulting phases are simple structures [1]. Classic Hume-Rothery rules define the factors affecting formation of solid solutions in binary systems, where only two element types are present. These factors include atomic size mismatch, valence electron concentration, electronegativity, enthalpy of mixing and configurational entropy, amongst other things [16]. Studies done on phase formation of MPEAs obtained conclusions similar to Hume-Rothery rules for binary systems [43], [106].

Unlike parametric approaches utilized in the past [43], machine learning based phase prediction is a robust tool which can offer insight from a given data with relatively low bias and high efficiency. Accurate and fast algorithms have been developed in recent years for purpose of material discovery [107], [108], [109]. These approaches serve to guide experimental analysis of complex multi-dimensional systems such as MPEAs and have become a vital tool for material discovery in recent years.

3.2.2 Description of Data Set

Sources of MPEA Phase Data

The data utilized in this study is obtained from various sources in literature [4], [59], [61], [7]. After removal of duplicates and alloys with radioactive elements, the merged dataset consists of 2218 compositions ranging from binary to multi-component alloys. In the compiled dataset, each data point represents an individual alloy composition, along with their respective chemical features and atomic phase.

While the size of the dataset is comparable to many publicly available MPEA datasets, it is small from the perspective of machine learning, where typical datasets could have thousands to millions of samples. However, unlike image or text datasets,

material science datasets are generated through expensive and time-consuming experiments which makes obtaining large datasets difficult. The compositional space of MPEAs is extremely large, with millions of possible alloy compositions resulting from combinations of multiple principal elements with varying concentrations. Hence, the available dataset only represents a small fraction of the feasible search space. Also, the dataset exhibits class imbalance, with some alloy phases represented by limited number of samples.

Characteristics of Features

We have considered ten features and following equations are used to obtain their numerical values ([43], [110], [111], [112]):

$$\bar{Z} = \sum_{i=1}^n c_i Z_i \quad (3.1)$$

$$\bar{r} = \sum_{i=1}^n c_i r_i \quad (3.2)$$

$$\bar{m}_a = \sum_{i=1}^n c_i m_{a,i} \quad (3.3)$$

$$\text{VEC} = \sum_{i=1}^n c_i \text{VEC}_i \quad (3.4)$$

$$\Delta\chi = \sqrt{\sum_{i=1}^n c_i (\chi_i - \bar{\chi})^2} \quad (3.5)$$

$$\delta = 100 \sqrt{\sum_{i=1}^n c_i (1 - r_i/\bar{r})^2} \quad (3.6)$$

$$T_{\text{melt}} = \sum_{i=1}^n c_i T_i \quad (3.7)$$

$$\Delta S = -R \sum_{i=1}^n c_i \ln c_i \quad (3.8)$$

$$\Delta H_{\text{mix}} = \sum_{i=1, i < j}^n 4H_{ij} c_i c_j \quad (3.9)$$

where n is number of elements in an alloy, c_i is atomic concentration of i^{th} element, $\bar{Z}$ is mean atomic number, Z_i is atomic number of i^{th} component, $\bar{r}$ is mean atomic radius, r_i is atomic radius of i^{th} element, $\bar{m}_a$ is mean atomic mass, $m_{a,i}$ is atomic

mass of i^{th} element, VEC is mean valence electron concentration, VEC_i is valence electron concentration of i^{th} element, $\Delta\chi$ is Pauling electronegativity difference, χ_i is electronegativity of i^{th} element, δ is atomic size mismatch, T_{melt} is mean melting temperature, T_i is melting temperature of i^{th} element, ΔS is configurational entropy, R is ideal gas constant, ΔH_{mix} is enthalpy of mixing and H_{ij} is enthalpy of binary alloy i, j . Values of binary mixing enthalpy are obtained from Miedema model [113], [114], and required elemental data for calculation of features is obtained from [106].

Characteristics of Classes

The phases are categorized as follows: face-centered cubic (FCC), body-centered cubic (BCC), hexagonal closed pack (HCP), mixed phases (MP), amorphous (AM), intermetallic (IM) and BCC+FCC. We can see that there is a distinct imbalance present in our class distribution. This is due to the fact that some alloy phases are studied more than others due to their potential applications and interest [4]. For MPEAs, single-phase solid solutions (FCC, BCC, HCP) are the target of interest because their presence correlates with excellent mechanical properties, whereas presence of intermetallics is considered undesirable due to their negative influence on mechanical behaviour [86], [46]. The details of data pre-processing are presented in Section 3.4.

Problems of Dataset

The compiled alloy dataset contains class imbalance and outliers. The class imbalance is depicted in Fig. 3.1 which illustrates distribution of phases. The MP class contains majority of alloy samples, and IM class contains the least. Also worth noting are BCC and FCC classes which are single-phase solid solutions with significant number of alloys. Due to the nature of metallurgic experiments and presence of systematic and random errors in experimental analysis, MPEA data becomes noisy and unsuitable for predictive modelling. Instead of choosing to modify noise, in this study we have chosen to identify and remove outliers. The physical nature of alloys makes it difficult to modify experimental data using statistical methods. Changing values of physical parameters might lead to unphysical depiction of alloy systems. For this purpose, we instead focus on outlier detection which constitutes removal of data points instead of any modification. This approach will make our phase prediction model reliable for use by metallurgists in search for MPEAs with desirable

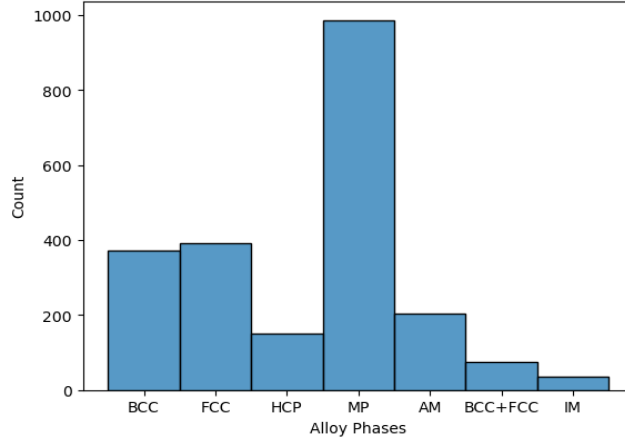


Figure 3.1: Distribution of classes in alloy dataset depicting imbalance

properties.

3.3 Methodology

3.3.1 ANN based Classification

To create a framework of phase prediction, we formulate a supervised classification approach where each alloy composition has ten distinct features as inputs and phase as output class. ANN architecture can be described with consecutive input, hidden and output layers. Hence, an initial network is proposed with two hidden layers and dropout regularization. Dropout is a regularization method used to prevent overfitting and is easy to implement without much overhead in computation [115]. Each hidden layer has ReLU as activation function [116]. Cross-entropy is chosen as loss function in the network. This function can be defined as:

$$\text{loss} = -\frac{1}{N} \sum_{n=1}^N \sum_{i=1}^K T_{ni} \ln Y_{ni} \quad (3.10)$$

where N is number of observations, K is number of classes, T_{ni} is target value and Y_{ni} is predicted value of the network.

Training process of ANN is an iterative process where series of linear operations between consecutive layers are used to improve the performance of a network by tuning network parameters i.e., weights and biases. The operation process can be defined as:

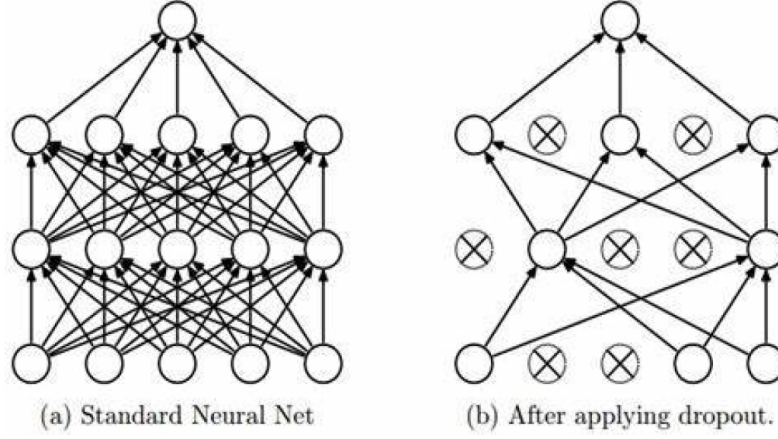


Figure 3.2: Depiction of dropout in neural networks by comparing networks with and without applying dropout [118].

$$x_j^{(l)} = f \left(\sum_i w_{i,j}^{(l)} x_i^{(l-1)} + b_j^{(l)} \right) \quad (3.11)$$

where $x_j^{(l)}$ is the output of j^{th} neuron in l^{th} layer, $x_i^{(l-1)}$ is input of i^{th} neuron in previous layer, whereas $w_{i,j}^{(l)}$ and $b_j^{(l)}$ are weights and biases of j^{th} node in l^{th} layer. f represents a non-linear activation function.

The architecture of ANN is crucial in determining its efficacy and implementation [117]. The architecture chosen for the current study contains an input layer of empirical Hume-Rothery features, two hidden layers and an output layer with each node providing the probability of each class/phase.

A key point to note here is that our network is not fully-connected due to the presence of dropout regularization, which removes connections between nodes randomly via a fixed probability $p_{dropout}$. This process is vital in overcoming overfitting, which occurs when a model fits too closely to its training data, depreciating its ability to predict testing data. An added benefit of using dropout is that it makes ANN lightweight and significantly decreases computation complexity by removing node connections. This process is illustrated in Figure 3.2.

3.3.2 Genetic Algorithm

As described in Section 2.4, Genetic algorithm (GA) is a type of optimization algorithm inspired by the process of natural selection and evolution. It is used to find solutions to optimization and search problems by mimicking the process of natural selection to evolve a population of solutions over successive generations.

Improvement of solutions via GA hinges on two key operators, crossover and mutation. Crossover creates new solutions by mixing solutions from previous generation, and mutation randomly changes a decision variable of a particular solution. This is an iterative process where each iteration is termed as a generation. In every generation, evaluation, selection, crossover and mutation occur and generations proceed until a termination criterion is satisfied, which could be number of generations or minimum tolerance in fitness value. The process of GA is illustrated in Fig. 3.3 in the form of a flowchart. The performance of GA can be sensitive to choice of parameters such as population size, crossover rate, mutation rate and selection criterion. In the current study we will employ GA for hyperparameter optimization by incorporating a novel strategy for dealing with variable-sized chromosomes.

Encoding data within the framework of genes is imperative for GA. Vast majority of GA studies employ binary encoding, where data is translated into 0s and 1s, but in our case, due to the presence of real numbered as well as categorical data, it is more suitable to utilize real encoding instead. Each chromosome is formed with N genes, where N is the number of hyperparameters in our problem. For the current study, $N = 7$, as depicted in Table 3.1.

A comprehensive depiction of GA employed in the current study is shown as a pseudo code in Algorithm 1, where, r is a uniformly distributed random number between 0 and 1, and p_{mut} is the probability of mutation, such that mutation is accepted only when $r < p_{mut}$. Both Fig. 3.3 and Algorithm 1 in conjunction describe the process of GA as it aims to improve the quality of solutions iteratively and move towards global minima of the search space under consideration.

Algorithm 1 The Procedure of Genetic Algorithm

- 1: Initialize population of hyperparameters with population size P .
 - 2: Calculate fitness of each individual.
 - 3: Arrange population in increasing order of fitness
 - 4: Select two individuals. Select a random point of intersection and apply crossover at that point to generate new offspring.
 - 5: If $r < p_{mut}$, select a random decision variable and initialize it with a different value within its respective range.
 - 6: Select the best 10% of current population and keep them for the next generation
 - 7: Terminate the algorithm if maximum number of generations is reached. If not, repeat Steps 2-6.
-

In the current study, we employ single-point crossover and single-point mutation with fixed probability $p_{mut} = 0.1$. Selection is based on the fitness of each individ-

ual, and instead of replacing all members of previous generation, we keep the best few individuals from the current generation to pass onto the next. This method is known as elitism and it ensures that the best solutions found so far are preserved across generations, preventing the loss of valuable genetic information due to random variation. Elitism helps maintain a certain level of stability and ensures that promising solutions are not lost prematurely. It can be particularly useful when the fitness landscape is dynamic or when the algorithm is prone to premature convergence, where the population converges prematurely to suboptimal solutions. By preserving the best individuals, elitism helps guide GA towards better solutions over time.

Dynamic Length Chromosome Operators

Canonical GA and its respective operators are designed on the basis of fixed number of decision variables, where each variable represents a gene in a fixed-length chromosome, and multiple chromosomes form a collective population. A major flaw in this approach is that GA is not suited to handle variable length data, whether its variable number of genes or variable length of one particular gene. To overcome this deficiency, studies have been conducted in the past that developed approaches for variable length chromosome GA by modifying crossover and mutation operators to suit the problem, such as topology design [119] [120], road traffic problem [121] and hyperparameter optimization [103].

GA strategies exist for both fixed and variable length chromosomes. However, a gap in current research knowledge is the ability to deal with variable length genes i.e., the capability to handle chromosomes containing genes with single and multiple values. An illustration of this phenomenon is shown in Fig. 3.4, where a normal chromosome is contrasted with a variable gene length chromosome.

Looking at Fig. 3.4, we can discern that there is a need of modified crossover and mutation operators capable of handling a chromosome containing both single-valued and multi-valued genes, where the length of multi-valued gene can vary. This is imperative for the current approach because feature subset selection can have variable number of elements, whereas all the other hyperparameters are single-valued. For this purpose, a robust framework of genetic operators is developed which will be able to effectively perform crossover and mutation for our problem.

To make the problem easy to deal with, we define two different crossover and

mutation operators, one for single length genes and other for the variable length gene. Single length genetic operators are the same as normal crossover and mutation, where parents are segmented and joined together as well as randomly mutated at a single point. For variable length operators, we deal with the only gene which has variable array length. Hence, both single and variable length operators are functioning simultaneously for an individual in a population.

Variable length crossover functions by selecting variable length gene from two parents, ensuring that length of parent 1 is less than length of parent 2 i.e., $len(P_1) < len(P_2)$. Next, we define length of offspring within range $[len(P_1), len(P_2)]$, hence length of offspring always lies between length of both parents. Then we select point of crossover c within range $[1, len(P_1)]$ such that the point lies within bounds of P_1 . The first half of offspring O_1 is initialized by selecting values within P_1 until point of crossover c , such that $O_1 = P_1(1, \dots, c)$. The second half of O_1 is defined by checking each member of P_2 . If the member is not present in O_1 , add it to O_1 , otherwise reject it. Via this process, we avoid repetitions and ensure the characteristics of both parents are represented within the resulting offspring. Depiction of this process is shown as pseudo code in Algorithm 2.

Algorithm 2 The Procedure of Variable Length Crossover

- 1: Select parents P_1 and P_2 such that $len(P_1) < len(P_2)$.
 - 2: Define length of offspring O_1 within $[len(P_1), len(P_2)]$ range.
 - 3: Select point of crossover c within range $[1, len(P_1)]$.
 - 4: Initialize first half of O_1 with $P_1(1, \dots, c)$.
 - 5: Initialize second half of O_1 by looping over P_2 and selecting members not present in O_1 .
-

The process described in Algorithm 2 is a novel contribution mentioned previously in Section 1.5.3, where a systematic approach of handling crossover with variable length solutions will allow us to incorporate feature subset selection within the framework of hyperparameter optimization, illustrated in Fig. 3.4.

For variable length mutation, we apply the same strategy as mutation for traveling salesman problem, where we simply generate a random permutation of length l , where l is a randomly selected length within the bounds of decision variable. Hence, we are sampling from a distribution of random permutations of variable length decision variable.

3.3.3 Hyperparameter Optimization

In ANN, there exist many tunable parameters, termed as hyperparameters which influence its performance and control aspects of learning process itself. These parameters are usually selected on ad hoc basis, however this is an ineffective practice, limiting the capabilities of ANN. Hence, hyperparameter optimization is employed to select optimum parameters via an optimization technique which will serve to maximize network performance, in other words testing accuracy.

For this study, hyperparameter optimization is designed as a minimization problem, where its aim is to find hyperparameters of a given model that return the best performance when measured on a test set. This process can be represented mathematically as:

$$x^* = \arg \min_{x \in \Theta} f(x) \quad (3.12)$$

where $f(x)$ is the fitness function to be minimized; x^* is the set of hyperparameters which yield the lowest value of objective function, where x can take any value in the domain Θ which consists of all possible combinations of hyperparameters. The fitness function in this study is $f(x) = 1/T_{acc}$, where T_{acc} is testing accuracy of the ANN. By defining objective function as inverse of testing accuracy, we create a minimization problem to be solved by GA.

Canonical hyperparameters for ANN considered in this study are:

- Learning rate
- Batch size
- Number of epochs
- Number of nodes in each hidden layer
- Rate of dropout

These parameters have been widely used in literature with varying types of optimization techniques and datasets, as well as problems of both classification and regression [122], [123], [124].

In this study, we incorporate two new ideas as hyperparameters in the framework of optimization: outlier detection and feature subset selection. This is a novel contribution to the field of hyperparameter optimization as well as data pre-processing previously mentioned in Section 1.5.3, where a methodology is developed such that

data cleaning and architecture optimization can be performed simultaneously and incorporated within the hyperparameter optimization framework as additional tunable hyperparameters. These topics are discussed further in the following subsection.

Outlier Detection

Outlier detection is the process of identifying data points or observations that deviate significantly from the majority of the data in a dataset. These outliers are typically observations that are unusual compared to the rest of the data and may indicate anomalies or errors. A well known statistical approach to identify outliers in a distribution is the boxplot method [125]. A boxplot provides visual summary of statistics for sample data and contains information about 25th and 75th percentile, maximum and minimum as well as median of the distribution. Outliers in terms of boxplots are defined as data points outside the interquartile range times a constant k_{iqr} , where interquartile range is the range between 25th and 75th percentile, such that $25^{th} \text{percentile} \geq \text{outlier} \geq 75^{th} \text{percentile}$. The constant value essentially determines the range beyond which any data point is considered an outlier. A large value of constant will consider less data points as outliers and vice versa. This makes the decision of k_{iqr} important since we want to eliminate the maximum number of outliers without affecting the information contained within each distribution. A visual depiction of boxplot is illustrated in Fig. 3.5.

The trade-off associated with k_{iqr} and number of observations considered by the model is illustrated in Fig 3.6, where the value of k_{iqr} is influencing number of outliers considered by boxplot method. It is quite evident that low k_{iqr} is extremely undesirable, since it is identifying large chunks of data as outliers and eliminating them from the dataset. A reasonably moderate to high value of k_{iqr} is needed for efficient identification and removal of outliers, hence k_{iqr} is considered as a hyperparameter for our model and its value will be determined via hyperparameter optimization.

Feature Subset Selection

Feature subset selection is a process where you choose a subset of relevant features (variables, predictors) from a larger set of features to build a model. The goal is to improve the model's performance by reducing overfitting, simplifying the model, and potentially improving interpretability. In many real-world datasets, there may be numerous features, but not all of them contribute equally to the predictive capability

of the model. Some features may be redundant or irrelevant, adding noise to the model or causing overfitting. Feature subset selection methods aim to identify and select the most informative and relevant features while discarding the less useful ones. Feature subset selection is crucial for building efficient and effective machine learning models, especially when dealing with high-dimensional data.

Many feature subset selections methods have been utilized in the past, such as correlation-based filtration [126], wrapper methods [127] and embedded methods such as LASSO [128]. Each method has its merits and demerits, however as of yet feature subset selection has not been incorporated within the framework of hyperparameter optimization, which will make selection of optimum features in conjunction with selection of suitable network hyperparameters, removing extra computational overhead and optimizing selection to maximize testing accuracy.

A comprehensive list of hyperparameters considered in this study as well as their respective ranges are depicted in Table 3.1. Here, range of feature subset signifies lowest and highest number of possible features to be considered by the algorithm, which in our case is 5 and 10.

Table 3.1: List of hyperparameters and their respective ranges.

Hyperparameter	Range
Learning rate	$[10^{-4}, 10^{-1}]$
Batch size	[5, 200]
Number of epochs	[10, 1000]
Number of nodes	[5, 50]
Dropout rate	[0, 1]
k_{igr}	[1.16, 3]
Feature subset	[5, 10]

3.4 Experimental Results

3.4.1 Data Statistics

For an accurate depiction of data landscape, descriptive statistics are a useful tool for summarizing and describing the main features of a dataset. They provide a concise overview of the data, allowing for easy interpretation and understanding. Descriptive statistics of features considered in this study are listed in Table 3.2, where the statistical parameters include mean, median, mode, range, variance, standard deviation, skewness and kurtosis. Mean, median and mode are used to describe

central tendency and provide information about the center and average of a feature. Range, variance and standard deviation describe variability of features and helps to understand how much the data points vary from the center. Skewness and kurtosis describe distribution of data and provide insight into symmetry and shape of distribution. From Table 3.2, we can see that ΔH_{mix} and $\bar{r}$ have high positive skewness, which is evident from the fact that majority of binary H_{mix} values lie within range $[-50, 50]$, but the long positive tail of distribution signifies presence of alloys with large positive ΔH_{mix} . The long positive tail of $\bar{r}$ can be explained from the presence of alloys with elements having large metallic radii, such as potassium, rubidium and caesium. Features with low skewness indicate centralized distribution with low asymmetry, as is the case for ΔS . This can be attributed to the presence of both low entropy conventional alloys and high entropy MPEAs, balancing out the distribution on either sides of the median. Presence of features with high positive kurtosis, such as ΔH_{mix} will most likely generate large number of outliers, whereas platykurtic features with low negative kurtosis are least likely to contain outliers since their distribution is closer to thin-tailed normal distribution, as is the case for n and ΔS .

Table 3.2: Descriptive Statistics values of features in our dataset.

Feature	Mean	Median	Mode	Range	Variance	Standard Devia- tion	Skewness	Kurtosis
n	3.71	4	2	8	3.13	1.77	0.37	-1.19
$\bar{Z}$	97.31	32.48	27	78.25	236.55	15.38	0.92	0.06
$\bar{m}_a$	8.7×10^1	7.39×10^1	3.22×10^1	2.17×10^2	1.65×10^3	4.06×10^1	9.88×10^{-1}	1.58×10^{-1}
$\bar{r}$	1.41	1.36	1.29	1.45	2.86×10^{-2}	1.69×10^{-1}	2.18	6.79
VEC	6.65	7.0	8	13	4.31	2.08	-1.67×10^{-1}	-2.35×10^{-1}
$\Delta\chi$	1.74	1.76	2.2	1.7	7.83×10^{-2}	2.8×10^{-1}	-1.81×10^{-1}	4.91×10^{-1}
δ	6.27×10^{-2}	5.16×10^{-2}	4.34×10^{-3}	3.36×10^{-1}	2.70×10^{-3}	5.19×10^{-2}	1.62	3.32
T_{melt}	1.88×10^3	1.82×10^3	1.12×10^3	3.37×10^3	3.54×10^5	5.95×10^2	1.57×10^{-1}	4.63×10^{-1}
ΔS	9.05	9.13	5.76	1.83×10^1	1.83×10^1	4.27	5.8×10^{-2}	-1.49
ΔH_{mix}	-6.33	-5.89	0	2.07×10^2	3.33×10^2	1.82×10^1	2.51	1.6×10^1

As depicted in Fig. 3.1, the imbalance present in our dataset will severely affect distribution and sampling of data in training and testing sets. To rectify this and

prevent skewed sampling, we utilize stratified sampling when separating data into training and testing sets. This ensures that the proportions of each class present in both sets is the same, and every minor class is represented in testing set, which might not be the case when using normal splitting, since minority class is less likely to appear in both sets. A sample distribution illustrating stratified sampling of testing and training sets is depicted in Fig. 3.7. We can observe that class distribution is the same in both sets, demonstrating the effectiveness of stratified sampling with imbalanced dataset. Here, training testing split is 70:30.

3.4.2 Data Pre-Processing

As mentioned previously in Section 3.2.2, there are ten features associated with each alloy in the dataset. In order to avoid unneeded influence among different features, it is necessary to normalize every instance in each feature so that they can be treated equally within the training process. Thus, Min-Max normalization is carried out for the purpose of pre-processing. Min-Max normalization is defined as:

$$X_{N_i} = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \quad (3.13)$$

where X_{N_i} denotes i^{th} normalized data point of variable X ; $X_{\min}$ and $X_{\max}$ represent minimum and maximum value of X , respectively.

3.4.3 Hyperparameter Optimization

The CPU used in these experiments is Intel Core i7 - 11700 with 16GB RAM. The programming platform used here is Python 3.12.2 with Jupyter Notebook as IDE, for easy execution of algorithm and visualization of results. Pytorch framework is utilized for generating ANNs.

To ensure that the framework of ANN is functioning accordingly, we train a network with random hyperparameters and visualize its training behaviour, illustrated in Fig. 3.8. As expected, training loss improves rapidly in the initial phase of training and eventually reaches saturation and shows oscillations. At this stage, further epochs will not improve the capability of ANN, and might even contribute towards overfitting. This is precisely the reason number of epochs is included as a hyperparameter and its selection will be such that it maximizes testing accuracy.

For implementation of hyperparameter optimization via GA, we initially choose

population size as 20 and number of generations to be 5, with elitism equal to 2. The results of fitness evolution with respect to generations is depicted in Fig. 3.9. Fitness initially decreases rapidly, and seems to oscillate at the end, suggesting presence of local minima. Since population size and number of generations is relatively small, the algorithm is not sampling a large enough search space and is more likely to get stuck within a local minima.

To improve the capabilities of hyperparameter optimization, we run the experiment again with population size as 40, number of generations as 20 and elitism as 3. The results of fitness evaluation through each successive generation of GA with the updated parameters is illustrated in Fig. 3.10. Taking a look at the behaviour of fitness, it seems to not change beyond the second generation, however this is not the case, as demonstrated by a zoomed in view enclosed within the plot. This view illustrates that fitness is still improving, albeit at a slower rate since it is close to saturation. Here, we can observe that a larger population size tends to give a very large starting fitness function value, depicting diverse sampling of the search space which is crucial for obtaining good quality hyperparameters and will result in phase prediction model with high accuracy, as is the aim of this study.

To avoid contingency and ensure fair results, 10 experiments are run with population size = 40, number of generations = 20 and elitism = 4, where each run has a different random seed. The best accuracy obtained from these experiments is a fitness value of 1.1320 which equates to testing accuracy of 88.34% via $\text{fit} = 1/T_{acc}$. The resulting set of hyperparameters obtained from experiment with the best fitness is described in Table 3.3. A look at the selected feature subset will make it evident that from the set of ten features, only eight are selected by the algorithm. This makes it evident that the other two features are considered either redundant or their contribution to testing accuracy is minimal compared to the other features in the universal set. The approach of incorporating feature subset selection within the framework of hyperparameter optimization ensures that the right features are selected to maximize the performance of phase prediction and accurately represent the search space of MPEA phase dataset.

3.4.4 Analysis of Results

To analyze the prediction quality of our network we utilize confusion matrix, illustrated in Fig. 3.11. This matrix depicts correct and incorrect predictions of

Table 3.3: List of hyperparameters with the best testing accuracy

Hyperparameter	Range
Learning rate	1.3036×10^{-2}
Batch size	1.89×10^2
Number of epochs	6.09×10^2
Number of nodes	44
Dropout rate	9.5782×10^{-2}
k_{igr}	2.5989
Feature subset	$[\bar{Z}, \bar{r}, \text{VEC}, \Delta\chi, \delta, T_{melt}, \Delta S, \Delta H_{mix}]$

individuals in testing set. AM and IM classes have the best prediction quality with low false positives and false negatives. This is surprising since IM is the smallest minority class. Hence the current network with hyperparameters from Table 3.3 is suitable for minority classes. BCC+FCC class has the worst prediction quality, with 13% false positives and 50% false negatives, where the network is incorrectly labelling MP as BCC+FCC 35% of the time. This can be attributed to the fact that BCC+FCC can be considered as a mixed phase itself, specially when phases other than BCC and FCC are also present in the microstructure. However, we must be careful when considering removal of a class, since the aim is to generate an accurate depiction of phase search space, and oversimplification might lead to unphysical depiction of physical systems.

3.4.5 Validation of Prediction Model

To validate the phase prediction framework, we use the resulting ANN to predict phases of additional MPEA alloys reported in recent studies [129], [130], [131], [132], which have not been included in our dataset. The results of predictions are listed in Table 3.4. It is observed that our model succeeds in estimating all five validation MPEAs, making the approach developed in the current study as a viable strategy for phase prediction of MPEAs.

Table 3.4: Prediction results of new alloys for validation of model

Alloy	Actual phase	Predicted phase
$\text{Fe}_{4.5}\text{Co}_3\text{CrVMn}_{0.5}$ [129]	SS	SS
$\text{CoCu}_2\text{Mn}_3\text{Ni}_4$ [130]	SS	SS
$\text{Al}_{0.75}\text{Cu}_{3.0833}\text{Fe}_{3.0833}\text{Mn}_{3.0833}$ [131]	SS	SS
$\text{Cu}_{3.33}\text{Fe}_{3.33}\text{Mn}_{3.33}$ [131]	SS	SS
$\text{Co}_{1.75}\text{Cr}_{1.25}\text{Fe}_{5.5}\text{NiMo}_{0.3}\text{C}_{0.2}$ [132]	IM	IM

To further validate the approach chosen in this study, we compare the best re-

sults of our algorithm with some commonly utilized classification algorithms. 3.5 lists the outcomes of the aforementioned comparisons, containing name of the technique and the corresponding best testing accuracy for 10 independent runs. It is evident that our model of using ANN with hyperparameter optimization including outlier detection and feature subset selection as additional hyperparameters give the best prediction accuracy compared with five widely used multi-class classification techniques when applied to phase prediction of MPEAs. This comparison makes it apparent that conventional classification techniques are unfit to handle complex and high-dimensional MPEA data containing outliers and imbalance.

Table 3.5: Comparison among various classification algorithms

Technique	Best Accuracy
ANN with HPO and additional hyperparameters	0.84
Support Vector Machines	0.68
Decision Tree	0.75
Random Forest	0.71
XGBoost Classifier	0.79
Naive Bayes	0.59

3.4.6 Ablation Studies

In context of machine learning, ablation studies are defined as scientific examination of a machine learning system by removing its components in order to gain insight into its performance. These studies are useful in illustrating the importance of building blocks in an ANN model. For this purpose, we utilize ablation studies to signify the importance of novel additional parameters in our study, namely outlier detection and feature subset selection.

To perform ablation studies, we compare the performance of our model with and without key additional hyperparameters, as well as comparison between the two additional parameters. Each comparison comprises of 20 runs, and the results of mean testing accuracy are reported in Table 3.6. Here, O denotes outlier detection hyperparameter and F denotes feature subset selection hyperparameter.

From Table 3.6, we can assess the performance of our model based on comparison between presence and absence of O and F . Here, absence of O and F means all other five hyperparameters are still present. In the first column, we can observe that the model performs significantly better when only O hyperparameter is added to the set of hyperparameters. This supports our initial hypothesis of including outlier

Table 3.6: Comparison results of ablation studies

Setting 1	Accuracy	Setting 2	Accuracy
Without O and F	0.73	With only O	0.84
Without O and F	0.73	With only F	0.76
With only O	0.84	With only F	0.76
Without O and F	0.73	With both O and F	0.85

detection as a way to increase the quality of data in MPEA dataset. The second comparison is surprising, since it shows slight improvement when adding F to the list of hyperparameters. The third comparison between presence of only O and only F indicates that inclusion outlier detection is far superior in terms of testing accuracy as compared to feature subset selection. The lacklustre performance of feature subset selection necessitates further analysis on physical depiction of all features in the dataset. Addition of other chemical properties as features in future studies may lead to further insight into the importance of feature selection in phase prediction of MPEAs. The fourth and final comparison with and without both O and F further cements the success of our approach of using additional hyperparameters which lead to significant improvement in the mean testing accuracy. Combining results of ablation studies and validation using new alloys in 3.4, we can conclude that our model is suitable for predicting phases of alloys given their respective compositions, and will be helpful for discovering new MPEAs with desirable properties.

3.5 Conclusion

In this chapter, a hyperparameter optimization approach of ANNs is proposed for phase prediction of MPEAs. Specifically, we incorporate outlier detection and feature subset selection as novel additions to hyperparameter optimization, and utilize GA as the optimization algorithm. To make GA suitable for feature subset selection, a new strategy is developed to deal with variable length chromosome, where a single gene has variable length. Experimental results have shown viability of this approach as a phase prediction framework. Analysis of GA results depict successful optimization of both real-valued variables as well as dynamic sized categorical array. To validate the current approach, best obtained model is utilized to successfully predict phases of new MPEAs not present in the dataset. Ablation studies performed on our model indicate that outlier detection performs better than feature subset selection, and both in conjunction serve to increase the prediction accuracy significantly.

This will contribute towards high-throughput development of MPEAs via assisting experimental design through data science, as well as a deeper understanding of their physical characteristics.

Future work can be summarized as follows: 1) inclusion of further chemical and mechanical features beyond Hume-Rothery rules; 2) employing different selection, crossover and mutation strategies; 3) designing a similar strategy for supervised regression and applying it to predict mechanical properties of MPEAs; 4) creating a framework of multi-objective optimization, where ANN phase prediction is utilized as a search space to find alloys with desirable properties; 5) applying the proposed hyperparameter optimization to other material science as well as data analysis applications.

The use of GA for hyperparameter optimization lead to enhanced performance, however the presence of tunable parameters in GA hinder its capabilities. Crossover type, mutation rate, selection mechanism are a few tunable parameters that if chosen incorrrrectly, may lead to suboptimal performance irrespective of data quality. A possible solution could be to extensively run tests for a large set of values of tunable parameters and choose the best subset by gauging network performance, but this is a computationally expensive process with limited prospects of improvement. A more suitable solution could be to employ an alternative optimization technique with less or no tunable parameters. Chapter 4 will explore the use of a new population-based heuristic algorithm for the purpose of MPEA phase prediction.

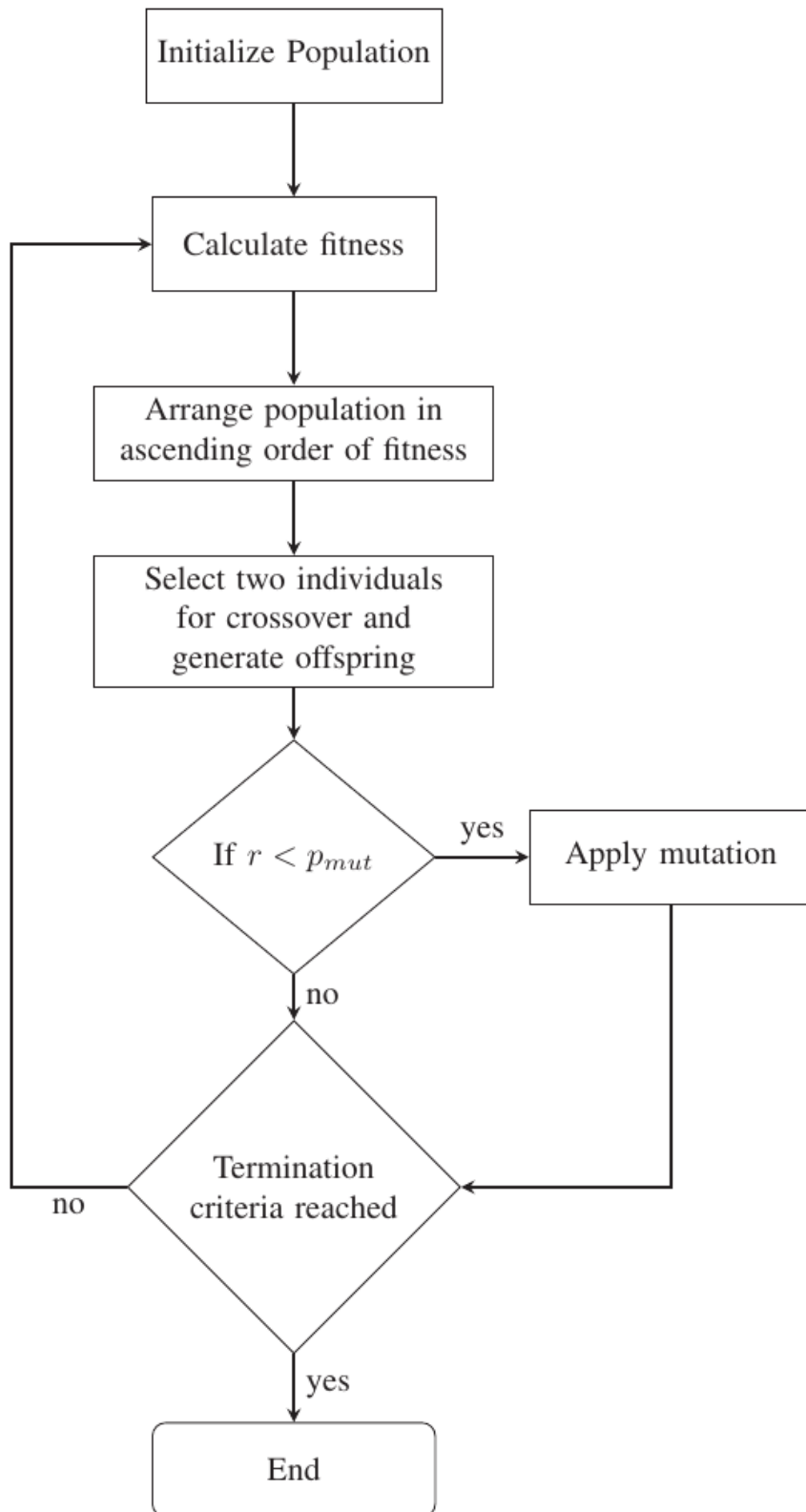


Figure 3.3: Flowchart of Genetic Algorithm. p_{mut} is mutation probability and r is a uniformly distributed random number between 0 and 1.

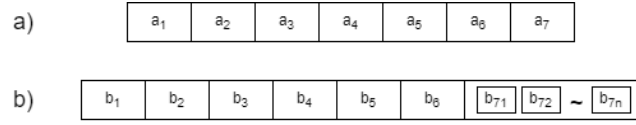


Figure 3.4: Depiction of chromosomes where a) shows normal chromosome with fixed length genes; b) shows chromosome having a single variable length gene which consists of values $b_{71}, b_{72} \dots b_{7n}$, where n is an integer which can vary between a lower and upper bound.

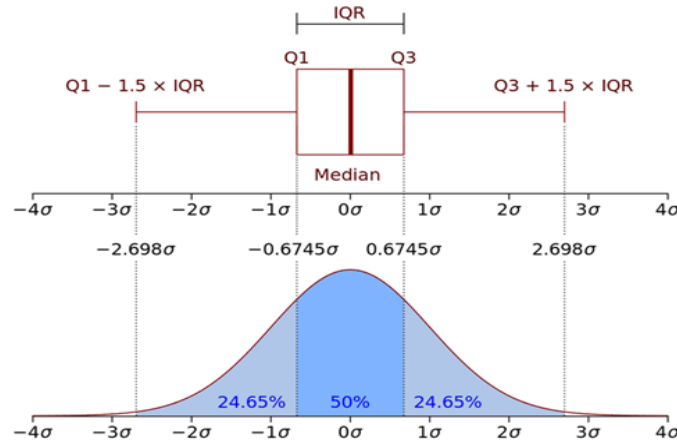


Figure 3.5: Depiction of boxplot with interquartile distance.

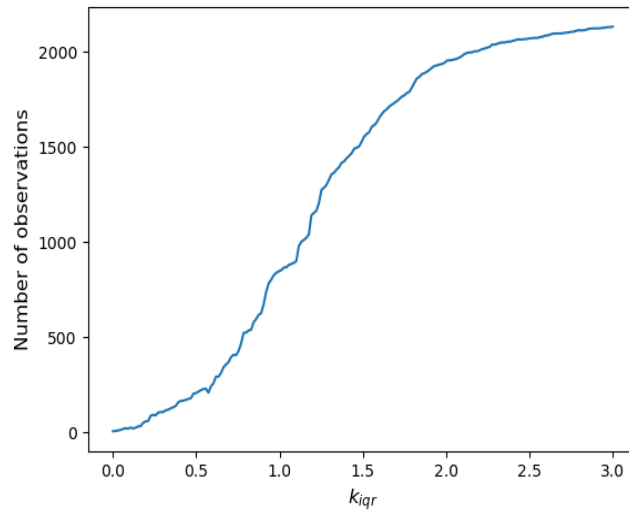


Figure 3.6: Variation of number of outliers with k_{iqr} .

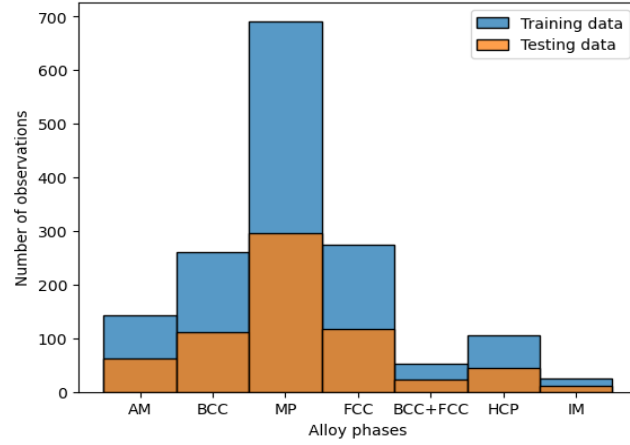


Figure 3.7: Distribution of classes within training and testing sets when using stratified sampling.

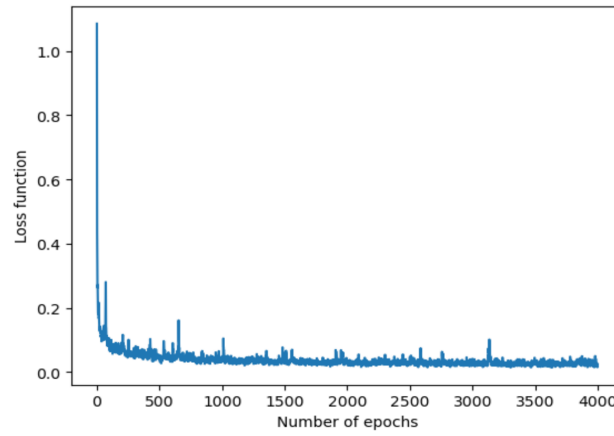


Figure 3.8: Training process of ANN, depicting evolution of loss function with number of epochs.

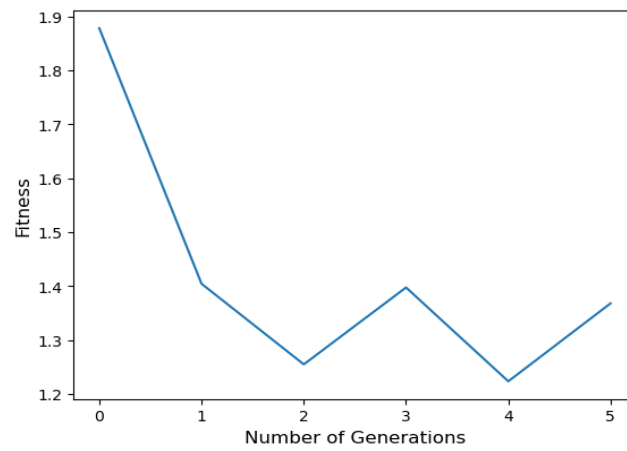


Figure 3.9: Fitness plot of GA with respect to number of generations. Here, Population size = 20, Number of Generations = 5 and Elitism = 2

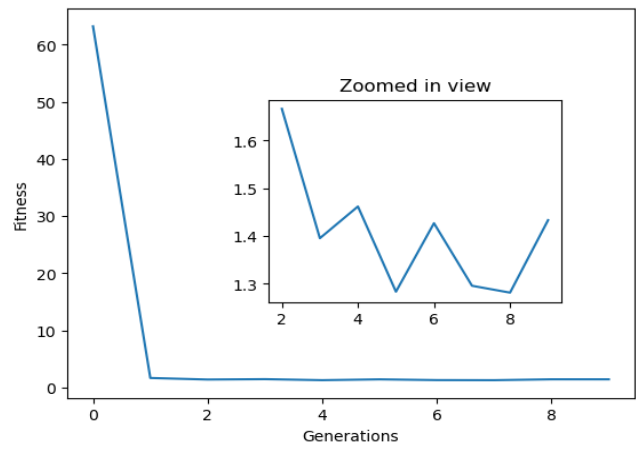


Figure 3.10: Fitness plot of GA with respect to number of generations, with a zoomed in view enclosed within the plot depicting change in fitness close to saturation. Here, Population size = 40, Number of Generations = 20 and Elitism = 3.

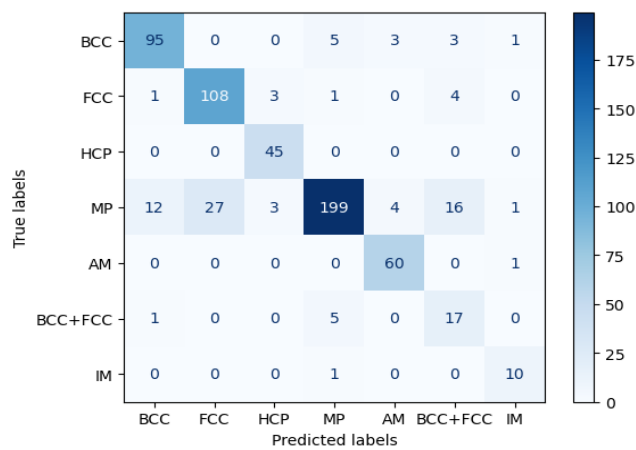


Figure 3.11: Confusion matrix of testing data

Chapter 4

Teaching Learning Based Optimization

4.1 Introduction

As discussed in Section 2.2, MPEAs have multiple principal elements and exhibit excellent physical and mechanical properties, making their study an important task for metallurgists and theoretical material scientists. Discovery and design of MPEAs is critical to their study, but it can not be accomplished by trial-and-error techniques as well as naive theoretical conjectures based on expert knowledge. Prediction framework of MPEAs has lead to the usage of data science techniques which are well suited for complex and high-dimensional systems.

As previously mentioned in Section 2.3, advances in algorithms and larger data availability has led to an increased interest in the fields of data science and machine learning applied to various aspects of material science, from material discovery and property analysis to quantum chemistry and development of accurate theoretical framework [133] [134] [135] [136]. Among these studies, phase prediction of MPEAs has been of critical importance due to the linkage between alloy phases and their physical properties.

The problem of phase prediction of MPEAs can be studied as a supervised multi-class classification problem, where alloy phases are represented as unique and mutually exclusive classes. Various studies have been published in this domain utilizing different approaches. These strategies involve use of random forests, support vector machines, feed-forward neural networks and generative adversarial networks. Among these techniques, artificial neural networks (ANNs) have received consider-

able interest due to their suitability for handling high-dimensional data.

Performance of ANNs hinges on selection of suitable hyperparameters, as discussed in Section 3.1. Building upon the motivation and findings discussed in Chapter 3, we aim to further improve upon the existing mechanism of GA by using a simpler and more flexible heuristic optimization technique which can overcome the disadvantages of GA as described in Section 2.4.2, and provide a better phase prediction mechanism of MPEAs.

Among the myriad of recently developed optimization techniques, Teaching Learning Based Optimization (TLBO) is a flexible and dynamic approach which has garnered interest due to the notable absence of tunable parameters and fast convergence. It is a population-based technique which operates similar to genetic algorithm and particle swarm optimization. However, in contrast to genetic algorithm which requires control parameters such as crossover rate and mutation rate for effective application, TLBO does not contain such parameters and hence is well-suited for engineering design applications and hyperparameter optimization.

The lack of tunable parameters is unfortunately not a sufficient advantage of TLBO over genetic algorithm which is used more frequently. The reason is due to imbalance between exploration and exploitation mechanisms in TLBO, leading to premature convergence. This problem is exacerbated in multi-modal and high-dimensional functions which have large number of local optima. Numerous modifications have been applied to the original TLBO to improve its performance in terms of knowledge sharing within population and reliable convergence criteria. This has led to application of TLBO in various domains such as electrical engineering [137][138], power engineering [75], internet of things [139] and material design [140] .

Motivated by the discussions above, the objective of this chapter is to develop a framework for multi-class classification via hyperparameter optimization, where the optimization algorithm utilized is TLBO with certain modifications to provide aid in exploration and slow down rate of convergence. The details of these modifications are presented in Section 4.3. This strategy is employed to boost the ability of ANNs when dealing with high-dimensional data. The proposed methodology is applied to phase prediction of MPEAs.

The remaining sections of this study are organized as follows. In Section 2, the background of MPEA classification is discussed, alongside a description of the datasets used in this study. Section 3 introduces the structure of TLBO and the proposed TLBO modification strategies, alongside an explanation of hyperparameter

optimization. In Section 4, experimental results for phase prediction of MPEAs are exhibited. Finally, Section 5 provides a conclusion of this study as well as suggestions for appropriate future work.

4.2 Background

4.2.1 Multi-Principal Element Alloys

Multi-principal element alloys (MPEAs) are defined as alloys comprising three or more principal elements, with each element present in 5-35% atomic concentration. Presence of larger concentrations of each element contribute toward unique properties. Experimental analysis of MPEAs have depicted these alloys as having excellent functional and mechanical properties. This behaviour can be attributed to the fact that MPEAs tend to exist in stable solid solution phases.

Since physical properties of MPEAs have been shown to depend upon atomic phases, phase prediction and identification have been key factors in their study. Given an atomic composition of elements, the ability to foretell alloy phase has been a focus of high-throughput design of MPEAs. Factors influencing formation of key phases can be inferred from Hume-Rothery rules which determine formation of solid solutions in binary alloys. Experimental studies have verified the theoretical conjectures of Hume-Rothery rules, making them a useful tool for phase identification prior to synthesis.

Advances in computational power and flexible algorithms have led to replacement of naive numerical models of alloys with machine learning-based prediction models which are data-driven and effective at extracting information from features. Numerous powerful techniques have been developed recently and have been used in material design and discovery. These are used as guiding tools for metallurgic analysis of MPEAs and aid in discovering applicable alloys.

4.2.2 Data Set Description

Data used in the current study has been gathered from multiple sources in literature. The combined dataset contains 2,218 alloy compositions, alongside their respective phase and elemental features. Ten chemical features are attributed to each alloy, listed in eqs. (3.1) to (3.9). Values of binary mixing enthalpies are taken from the

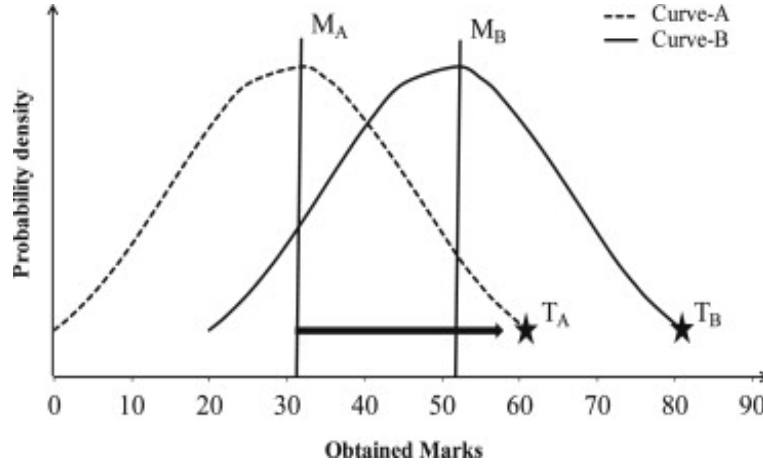


Figure 4.1: Distribution of marks obtained for a group of learners, depicting the formulation of TLBO [73]

Miedema model [141]. Number of principal elements in an alloy is also considered as a feature, since entropy of MPEAs depend upon number of principal components.

Each alloy is categorized into one of seven classes: body-centered cubic (BCC), face-centered cubic (FCC), hexagonal closed pack (HCP), intermetallic (IM), amorphous (AM), multi-phase (MP) and BCC+FCC. These classes, also known as atomic phases, dictate the arrangement of atoms within the lattice of an alloy and are responsible for distinct physical attributes such as ductility, strength and hardness.

4.3 TLBO Algorithms

4.3.1 Basic TLBO Algorithm

The TLBO algorithm developed by Rao et. al. [142] is a population-based evolutionary optimization technique where each discrete member of the population represents a possible solution set, similar to other well-known population-based algorithms such as particle swarm optimization [143] and genetic algorithm [144]. A detailed description of the algorithm has been presented in Section 2.5.

A qualitative depiction of TLBO can be thought of as a teacher (best student) improving the rest of the class (population) by moving the mean grades (fitness) towards itself. This description is illustrated in Fig. 4.1, where T_A and T_B depict teachers of grade distributions A and B , with means M_A and M_B . The teacher T_A will move students towards the right, thereby changing the distribution with an improved mean and a better teacher after every iteration of learning.

TLBO consists of two distinct phases: teaching phase and learner phase. The

solution update mechanism in teaching phase is described in eq. (2.3), and update mechanism in learner phase is mentioned in eq. (2.5) in Chapter 2.

4.3.2 Popular TLBO Variants

Since 2011, TLBO has been widely extended and modified to improve performance, specifically on complex and high-dimensional optimization problems. Popular variants of TLBO in literature are as follows.

Improved and Self-Adaptive TLBO Variants

These variants modify the conventional TLBO mechanism e.g. adding new phases, adaptive behaviour or improved search operators. Improved TLBO (ITLBO) is a variant which introduced self-feedback phase to balance exploration and exploitation by utilizing historical knowledge of population[145]. Autonomous TLBO (ATLBO) added a self-learning phase based on fuzzy rules which enhanced learner updates compared to the original update mechanism of teacher and learner phases[145]. Balanced TLBO incorporated restart and weighted averages to improve convergence with excellent results[146]. Bare-bones TLBO used a hybridization approach by using Gaussian sampling with neighbourhood search to improve both teacher and learner phase updates[147].

Elitist and Multi-teacher TLBO Variants

Elitist TLBO function in the same manner as elitism in GA, by preserving high quality solutions across generations. By ensuring that the best individuals are retained, elitist TLBO improves convergence speed and solution accuracy. However, excessive use of elitism may lead to reduced population diversity and hence must be carefully monitored [80].

Multi-teacher TLBO improve upon the original framework by extending the classroom analogy by assigning different teachers to subgroups of learners, each emphasizing different knowledge components. This strategy improves parallel exploration of the search space and has been shown to improve performance in multimodal and multi-objective optimization problems [148].

Hybrid TLBO Variants

These modified variants combine the approach of TLBO with other metaheuristic algorithms to improve their performance. Harmony search was combined with TLBO using self-adaptive selection strategy with the aim to balance global and local search. Harmony search focused mainly on exploration and TLBO on exploitation, creating a synergistic approach which combines the strength of both algorithms[149]. Grey wolf optimizer was integrated with TLBO where Grey wolf acted similar to harmony search with its focus on exploration and TLBO was used to fine-tune the search close to a local minimum[150].

Binary and Discrete TLBO Variants

The original TLBO formulation is designed for continuous search spaces. To extend its applicability, binary and discrete TLBO variants have been proposed in literature, using transfer functions or problem-specific encoding schemes. These adaptations enable TLBO to address feature selection, scheduling, and combinatorial optimization problems. However, the effectiveness of such variants depends strongly on the choice of discretization strategy [151].

Chaotic TLBO Variants

More recent variants incorporate chaos theory to enhance population diversity and escape local optima. Chaotic TLBO replaces random number generators with chaotic maps to improve generalizability, while also accelerating convergence. These approaches have shown promise in highly nonlinear and multimodal optimization landscapes [152].

4.4 Novel TLBO Algorithm

Taking inspiration from modifications done for PSO and DE, we have formulated three novel strategies for tackling the problem of fast convergence and lack of exploration-exploitation balance in canonical TLBO. These modifications are mentioned below.

4.4.1 Incorporation of Previous Knowledge in TLBO

The use of historical knowledge of solutions to modify and improve current solutions is a tactic employed for similar population-based optimization algorithms such as PSO, GA and DE in literature. Previous position of an individual in a population can help guide the direction of its future modification and help avoid getting stuck in a local minima. This mode of self feedback is essential in creating a balance between exploration and exploitation.

To this effect, we design a self feedback loop as follows.

$$X_i^{k+1} = X_i^k + r_i(X_i^k - X_i^{k-1}) \quad \text{if} \quad f(X_i^k) < f(X_i^{k-1}) \quad (4.1)$$

where i represents a particular individual within the population and k denotes the current number of generation. This self-feedback improvement will occur only when the previous solution is better than the current solution, ensuring improvement in the subsequent generation. The fitness function chosen in the current study is $f(X) = 1/T_{acc}$, where T_{acc} is testing accuracy of MPEA phase prediction model dependent on X hyperparameters.

To further improve canonical TLBO, a weighted fitness-based mean is introduced within the teaching phase to better reflect the characteristics of the population with respect to fitness of each individual solution. This is denoted mathematically as follow.

$$M_{weighted} = \frac{\sum_{i=1}^N \frac{X_i}{fit(X_i)}}{\sum_{i=1}^N \frac{1}{fit(X_i)}} \quad (4.2)$$

Self-adaptive strategies have been utilized in TLBO in prior research, but a combination of self-feedback and weighted mean is a novel approach with the aim to incorporate historical knowledge of previous generations to guide the direction of search, and the weighted mean serves as a better representation of population with good learners given precedence in the teacher phase. Both strategies serve to direct each individual in a more promising area of the search space while avoiding local minima, improving upon the original formulation of TLBO. The modified algorithm is depicted as a flowchart in Fig. 4.2, where the current modified version is termed as Variant 1.

here, $M_{weighted}$ is the weighted mean, X_i and $fit(X_i)$ are values of learner and fitness of learner respectively, and N is the total number of individuals in a population.

The combined approach of weighted mean with self-feedback learning can aid in exploration while reducing the chances of premature convergence, the critical ailment of conventional TLBO.

4.4.2 Noisy TLBO with Weighted Teaching

Inspired by approaches adopted in research for PSO and DE, this modification incorporates a method of weighted teaching by defining a weight as follows.

$$w_i = w_{max} - \left(\frac{w_{max} - w_{min}}{i_{max}} \right) i \quad (4.3)$$

where w_i is the calculated weight within the interval $[w_{min}, w_{max}]$, i is the current iteration and i_{max} is the maximum number of iterations. The modified teaching phase with weighted learning is depicted as follows.

$$X_i^{k+1} = w_i X_i^k + r_i (T_i^k - T_F M^k) \quad (4.4)$$

here, i denotes a particular solution from the population and k depicts the current iteration. The addition of linearly decreasing weight with respect to number of iterations serves as a crucial balance between exploration and exploitation. The learner phase is similarly modified below.

$$X_i^{k+1} = \begin{cases} w_i X_i^k + r_i (X_i^k - X_j^k), & \text{if } f(X_i^k) < f(X_j^k) \\ w_i X_i^k + r_i (X_j^k - X_i^k), & \text{if } f(X_i^k) > f(X_j^k) \end{cases} \quad (4.5)$$

The addition of weights to the update process of TLBO controls how strongly solutions are influenced. The objective is to control step size during updates to balance exploration and exploitation. This benefit comes at the cost of introducing additional parameters that need to be adapted carefully, undermining the simplicity of original TLBO which stems from its parameter-free design. The current modified version of TLBO is termed as Variant 2, and its flowchart is illustrated in Fig. 4.3.

In addition to weighted teaching and learning, white noise is added as a final step after teaching and learning phases. This noise is the addition of normally distributed random number $N(0, 1)$ with zero mean and unit variance. Combined with weighted teaching, this approach will slow down convergence and focus on careful exploitation instead of fast exploration, as opposed to the characteristic of canonical TLBO which favours exploration over exploitation.

Similar improvements have been implemented for inertial weights in PSO, but the modification of adding white noise to a weighted teaching scheme is a novel addition to TLBO, serving to improve its performance by focusing on exploration phase in the beginning and gradually changing focus to local search at later stages of the algorithm. This process is aided by a linearly increasing weight, while white Gaussian noise functions similarly to mutation operator in GA by allowing random jumps in search, increasing the probability of escaping unfavourable regions.

4.4.3 TLBO with Gaussian Sampling and Increased Knowledge Sharing

Developed by taking ideas applied to PSO and DE, Gaussian sampling is introduced heuristically by incorporating normal distribution in the teaching and learning phases. This is motivated by the statistical properties of normal distribution, which concentrates probability around the mean while still allowing occasional large deviations. This characteristic is particularly well suited for optimization where most search steps should refine the current solution, while allowing sporadic large jumps to escape local minima. The process of Gaussian sampling in the teacher phase is described as follows.

$$X_i^{k+1} = \begin{cases} N(\frac{X_i^k + T_i^k}{2}, |T_i^k - X_i^k|), & \text{if } r_{new} > 0.5 \\ X_i^k + r_i(T_i^k - T_F M_i^k), & \text{otherwise} \end{cases} \quad (4.6)$$

here, X_i^k represents a particular learner i at iteration k , T_i^k depicts the best learner in a population and $N(\mu, \sigma^2)$ represents normal distribution with μ mean and σ^2 standard deviation, with r_i and r_{new} representing uniformly distributed random number between 0 and 1. T_F and M_i^k depict teaching factor and mean of the population respectively.

The teaching phase in this modification has the ability to switch freely between gaussian distribution and uniform distribution of original TLBO. This can lead to improved heuristic search as well as adaptability. The switching probability of 0.5 can also be tuned depending upon the problem, as well as the stage of optimization, as the algorithm may want to explore more in the early stages and refine its search locally in later stages of optimization. We decided to keep the value of switching probability as 0.5 rather than modify it, to preserve the original benefit of TLBO

which is lack of tunable parameters. Adding a hyperparameter of switching probability may improve the algorithm but only if it is chosen carefully or else it may hinder the search process significantly.

Similar modification of learner phase is employed, with an added modification of the use of four learners instead of two in the original TLBO. This modification serves to incorporate knowledge of multiple students simultaneously when updating the current learner. In a classroom setting, multiple students have the capability to learn from each other and improve together, and this modification aims to embody this process. Hence, learner phase can be described as.

$$X_i^{k+1} = \begin{cases} N(\frac{X_i^k + X_j^k}{2}, |X_i^k - X_j^k|), & \text{if } r_{new} > 0.5 \\ X_i^k + r_1(X_j^k - X_i^k) + r_2(X_m^k - X_n^k), & \text{otherwise} \end{cases} \quad (4.7)$$

where X_i^k is the chosen learner and X_j^k , X_l^k , X_m^k and X_n^k are four distinct solutions from the current population which serve to pool their knowledge in order to improve the solution of X_i^k . Here, r_1 and r_2 are uniformly distributed random numbers. Similar to the teacher phase, learner phase has the ability to switch between normal mode of update and pooling mode with equal probability. This version of TLBO is termed as Variant 3, and the algorithms is depicted as a flowchart in Fig. 4.4.

Gaussian sampling has been implemented in TLBO before, but the addition of knowleedge sharing with multiple learners makes this a novel modification to original TLBO. Gaussian sampling shifts a small amount of local optimum to explore new solutions, preventing conventional TLBO from sticking in a local optima. Increased knowledge sharing delivers improvement from multiple sources as opposed to one, contributing to an overall improvement in learner phase. This novel modification scheme takes TLBO closer to PSO, where knowledge sharing among multiple individuals acts in a similar way as neighbourhood information in swarms of PSO.

4.5 Experimental Results

4.5.1 Benchmark Results

To justify the choice of using TLBO over GA, we test them on several well known benchmark functions. Table 4.1 depicts the definitions of several benchmark functions used in this study, and Table 4.2 lists a detailed mathematical depiction of each

function. We have considered a combination of unimodal and multimodal benchmark functions for the purpose of testing TLBO on a variety of different functions to gauge its performance, specifically compared with GA.

Table 4.1: List of benchmark functions considered in this study

Function number	Function name	Search range
$F_1(x)$	Sphere Function	$[-100, 100]$
$F_2(x)$	Rosenbrock Function	$[-30, 30]$
$F_3(x)$	Step Function	$[-100, 100]$
$F_4(x)$	Bent Cigar Function	$[-10, 10]$
$F_5(x)$	Rastrigin Function	$[-5.12, 5.12]$
$F_6(x)$	Zakharov Function	$[-5, 5]$
$F_7(x)$	Levy Function	$[-10, 10]$
$F_8(x)$	Ackley Function	$[-32, 32]$
$F_9(x)$	Griewank Function	$[-100, 100]$

Table 4.2: Mathematical description of benchmark functions

Function number	Function name	Mathematical formula
$F_1(x)$	Sphere Function	$F_1(x) = \sum_{i=1}^D x_i^2$
$F_2(x)$	Rosenbrock Function	$F_2(x) = \sum_{i=1}^{D-1} (100(x_{i+1} - x_i)^2 + (x_i - 1)^2)$
$F_3(x)$	Step Function	$F_3(x) = \sum_{i=1}^D (\lfloor x_i + 0.5 \rfloor)^2$
$F_4(x)$	Bent Cigar Function	$F_4(x) = x_1^2 + 10^6 \sum_{i=2}^D x_i^2$
$F_5(x)$	Rastrigin Function	$F_5(x) = \sum_{i=1}^D (x_i^2 - 10 \cos 2\pi x_i + 10)$
$F_6(x)$	Zakharov Function	$F_6(x) = \sum_{i=1}^D x_i^2 + \left(\sum_{i=1}^D 0.5 i x_i \right)^2 + \left(\sum_{i=1}^D 0.5 i x_i \right)^4$
$F_7(x)$	Levy Function	$F_7(x) = \sin^2 \pi w_1 + \sum_{i=1}^{D-1} (w_i - 1)^2 \left(1 + 10 \sin^2 (\pi w_i + 1) \right) + (w_D - 1)^2 \left(1 + \sin^2 2\pi w_D \right), \quad w_i = 1 + \frac{x_i - 1}{4}$
$F_8(x)$	Ackley Function	$F_8(x) = -20 \exp \left(-0.2 \sqrt{\frac{1}{D} \sum_{i=1}^D x_i^2} \right) - \exp \left(\frac{1}{D} \sum_{i=1}^D \cos 2\pi x_i \right) + 20 + e$
$F_9(x)$	Griewank Function	$F_9(x) = \sum_{i=1}^D \frac{x_i^2}{4000} - \prod_{i=1}^D \cos \frac{x_i}{\sqrt{i}} + 1$

The performance of both GA and original TLBO are compared on the nine benchmark functions in Figs. 4.5 to 4.13. To ensure a fair comparison, multiple runs are conducted for each function, with the comparison parameters given in Table 4.3. For every benchmark, TLBO heavily outperforms GA due to its faster convergence to the global solution, irrespective of unimodal or multimodal functions. For $F_3(x)$, $F_5(x)$ and $F_9(x)$, TLBO finds the global minima before GA even converges to the solution. For $F_1(x)$, $F_2(x)$, $F_4(x)$, $F_6(x)$, $F_7(x)$ and $F_8(x)$, TLBO converges towards the solution much faster than GA. The fast convergence rate of TLBO is the principal cause of its superior performance since the crossover and mutation

operations in GA serve to balance exploration and exploitation but also lead to slower rate of convergence.

Table 4.3: List of comparison parameters for benchmark functions

Parameter Name	Parameter Value
Number of runs	30
Number of dimensions	10
Maximum iterations	500
Population size	50

Faster convergence may seem like an advantageous aspect of TLBO, but the focus of this study is to create a prediction model for phases of MPEAs, a highly non-linear and complex search space. To accurately gauge the performance of TLBO utilized in hyperparameter optimization for phase prediction model of MPEAs, we compare its accuracy with GA by performing 20 independent runs with population size of 30 for both GA and TLBO. The results of testing accuracy comparison are depicted in Table 4.4.

Table 4.4: Statistical results of testing accuracy comparison between GA and TLBO when applied to MPEA data.

Statistical quantity	GA	TLBO
Best value (Maximum)	85.71	72.34
Mean	79.05	61.33
Minimum	74.11	53.59
Range	11.60	18.75

We can derive the conclusion from the results obtained by comparison of GA and TLBO on benchmark and MPEA data is that TLBO outperforms GA on benchmarks, but severely underperforms on alloy data, due to poor convergence and lack of search diversity. This begs the need for improvements in canonical TLBO such that we can reap the benefits of having no tunable parameters while also obtaining a good prediction model.

4.5.2 Hyperparameter Optimization

The CPU used in these experiments is Intel Core i7 - 11700 with 16GB RAM. The programming platform used here is Python 3.12.2 with Jupyter Notebook as IDE,

for easy execution of algorithm and visualization of results. Pytorch framework is utilized for generating ANNs.

For implementation of hyperparameter optimization via modified TLBOs, we choose a population size of 50 and number of generations as 20. After 10 independent runs, statistical results of comparison between original TLBO and the three modified TLBOs are presented in Table 4.5

Table 4.5: Statistical results of testing accuracy comparison between original TLBO and its modified variants when applied to hyperparameter optimization of phase prediction for MPEA data.

Statistical quantity	Original TLBO	Variant 1	Variant 2	Variant 3
Best value (Maximum)	0.59	0.79	0.71	0.64
Mean	0.53	0.72	0.66	0.60
Minimum	0.51	0.68	0.64	0.62

The results depicted in Table 4.5 clearly illustrate that all modified TLBO variants performed better than the original TLBO, and Variant 1 gives the best testing accuracy out of the three variants. However, if we compare the result obtained from the best variant (0.79) with the best result obtained from GA (0.85), GA still outperforms it. Here we explore a novel idea of testing all the combinations of the TLBO variants and comparing their prediction accuracy. 5 independent runs are employed and the results of the best accuracy obtained in those 5 runs are depicted in Table 4.6. The combination of Variant 1 and Variant 2 creates the best model with high testing accuracy. Comparing it to the best accuracy obtained from GA which is 0.852831, we can surmise that the combined approach of Variant 1 and 2 outperforms GA by a narrow margin, while keeping the benefits of TLBO intact, which are the lack of tunable parameters compared with GA. We have chosen to call the best modified TLBO model as enhanced TLBO.

Table 4.6: Statistical results of testing accuracy comparison between original TLBO and its modified versions when applied to hyperparameter optimization of phase prediction for MPEA data.

Presence of Variant 1	Presence of Variant 2	Presence of Variant 3	Accuracy
Yes	Yes	No	0.853941
Yes	No	Yes	0.736504
No	Yes	Yes	0.790513
Yes	Yes	Yes	0.804011

4.5.3 Analysis of Results

The fitness characteristic of enhanced TLBO can be illustrated by visualizing the variations of fitness with increase in iterations, as shown in Fig. 4.14. As opposed to fast convergence, the fitness plot of enhanced TLBO exhibits slow and careful convergence, avoiding local minima with every few generations and improving its testing accuracy. This phase prediction model is shown to be robust and feasible, and can be utilized for high-throughput MPEA design.

To analyze the prediction quality of the enhanced TLBO, we utilize confusion matrix, depicted in Fig. 4.15. This matrix depicts correct and incorrect predictions of individuals in the testing set. IM and AM classes have the best prediction quality with low number of false positives and false negatives. This is a surprising statistic, since both IM and AM have low representation in the dataset, due to data imbalance. Due to the heuristic nature of enhanced TLBO, it may be difficult to clearly state the reason for superior performance, other than to say that the balance between exploration and exploitation offered by enhanced TLBO is leading to good representation of minority classes in the function space created by the optimized ANN. This shows that we can utilize this approach confidently for minority classes. Similar to the results of GA, the BCC+FCC class has the worst prediction accuracy, where the network is falsely labeling it as MP 34% of the time. From domain knowledge, we can surmise that BCC+FCC is a redundant phase, since it is a combination of two individual phases, whereas MP is also a combination of multiple phases. However, we can not hastily remove BCC+FCC phase from our study, since it is a microstructure of interest to the field of experimental metallurgy, and this study aims to incorporate domain knowledge to the statistical power of data-driven modelling. The prediction quality of each phase indicates the feasibility of this approach, as well as indicating room for further improvement.

4.6 Conclusion

In this chapter, a hyperparameter approach of ANNs is proposed for phase prediction of MPEAs. For this purpose, we utilize TLBO as a heuristic optimization algorithm without tunable parameters. Results of comparison tests on benchmark functions between GA and TLBO indicate that TLBO is significantly better, but does not outperform GA when applied to MPEA data. Original TLBO suffers from

premature convergence, therefore we propose three novel modifications to promote balanced search. Experimental results have shown viability of this approach as a phase prediction framework. Analysis of the three novel variants of TLBO depict successful improvement in optimization of hyperparameters of ANN compared with original TLBO. A novel approach of testing all possible permutations of modifications have indicated that variant 1 and 2 combined result in the best testing accuracy. We term this combination as enhanced TLBO and compare its performance to GA to further validate the current approach. Results of comparison indicate improvement in performance utilizing enhanced TLBO, while overcoming the limitations of original TLBO as well as keeping the benefits intact. This will contribute towards high-throughput development of MPEAs by assisting experimental design via data-driven approach, as well as contributing to theoretical framework of complex multi-dimensional alloy systems and their physical characteristics.

The results of GA and TLBO in Chapter 3 and current chapter indicate promising prospectives of using intelligent data analysis techniques for MPEAs. However, no special attention has been given to the process of feature design, which is imperative for a faithful depiction of MPEA composition space. Features chosen thus far emanate from Hume-Rothery rules, which provide rudimentary guidelines for phase formation. Since they were originally designed for binary systems, extension to multi-component systems prove to be cumbersome and inaccurate. This necessitates the use of an alternate feature design approach which can accurately depict the chemical and physical nature of each principal element in an MPEA, which in turn be used as inputs for training a phase prediction model. For this purpose, Chapter 5 explores a novel combination of a new alloy representation combined with hyperparameter optimization.

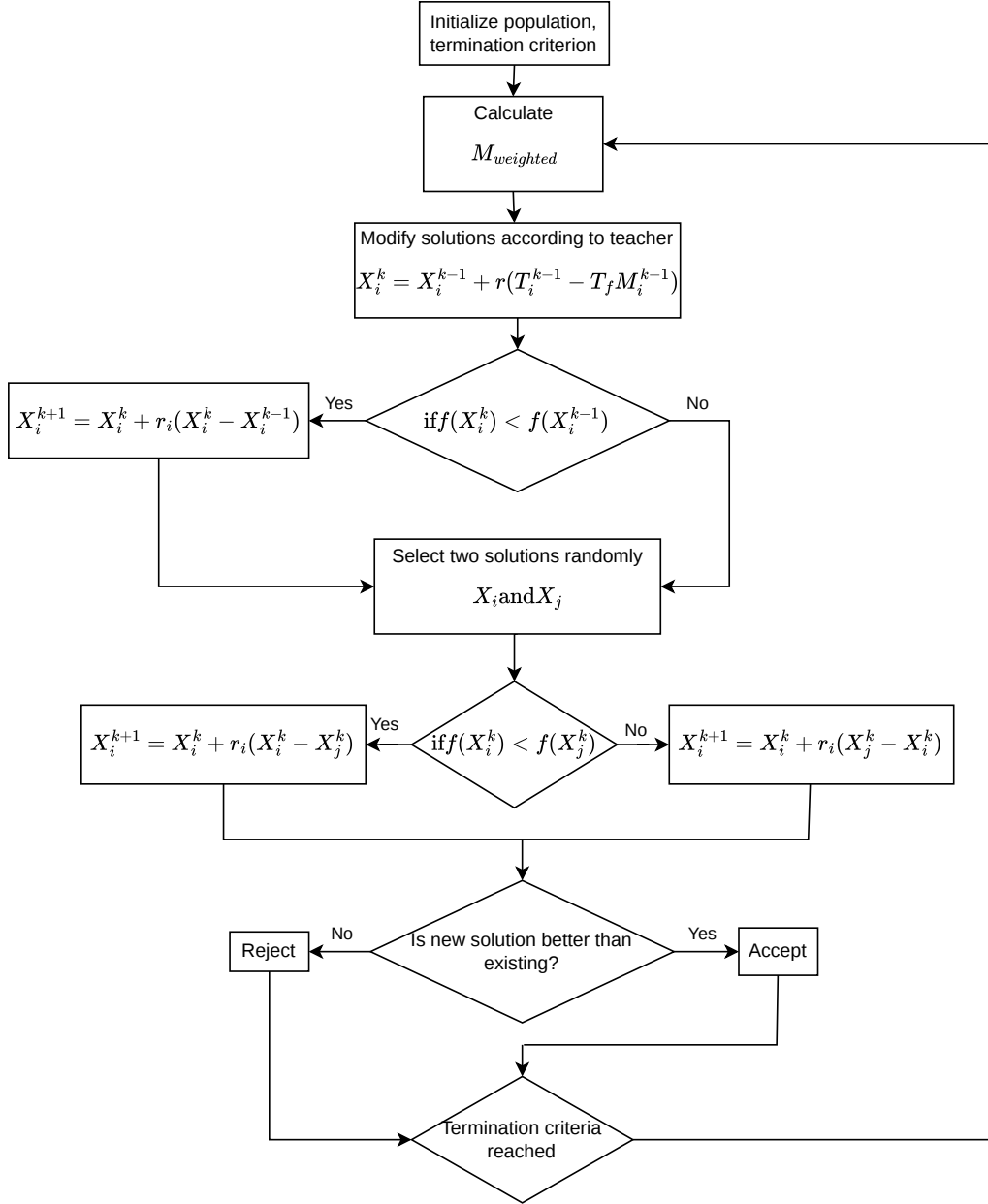


Figure 4.2: Flowchart of TLBO Variant 1, depicting teaching and learner phase.

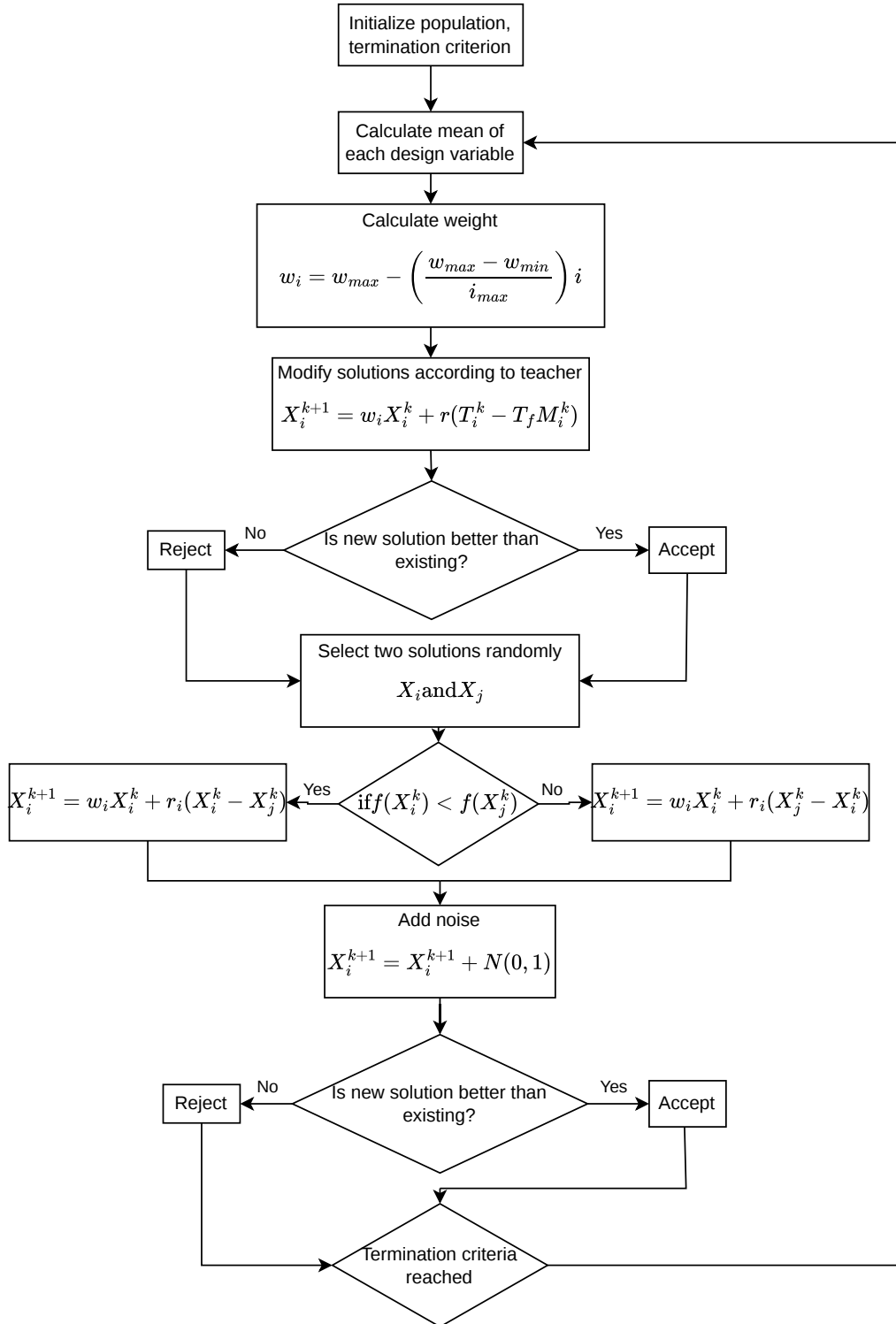


Figure 4.3: Flowchart of TLBO Variant 2, depicting teaching and learner phase.

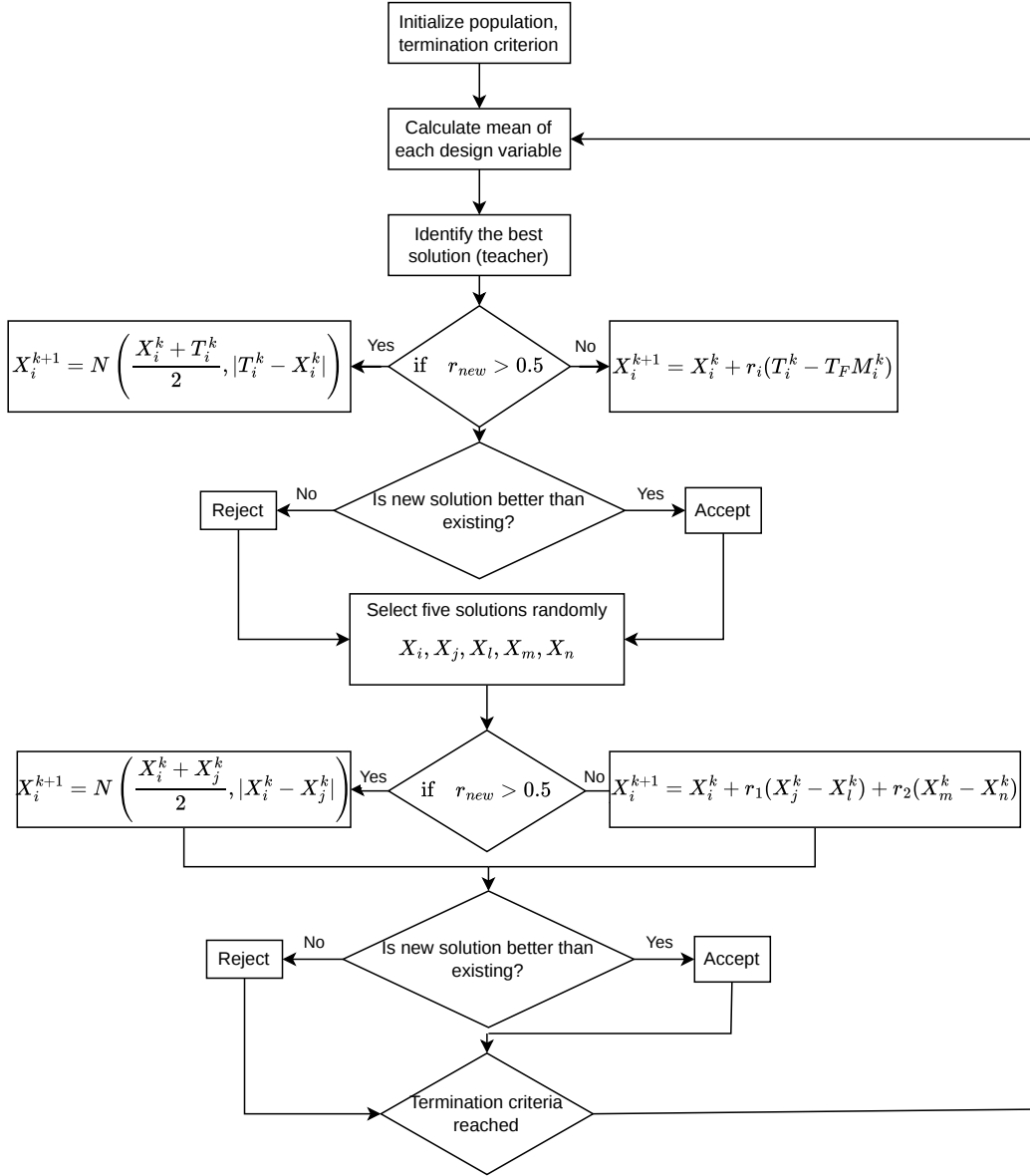


Figure 4.4: Flowchart of TLBO Variant 3, depicting teaching and learner phase.

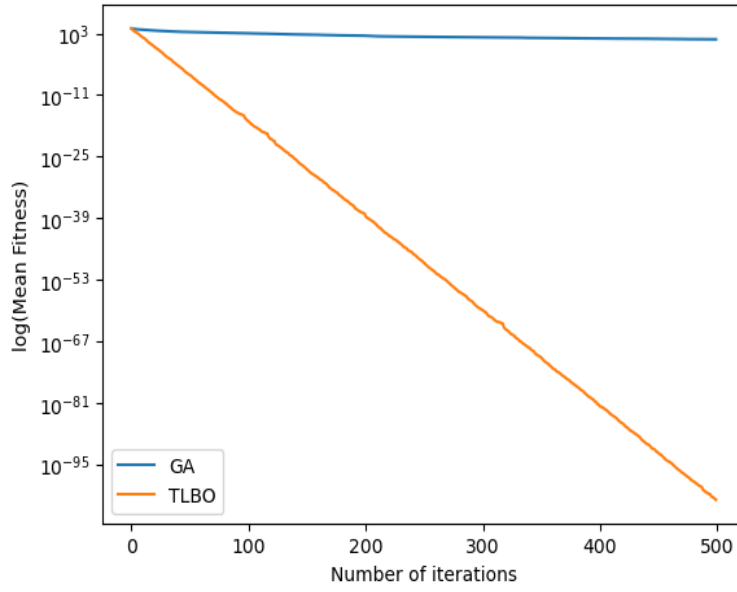


Figure 4.5: Comparison between optimization results of GA and TLBO for $F_1(x)$

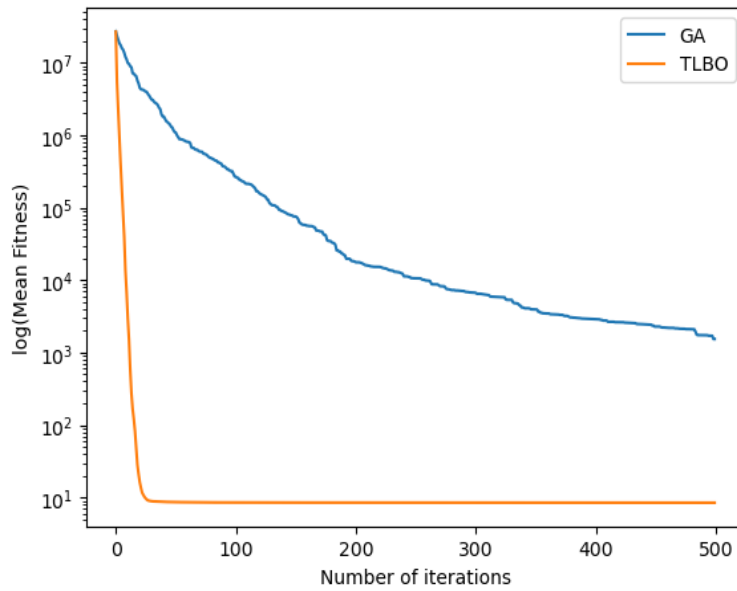


Figure 4.6: Comparison between optimization results of GA and TLBO for $F_2(x)$

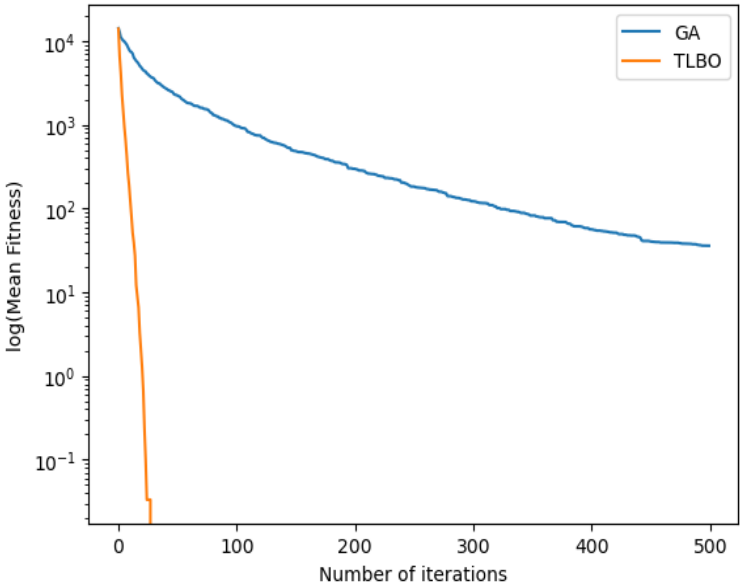


Figure 4.7: Comparison between optimization results of GA and TLBO for $F_3(x)$

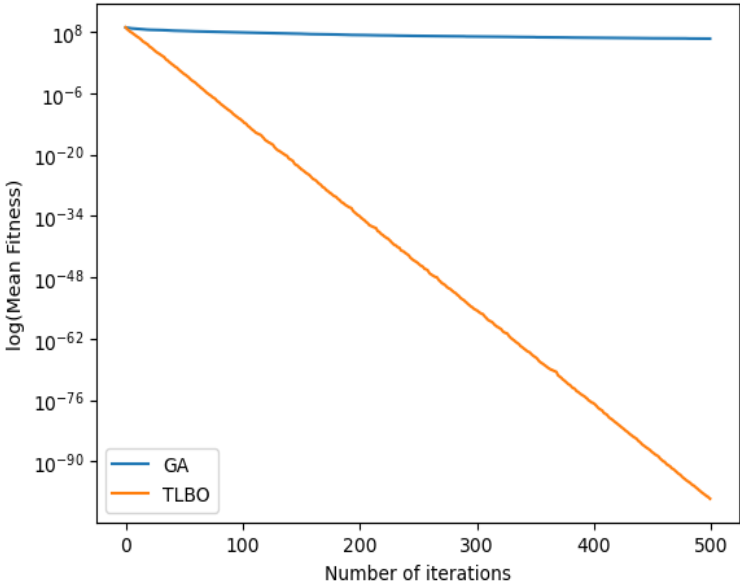


Figure 4.8: Comparison between optimization results of GA and TLBO for $F_4(x)$

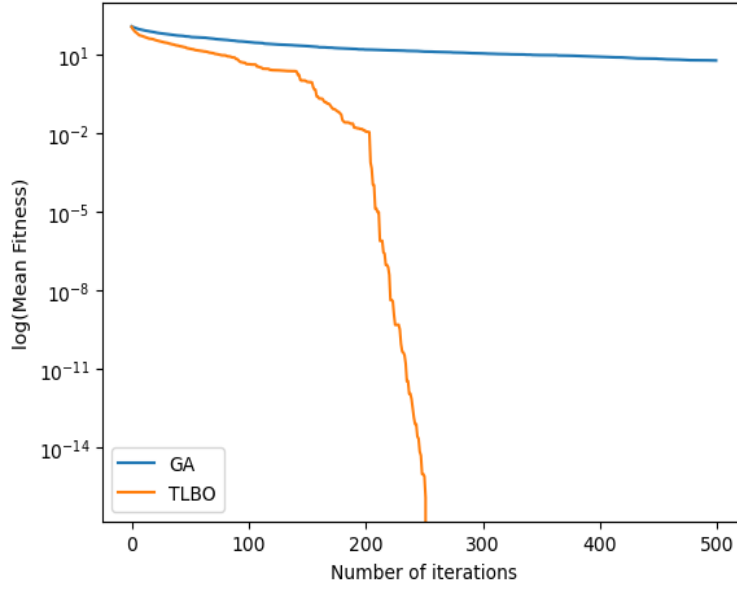


Figure 4.9: Comparison between optimization results of GA and TLBO for $F_5(x)$

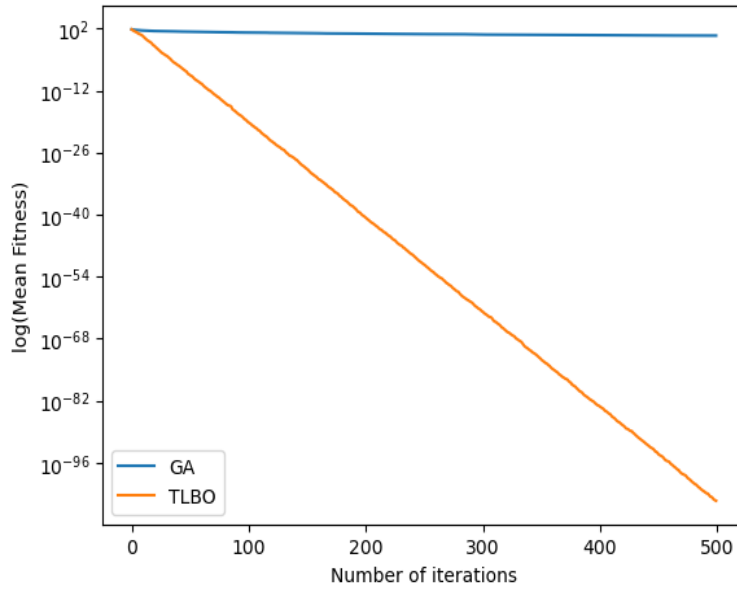


Figure 4.10: Comparison between optimization results of GA and TLBO for $F_6(x)$

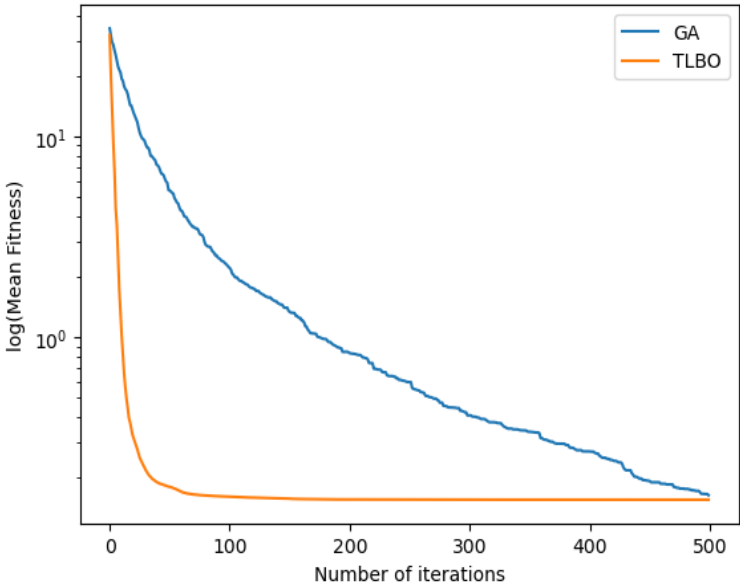


Figure 4.11: Comparison between optimization results of GA and TLBO for $F_7(x)$

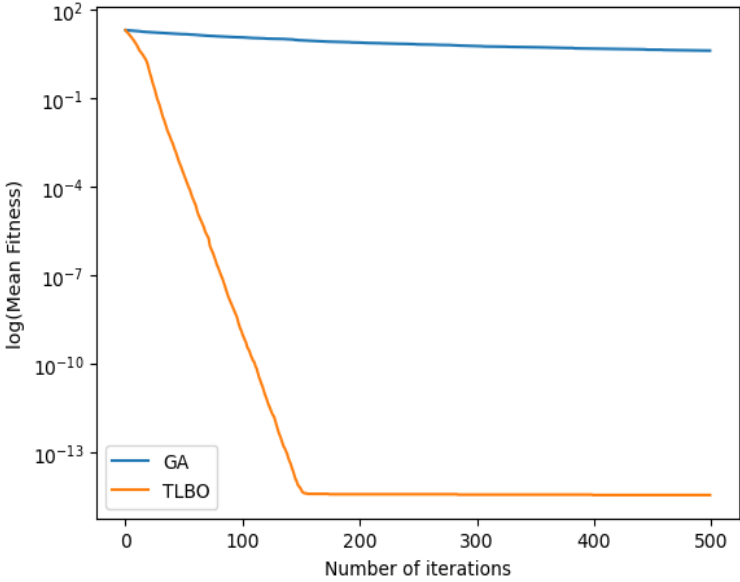


Figure 4.12: Comparison between optimization results of GA and TLBO for $F_8(x)$

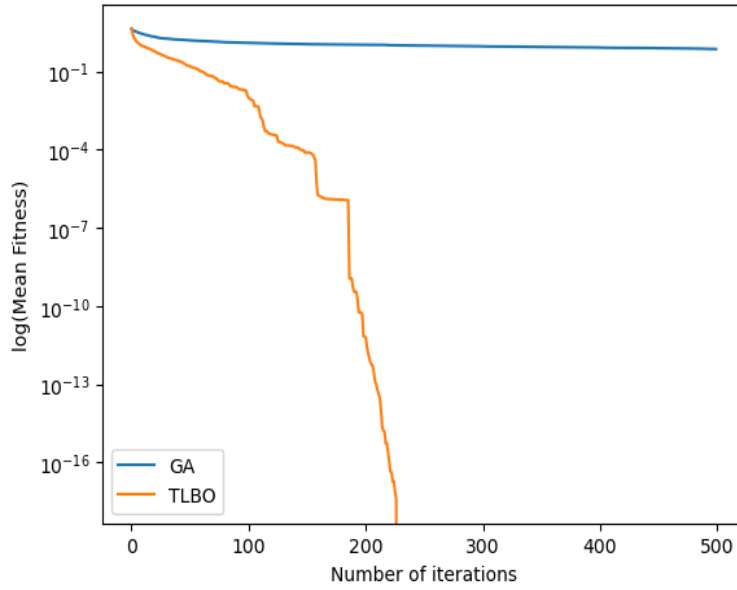


Figure 4.13: Comparison between optimization results of GA and TLBO for $F_9(x)$

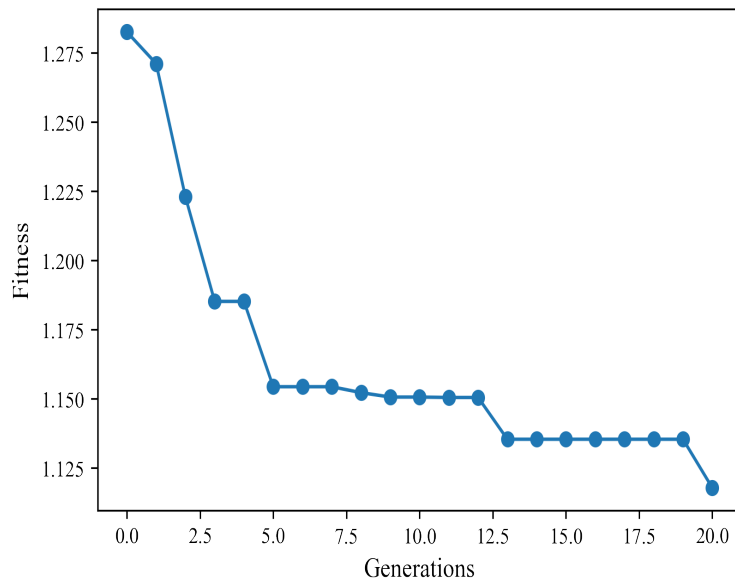


Figure 4.14: Fitness plot of enhanced TLBO

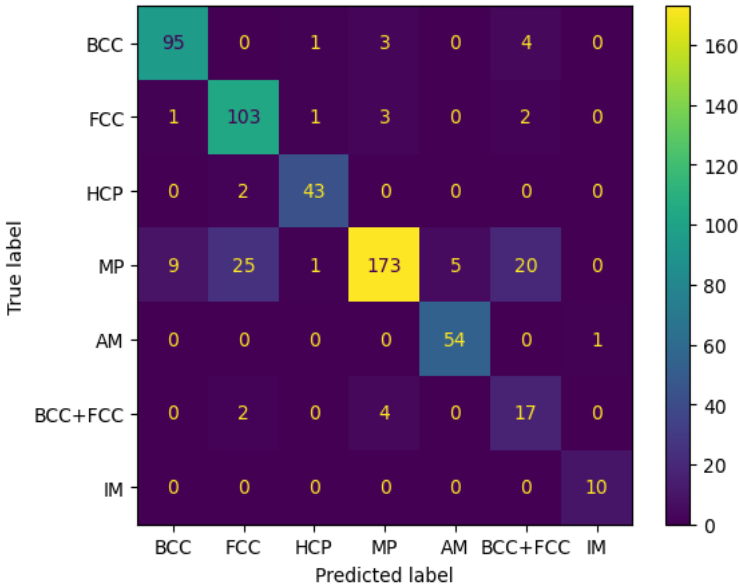


Figure 4.15: Confusion matrix of testing data for enhanced TLBO

Chapter 5

Periodic Table Representation

5.1 Introduction

5.1.1 Limitations of Hume-Rothery Rules in MPEAs

So far, the focus of feature selection in this study has been to designate appropriate chemical features based on Hume-Rothery rules inspired by studies in literature. However, despite their vast usage, Hume-Rothery rules exhibit several fundamental limitations that restrict their application in modern material science, mentioned in Section 2.2.3.

In light of these limitations, the Hume-Rothery rules are best regarded as historically significant, rather than a comprehensive guide for phase stability in complex alloy systems like MPEAs. They have been supplemented in literature by thermodynamic modelling and data-driven approaches that offer better predictive accuracy [22]. Inspired by the limitations, we aim to replace the Hume-Rothery rules altogether with a more comprehensive approach of extracting chemical information of complex alloys by utilizing the periodic table. The rest of this chapter describes in detail the methodology of periodic table representation and its implementation in phase prediction of MPEAs.

5.1.2 Qualitative Description of Periodic Table Representation

The periodic table representation occupies a central position in materials science as a framework that links atomic-scale electronic structure to macroscopic material properties. It is used as a predictive and interpretive tool for understanding the

relationships between structure and property across a wide range of material classes, such as metals, ceramics, semiconductors, and functional materials.

By definition, the periodic table organizes elements according to atomic number, electronic configuration, and recurring chemical trends in a matrix-like formation, where rows and columns are represented by groups and periods. These elemental descriptors form the basis for many fundamental materials parameters, such as atomic radius, electronegativity, valence electron concentration, and cohesive energy. These parameters directly influence crystal structure selection, bonding character, defect formation, diffusion kinetics, and thermodynamic phase stability [153] [36].

An key aspect of periodic table is the division into periods, groups, and electronic orbitals (s, p, d, and f). Periodic trends across rows and columns describe variations in atomic size, ionization energy, and electron affinity. These variations then in turn define trends in lattice parameters, elastic moduli, melting temperatures, and electrical and thermal transport properties. For example, the partially filled d-orbitals characteristic of transition metals give rise to their distinctive mechanical strength, magnetic behavior, and catalytic activity, making the d-block particularly important for alloy development and material design [18] [44].

An important thing to note is that the Hume-Rothery rules, which are experimentally derived guidelines, are themselves explainable by periodic table representation. This includes electronegativity-based criteria for compound stability, and valence electron concentration models for predicting intermetallic phases. In this context, the periodic table operates not merely as a descriptive chart but as a heuristic framework that enables qualitative prediction and rationalization of observed material behavior.

5.2 Methodology

5.2.1 Framework of Periodic Table Representation

To eliminate the need for feature hand-engineering via Hume-Rothery rules, in this study we adopt the periodic table representation (PTR) which has the ability to automatically engineer features by extracting necessary information from the matrix representation of PTR via a convolutional neural network (CNN). Instead of choosing alloy compositions as discrete variables from which alloy descriptors are calculated separately, PTR can correlate between different elements according to

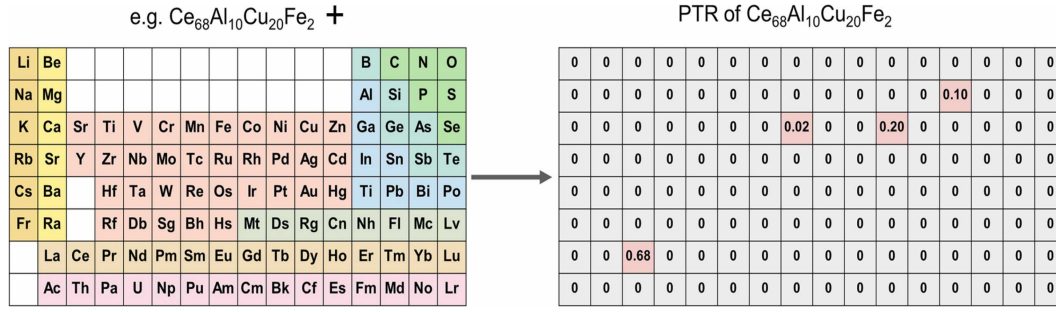


Figure 5.1: Illustration of periodic table representation with $\text{Ce}_{68}\text{Al}_{10}\text{Cu}_{20}\text{Fe}_2$ as an example [154]

their position and extract vital information necessary to describe the phase formation of the respective MPEA. To do this, we project the alloy composition onto the periodic table in the units of atomic number fraction and normalizing the sum of all fractions to 1. This creates an 8 by 16 pixel image where each pixel represents the atomic concentration of a particular element corresponding to its position in the periodic table. For simplification, we remove Group VII A, Group VIII A and Period I, because they do not contain any metals. Group VII A contains halogens, Group VIII A consists of nobel gases and Period 1 constitutes hydrogen and helium. As an example, in Fig. 5.1 we demonstrate the PTR matrix of $\text{Ce}_{68}\text{Al}_{10}\text{Cu}_{20}\text{Fe}_2$, a multicomponent alloy. The resulting matrix acts as a signature image of the corresponding alloy, and we can use it as a training input for a neural network for the purpose of designing a phase prediction framework.

A detailed description of PTR of an MPEA in the form of pseudo code is given in Algorithm 3, where c_i is the atomic concentration of element i and N is the total number of elements in the MPEA, as visually depicted in Fig. 5.1.

Algorithm 3 The Procedure of Periodic Table Representation

- 1: Initialize an 8 by 16 matrix P such that every element has value 0
 - 2: Normalize atomic concentrations, $\frac{c_i}{\sum_{i=1}^N c_i}$
 - 3: Select an element from the MPEA
 - 4: Identify its position in the matrix according to its period and group number in the periodic table
 - 5: Insert the value of c_i in the matrix with $P_{mn} = c_i$, where m is period number and n is group number of the element
 - 6: Select a different element and repeat Steps 3-5 until all elements have been considered
-

5.2.2 Convolutional Neural Network

Convolutional Neural Networks (CNNs) are a class of deep, feedforward ANNs designed specifically for grid-like data such as images. CNNs have become a foundational tool in the field of computer vision and they have been adopted in multiple scientific domains, including material science [155] [156], computational physics [157] [158] and medical imaging [159] [160]. Their increased usage is due to the fact that they can automatically learn hierarchal feature representation from raw data [161] [162].

The PTR dataset consists of an 8 by 16 pixel image of each MPEA, making CNNs a viable choice for supervised classification for this data. CNNs are particularly well-suited for image classification because they are designed to learn the hierarchal spatial patterns present in PTR images. Instead of requiring manually engineered features, CNNs automatically learn relevant features directly from raw pixel data.

The characteristic that defines CNNs is the use of convolution operations which replace the general matrix multiplication used in fully connected ANNs. In a convolutional layer, a set of filters are applied across the input domain to generate feature maps that encode local patterns of an image, such as edges and textures. This particular operation exploits local connectivity and parameter sharing which reduces the number of tunable parameters while preserving spatial relationships in the data [163].

We can represent a two-dimensional discrete convolution mathematically as

$$y(i, j) = \sum_m \sum_n x(i + m, j + n) \cdot w(m, n) \quad (5.1)$$

where x denotes the input feature map, w represents the convolutional kernel and y is the resulting output feature map. The summations are over the indices of the convolutional kernel m and n , whereas i and j represent the indices of the output matrix. Through successive convolutional layers, CNNs construct increasingly abstract representations of the input feature map.

In addition to convolutional layers, CNN architectures also include nonlinear activation functions such as rectified linear unit (ReLU) which introduces nonlinearity and improves gradient propagation during training. In between subsequent convolutional layers, pooling layers are incorporated to perform spatial downsampling. This decreases the computational cost and increases translational invariance. While pooling improves robustness, it can also lead to loss of fine spatial information. This

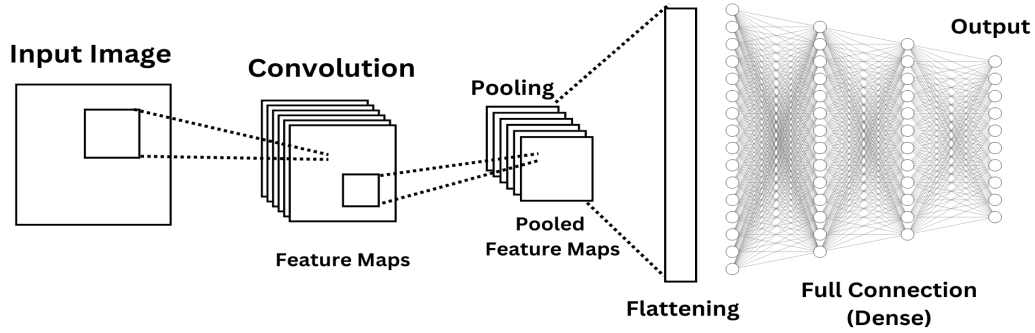


Figure 5.2: Depiction of a typical CNN architecture containing convolution, pooling and fully-connected layers.

prompted the development of alternative CNN architectures that reduce the use of pooling or eliminate it completely [164].

CNNs are trained using backpropagation in conjunction with gradient-based optimization algorithms such as stochastic gradient descent (SGD) or its adaptive variants (e.g., Adam). The number of stacked layers in a CNN has been shown to correlate with representational power as well as overfitting. Hence, innovations in architecture such as batch normalization have been proposed to address these issues [165]. These improvements enable the training of very deep CNNs with good convergence and performance [166].

Theoretically, CNNs can be interpreted as hierarchical feature extractors that encode global context while preserving local information. Their bias towards spatial locality makes them well suited for physical data where local interactions dominate the behaviour of the system. In material science, CNNs have been applied for microstructure classification, phase identification, defect detection and property prediction directly from images [167].

Figure 5.2 illustrates CNN architecture, depicting from right to left: input image, convolutional layers, pooling layers, flattening of 2D matrices into 1D arrays, and a fully-connected feed-forward ANN which delivers the desired output.

Despite their success, there are some limitations of CNNs. Their performance strongly depends on the availability of large, well-labeled datasets and their representations are often difficult to interpret physically, making their models a black-box. There is ongoing research which aims to improve model interpretability [168] as well

as incorporation of physical constraints [169]. There has also been work done on the integration of CNNs with physics-based modeling frameworks to enhance their reliability and generalization in scientific applications [170]. Another key limitation of CNNs is the risk of overfitting, where a dense architecture may lead to learning of irrelevant patterns and noise in the data. An overfitted model exhibits high accuracy on training data but performs poorly on testing data which it has not seen before. To mitigate this risk, we use dropout methodology in the fully-connected layer, illustrated in Chapter 3, Fig. 3.2.

CNNs are a powerful and versatile class of deep learning models that combine mathematical efficiency with inductive biases. Their ability to learn multiscale representation from structured data is the key motivation for their use in creating a phase prediction model using PTR as training input.

5.3 Experiments and Results

5.3.1 Data Collection and Preprocessing

The same dataset used in previous two chapters is utilized for PTR, with 1817 alloys and 7 possible classes: BCC, FCC, HCP, IM, AM, MP and BCC+FCC. Since we are using PTR instead of Hume-Rothery rules-based features, we do not need to calculate features for each input before feeding it into the network.

To calculate the normalized atomic concentration for PTR of a particular MPEA, we use the following formula.

$$c_{i_{norm}} = \frac{c_i}{\sum_{k=1}^N c_k} \quad (5.2)$$

where c_i and $c_{i_{norm}}$ denotes the concentration of original and normalized i^{th} element, and N is the number of total elements in an alloy.

5.3.2 Architecture of CNN

The architecture chosen in this study for CNN can be described as follows. There are six convolution layers, followed by a pooling layer, with the inputs of each convolution layer being passed through a ReLU function which serves as an activation function for each layer. The number of kernels in the convolutional layers is [16, 16, 32, 32, 64, 64] with unit stride and unit padding, to ensure that the informa-

tion on the edges of PTR matrix does not get eliminated. Max pooling is used as the pooling layer with kernel size=2 and stride=2.

Following the convolution and pooling layers, a fully-connected feed-forward network with two hidden layers is designed with variable number of neurons. Here, we incorporate the same strategy of hyperparameter optimization, where the hyperparameters considered for CNN architecture are number of epochs, batch size, learning rate, number of nodes in first fully connected layer, number of nodes in second fully connected layer and rate of dropout. These six hyperparameters will be tuned by various optimization algorithms discussed previously, and its performance will be analyzed via fitness plots. A comprehensive list of hyperparameters of CNN considered in this study as well as their respective ranges are depicted in Table 5.1.

Table 5.1: List of hyperparameters and their respective ranges.

Hyperparameter	Range
Number of epochs	[10, 1000]
Batch size	[5, 200]
Learning rate	$[10^{-4}, 10^{-1}]$
Number of nodes in 1 st hidden layer	[5, 50]
Number of nodes in 2 nd hidden layer	[5, 50]
Rate of dropout	[0, 1]

5.3.3 PTR Results

For implementation of PTR via GA, we initially choose population size as 20 and number of generations as 10, with elitism equal to 2. The results of fitness evolution with respect to number of generations is depicted in Figure 5.3. Fitness decreases in steps, depicting the existence of local minima, where each step takes the search to a lower minima, overall improving the testing accuracy of the model.

To validate the choice of this study, we compare fitness plots of GA with original TLBO and enhanced TLBO developed in the previous chapter, illustrated in Figure 5.4. The profile of the fitness plots clearly illustrate superior performance of enhanced TLBO compared with original TLBO and GA. This comparison further validates the novel modifications of TLBO utilized in the current study.

The values of hyperparameters obtained from the best result of PTR using enhanced TLBO for hyperparameter optimization are listed in Table 5.2.

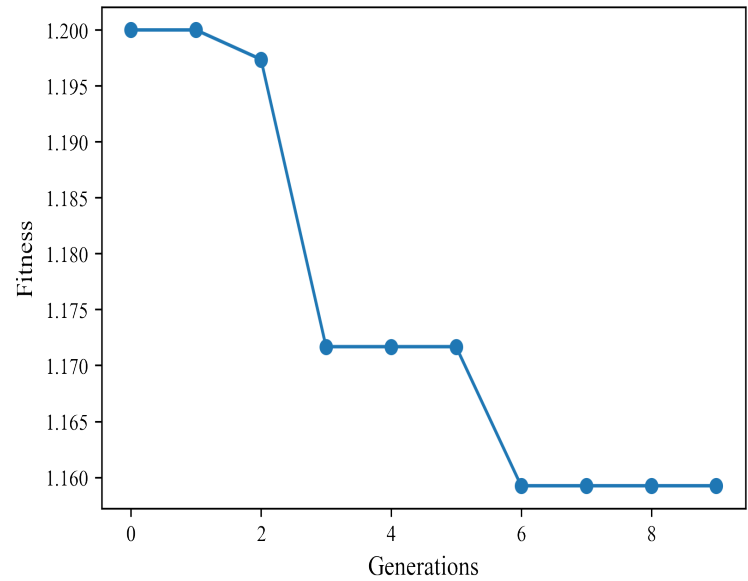


Figure 5.3: Fitness plot of GA with respect to number of generations for PTR. Here, Population size = 20, Number of Generations = 10, Elitism = 2

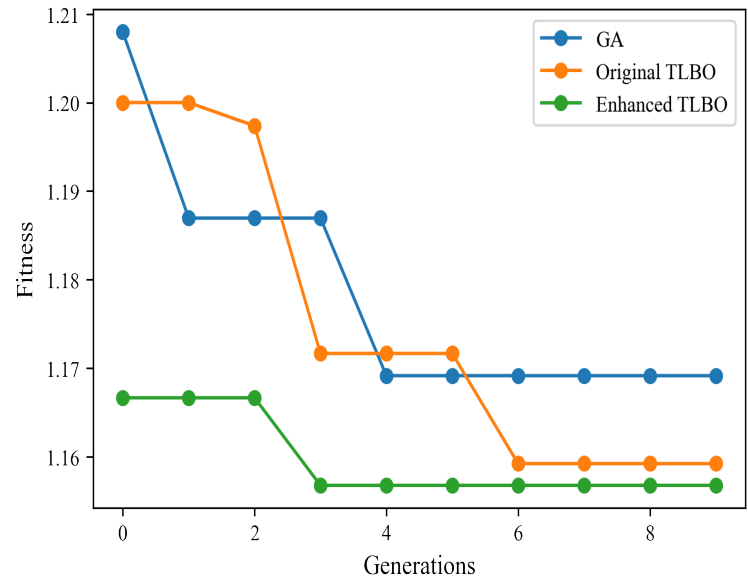


Figure 5.4: Comparison plot of fitness between GA, original TLBO and enhanced TLBO developed in Chapter 3

Table 5.2: List of hyperparameters with the best testing accuracy using PTR with enhanced TLBO

Hyperparameter	Range
Number of epochs	8.09×10^2
Batch size	1.93×10^2
Learning rate	1.4555×10^{-2}
Number of nodes in 1 st hidden layer	37
Number of nodes in 2 nd hidden layer	11
Rate of dropout	7.0545×10^{-2}

5.3.4 Validation of PTR Prediction Model

As discussed in Section 5.1, the limitations of Hume-Rothery rules have served as the key motivation for adopting periodic table representation of MPEAs. To validate this choice, three independent PTR-based phase prediction models are generated and their testing accuracy is compared with the best accuracies achieved from Hume-Rothery based features. Results of comparison are depicted in Table 5.3, where PTR-based approach outperforms models built with empirical rule-based features. Also, PTR with enhanced TLBO exhibits better testing accuracy compared with GA, further supporting the methodology developed in Chapter 4 of applying suitable modifications to TLBO to improve its capabilities and avoid premature convergence. Results obtained from this study indicate that PTR is a better indicator of MPEA microstructure as it contains all the relevant information without explicitly calculating them, which was the case for features guided by Hume-Rothery rules. PTR provides deeper insight into the chemical nature of multicomponent alloy systems and can be readily employed in data-driven prediction models for the purpose of high-throughput material discovery and design.

Table 5.3: Comparison of prediction accuracies between PTR-based optimization and use of Hume-Rothery features.

Optimization Strategy	Accuracy
Hume-Rothery features using GA	0.852831
Hume-Rothery features using enhanced TLBO	0.853941
PTR using GA	0.855311
PTR using enhanced TLBO	0.864468

5.4 Conclusion

Building upon the work presented in Chapter 3 and 4, in this chapter, use of periodic table representation is proposed with the novel application on MPEA phase prediction mechanism. The representation is tailored for MPEAs by replacing Hume-Rothery based features with a 2D image. Each pixel in the image represents atomic concentration of a principal element in an MPEA, essentially creating a unique matrix signature of each alloy. CNN is chosen for the purpose of creating a phase prediction mechanism for MPEAs and hyperparameter optimization is utilized with GA (Chapter 3) and enhanced TLBO (Chapter 4) to optimize the architecture of the CNN. Analysis of fitness plots for GA with elitism depict successful optimization and improvement in testing accuracy. Comparison among GA, original TLBO and enhanced TLBO indicate the superiority of enhanced TLBO in terms of rate of convergence and prediction accuracy. To validate the current approach, experiments are conducted with PTR-based feature engineering and Hume-Rothery rules-based features using both GA and enhanced TLBO. Comparison between testing accuracies illustrate that PTR with enhanced TLBO gives the best performance and can be readily utilized for creating a robust and accurate phase prediction framework for MPEAs.

Chapter 6

Conclusion

6.1 Introduction

In this chapter, first we provide a comprehensive summary of the work described in this thesis and the contributions of each chapter. Then, we provide indication of some future research directions that follow from this thesis.

6.2 Summary of the Research

In this thesis, we have developed new MPEA phase prediction algorithms by addition of novel-hyperparameters, modifications in existing optimization techniques and improved feature extraction via periodic table placement of metals.

In Chapter 1, we have introduced background and context regarding the work done in this thesis, beginning with an overview of the emerging field of MPEAs, which are attracting researchers with their promising combination of mechanical and functional properties, signifying the importance of their study from both a fundamental scientific perspective as well as their potential use as circular alloys. The use of data-driven techniques for high-throughput MPEA design has been discussed by highlighting their use in prediction of critical alloy properties, including atomic phases and mechanical properties. These supervised machine learning models can provide suggestive guidelines for MPEAs with tailored properties suitable for applications in harsh environments with high temperature and pressure conditions. The concept of hyperparameter optimization and population-based heuristic optimization algorithms has been introduced with emphasis on GA and TLBO. The main aims and problems tackled in this work have been discussed, with the focus

on research questions by addressing gaps in existing knowledge and limitations of current research directions. Core objectives of this work have been stated alongside the contribution to knowledge resulting from the current research. Scope and limitations are discussed, including the core focus on MPEAs as opposed to other complex alloy systems. A brief overview of the methodology has been presented while also discussing challenges faced in this work.

In Chapter 2, a comprehensive background of topics covered in the current research has been discussed, starting with MPEAs. Their definition, landmark work, high configurational entropy and multi-component nature is depicted with focus on finding single-phase solid solutions. The use of machine learning techniques to effectively create a MPEA design framework has been presented with focus on classification mechanisms for predicting alloy phases, and regression mechanisms for predicting physical and chemical properties. Flexibility and adaptive nature of data-driven algorithms is emphasized and supported by discussing recent and relevant research done in this domain of material science and machine learning. A brief theoretical description of GA and TLBO has been included, with the focus on their advantages and limitations. The global search capability and flexibility of both optimization algorithms are shown to make them prime candidates for high-throughput design of multi-component and high-dimensional alloy systems, not unlike MPEAs.

In Chapter 3, we have proposed a novel hyperparameter optimization method for ANNs with the addition of two novel hyperparameters: outlier detection and feature subset selection. These hyperparameters serve to simultaneously pre-process and train it to provide accurate alloy phase predictions. GA is utilized as the optimization algorithm with a novel dynamic length chromosome operator which can handle genetic operations for variable length individuals in the population. Results obtained from ablations studies have indicated significant improvement of accuracy with the inclusion of new hyperparameters, hence justifying our approach. We have tested this algorithm on recently published MPEAs, and the predicted alloy phases have been shown to be in agreement with actual phases, making this a viable approach for aiding experimental analysis of MPEAs.

In Chapter 4, a hyperparameter approach of ANNs has been proposed for phase prediction of MPEAs. For this purpose, we have utilized TLBO as a heuristic optimization algorithm without tunable parameters. Results of comparison tests on benchmark functions between GA and TLBO have indicated that TLBO is significantly better, but does not outperform GA when applied to MPEA data. Original

TLBO suffers from premature convergence, therefore we have proposed three novel modifications to promote balanced search. Experimental results have shown viability of this approach as a phase prediction framework. Analysis of the three novel variants of TLBO have depicted successful improvement in optimization of hyperparameters of ANN compared with original TLBO. A novel approach of testing all possible permutations of modifications have indicated that variant 1 and 2 combined result in the best testing accuracy. We have termed this superior combination as enhanced TLBO and have compared its performance to GA to further validate the current approach. Results of comparison have indicated improvement in performance utilizing enhanced TLBO, while overcoming the limitations of original TLBO and keeping the benefits intact. The contribution towards high-throughput development of MPEAs by assisting experimental design via data-driven approach have been highlighted, as well as contribution towards theoretical framework of complex multi-dimensional alloy systems and their physical characteristics.

In Chapter 5, use of periodic table representation has been proposed with the novel application on MPEA phase prediction mechanism. The representation is tailored for MPEAs by replacing Hume-Rothery based features with a 2D image. Each pixel in the image represents atomic concentration of a principal element in an MPEA, essentially creating a unique matrix signature of each alloy. CNN is chosen for the purpose of creating a phase prediction mechanism for MPEAs and hyperparameter optimization has been utilized with GA and enhanced TLBO to optimize the architecture of the CNN. Analysis of fitness plots for GA with elitism have depicted successful optimization and improvement in testing accuracy. Comparison among GA, original TLBO and enhanced TLBO have indicated the superiority of enhanced TLBO in terms of rate of convergence and prediction accuracy. To validate the current approach, experiments have been conducted with PTR-based feature engineering and Hume-Rothery rules-based features using both GA and enhanced TLBO. Comparison between testing accuracies have illustrated that PTR with enhanced TLBO gives the best performance and can be readily utilized for creating a robust and accurate phase prediction framework for MPEAs.

6.3 Future Research Directions

In this thesis, we have studied evolutionary optimization algorithms and supervised predictive learning via ANNs. To address the challenges, different approaches

of optimization and architecture have been proposed with the hopes of improving prediction accuracy while working with a small dataset. Although the developed approaches have proven to be successful when applied to MPEAs, there is still ample room for improvement. Below are some future research directions relevant to the work presented in this thesis.

- In this thesis, the effectiveness of the proposed approaches is evaluated on MPEA data. In the future, we aim to extend this approach to other types of alloys and materials where their physical properties are dependent on the structural order of atoms. These include bulk metallic glasses, complex concentrated alloys, eutectic alloys and refractory HEAs.
- The performance of population-based evolutionary algorithms is dependent on certain tunable parameters such as population size and initialization of population. This is a challenging task when dealing with high-dimensional systems where the aim is to discover optimum solutions in reasonable time. Under these circumstances, a future research direction is to carefully consider the initialization methodology of optimization algorithms and create new variants of TLBO with improved convergence.
- In this thesis, the data used for training and testing the prediction models is obtained from experimental analysis only. This is a major hurdle for generalization of the search space because experimental validation is limited and time consuming. Other studies have used simulated data obtained from thermodynamic calculation and ab-initio quantum mechanics as well as synthetic data from generative algorithms to increase the dataset. A future research approach could be the exploration of simulated and synthetic data and their performance on the proposed methodologies in our work.

Bibliography

- [1] B. Cantor, I. T. Chang, P. Knight, and A. Vincent, “Microstructural development in equiatomic multicomponent alloys,” *Materials Science and Engineering: A*, vol. 375, pp. 213–218, 2004.
- [2] K. L. Elder, J. Berry, B. Bocklund, S. K. McCall, A. Perron, and J. T. McKeown, “Computational discovery of ultra-strong, stable and lightweight refractory multi-principal element alloys. part i: Design principles and rapid selection,” *npj Computational Materials*, vol. 9, no. 1, p. 84, 2023.
- [3] F. Ghadami and M. A. Davoudabadi, “Ultra-high entropy alloys: Entropy-driven design, microstructural evolution, and pathways to extreme-environment resilience,” *Jouranal of Material Research and Technology*, vol. 41, pp. 2452–2466, 2026.
- [4] D. B. Miracle and O. N. Senkov, “A critical review of high entropy alloys and related concepts,” *Acta materialia*, vol. 122, pp. 448–511, 2017.
- [5] M. Hu et al., “Recent applications of machine learning in alloy design: A review,” *Materials Science and Engineering: R: Reports*, vol. 155, p. 100746, 2023.
- [6] N. H. Chau, M. Kubo, L. V. Hai, and T. Yamamoto, “Support vector machine-based phase prediction of multi-principal element alloys,” *Vietnam Journal of Computer Science*, vol. 10, no. 01, pp. 101–116, 2023.
- [7] K. Lee, M. V. Ayyasamy, P. Delsa, T. Q. Hartnett, and P. V. Balachandran, “Phase classification of multi-principal element alloys via interpretable machine learning,” *npj Computational Materials*, vol. 8, no. 1, p. 25, 2022.
- [8] J. Berry and K. A. Christofidou, “Supervised machine learning for multi-principal element alloy structural design,” *Materials Science and Technology*, vol. 41, no. 11, pp. 773–790, 2025.

- [9] E. Gerashi et al., “Machine learning-aided phase and mechanical properties prediction in multi-principal element alloys,” *Computational Materials Science*, vol. 243, p. 113 114, 2024.
- [10] C. Jin et al., “High-throughput calculation and machine learning-assisted prediction of the mechanical properties of refractory multi-principal element alloys,” *Advanced Theory and Simulations*, vol. 9, no. 1, e00784, 2026.
- [11] J. Ren and L. Xiao, “Physics-guided neural network for fatigue life prediction of fcc-based multi-principal element alloys,” *Scripta Materialia*, vol. 253, p. 116 307, 2024.
- [12] D. Theng and K. K. Bhoyar, “Feature selection techniques for machine learning: A survey of more than two decades of research,” *Knowledge and Information Systems*, vol. 66, no. 3, pp. 1575–1637, 2024.
- [13] P. Probst, A.-L. Boulesteix, and B. Bischl, “Tunability: Importance of hyperparameters of machine learning algorithms,” *Journal of Machine Learning Research*, vol. 20, no. 53, pp. 1–32, 2019.
- [14] M. Mitchell, *An introduction to genetic algorithms*. MIT press, 1998.
- [15] R. V. Rao, “Teaching-learning-based optimization algorithm,” in *Teaching learning based optimization algorithm: and its engineering applications*, Springer, 2015, pp. 9–39.
- [16] R. W. Cahn and P. Haasen, *Physical metallurgy*. Elsevier, 1996, vol. 1.
- [17] S. Nene, S. Sinha, D. Yadav, and A. Dutta, “Metallurgical aspects of high entropy alloys,” *Journal of Alloys and Compounds*, vol. 1005, p. 175 849, 2024.
- [18] W Hume-Rothery, “Atomic diameters, atomic volumes and solid solubility relations in alloys,” *Acta Metallurgica*, vol. 14, no. 1, pp. 17–20, 1966.
- [19] H Okamoto and T. B. Massalski, “Guidelines for binary phase diagram assessment,” *Journal of phase equilibria*, vol. 14, no. 3, pp. 316–335, 1993.
- [20] J.-W. Yeh et al., “Nanostructured high-entropy alloys with multiple principal elements: Novel alloy design concepts and outcomes,” *Advanced engineering materials*, vol. 6, no. 5, pp. 299–303, 2004.
- [21] E. P. George, D. Raabe, and R. O. Ritchie, “High-entropy alloys,” *Nature reviews materials*, vol. 4, no. 8, pp. 515–534, 2019.

- [22] S. Curtarolo, G. L. Hart, M. B. Nardelli, N. Mingo, S. Sanvito, and O. Levy, “The high-throughput highway to computational materials design,” *Nature materials*, vol. 12, no. 3, pp. 191–201, 2013.
- [23] M. H. Golbabaei, M. Zohrevand, and N. Zhang, “Applications of machine learning in high-entropy alloys: A review of recent advances in design, discovery, and characterization,” *Nanoscale*, vol. 17, no. 36, pp. 20 548–20 605, 2025.
- [24] Z. Rao et al., “Machine learning-enabled high-entropy alloy discovery,” *Science*, vol. 378, no. 6615, pp. 78–85, 2022.
- [25] E. Pickering and N. Jones, “High-entropy alloys: A critical assessment of their founding principles and future prospects,” *International Materials Reviews*, vol. 61, no. 3, pp. 183–202, 2016.
- [26] D. A. Porter and K. E. Easterling, *Phase transformations in metals and alloys (revised reprint)*. CRC press, 2009.
- [27] U. Mizutani, “The hume-rothery rules for structurally complex alloy phases,” in *Surface properties and engineering of complex intermetallics*, World Scientific, 2010, pp. 323–399.
- [28] J. K. Pickard, J. A. Carretero, and V. C. Bhavsar, “On the convergence and origin bias of the teaching-learning-based-optimization algorithm,” *Applied Soft Computing*, vol. 46, pp. 115–127, 2016.
- [29] C. Chen, H. Zhou, W. Long, G. Wang, and J. Ren, “Phase prediction for high-entropy alloys using generative adversarial network and active learning based on small datasets,” *Science China Technological Sciences*, vol. 66, no. 12, pp. 3615–3627, 2023.
- [30] Z. Yang, S. Li, S. Li, J. Yang, and D. Liu, “A two-step data augmentation method based on generative adversarial network for hardness prediction of high entropy alloy,” *Computational Materials Science*, vol. 220, p. 112 064, 2023.
- [31] T. Chen, T. Shun, J. Yeh, and M. Wong, “Nanostructured nitride films of multi-element high-entropy alloys by reactive dc sputtering,” *Surface and Coatings Technology*, vol. 188, pp. 193–200, 2004.

- [32] C.-Y. Hsu, J.-W. Yeh, S.-K. Chen, and T.-T. Shun, “Wear resistance and high-temperature compression strength of fcc cuconical0. 5fe alloy with boron addition,” *Metallurgical and Materials Transactions A*, vol. 35, no. 5, pp. 1465–1469, 2004.
- [33] P.-K. Huang, J.-W. Yeh, T.-T. Shun, and S.-K. Chen, “Multi-principal-element alloys with improved oxidation and wear resistance for thermal spray coating,” *Advanced Engineering Materials*, vol. 6, no. 1-2, pp. 74–78, 2004.
- [34] J.-W. Yeh et al., “Formation of simple crystal structures in cu-co-ni-cr-al-fe-ti-v alloys with multiprincipal metallic elements,” *Metallurgical and Materials Transactions A*, vol. 35, no. 8, pp. 2533–2536, 2004.
- [35] B. Murty, J. Yeh, S Ranganathan, and P. Bhattacharjee, “High-entropy alloys. butterworth-heinemann (elsevier),” 2014.
- [36] W. D. Callister Jr and D. G. Rethwisch, *Materials science and engineering: an introduction*. John wiley & sons, 2020.
- [37] A. R. Paul, M. Mukherjee, and D. Singh, “A critical review on the properties of intermetallic compounds and their application in the modern manufacturing,” *Crystal research and technology*, vol. 57, no. 3, p. 2100159, 2022.
- [38] Z. Feng, H. Geng, Y. Zhuang, and P. Li, “Progress, applications, and challenges of amorphous alloys: A critical review,” *Inorganics*, vol. 12, no. 9, p. 232, 2024.
- [39] S. Mandal, R. Ghosh, S. Kunar, R. Samanta, A. Sinha, and G. Mandal, “Current development of multiphase steel: Alloy design, properties and application,” *Journal of The Institution of Engineers (India): Series D*, vol. 105, no. 3, pp. 2023–2038, 2024.
- [40] D. R. F. West, *Ternary phase diagrams in materials science*. CRC press, 2017.
- [41] W. Wang and Z. Li, *CALPHAD as a Toolbox to Facilitate the Development of HEAs*. IntechOpen, 2022.
- [42] Y. Jien-Wei, “Recent progress in high entropy alloys,” *ann. Chim. sci. Mat*, vol. 31, no. 6, pp. 633–648, 2006.

- [43] Y. Zhang, Y. J. Zhou, J. P. Lin, G. L. Chen, and P. K. Liaw, “Solid-solution phase formation rules for multi-component alloys,” *Advanced engineering materials*, vol. 10, no. 6, pp. 534–538, 2008.
- [44] U. Mizutani, *Introduction to the electron theory of metals*. Cambridge university press, 2001.
- [45] Y. Zhang et al., “Microstructures and properties of high-entropy alloys,” *Progress in materials science*, vol. 61, pp. 1–93, 2014.
- [46] B. Cantor, “Multicomponent and high entropy alloys,” *Entropy*, vol. 16, no. 9, pp. 4749–4768, 2014.
- [47] H. Huo and M. Rupp, “Unified representation of molecules and crystals for machine learning,” *Machine Learning: Science and Technology*, vol. 3, no. 4, p. 045 017, 2022.
- [48] F. Häse, L. M. Roch, and A. Aspuru-Guzik, “Chimera: Enabling hierarchy based multi-objective optimization for self-driving laboratories,” *Chemical science*, vol. 9, no. 39, pp. 7642–7655, 2018.
- [49] J. P. Janet, C. Duan, T. Yang, A. Nandy, and H. J. Kulik, “A quantitative uncertainty metric controls error in neural network-driven chemical discovery,” *Chemical science*, vol. 10, no. 34, pp. 7913–7922, 2019.
- [50] J. Ling, M. Hutchinson, E. Antono, S. Paradiso, and B. Meredig, “High-dimensional materials and process optimization using data-driven experimental design with well-calibrated uncertainty estimates,” *Integrating Materials and Manufacturing Innovation*, vol. 6, no. 3, pp. 207–217, 2017.
- [51] S. Cuomo, V. S. Di Cola, F. Giampaolo, G. Rozza, M. Raissi, and F. Piccialli, “Scientific machine learning through physics-informed neural networks: Where we are and what’s next,” *Journal of Scientific Computing*, vol. 92, no. 3, p. 88, 2022.
- [52] M. Yao, K. G. Pradeep, C. C. Tasan, and D. Raabe, “A novel, single phase, non-equiatomic femnnicocr high-entropy alloy with exceptional phase stability and tensile ductility,” *Scripta Materialia*, vol. 72, pp. 5–8, 2014.
- [53] M. C. Gao and D. E. Alman, “Searching for next single-phase high-entropy alloy compositions,” *Entropy*, vol. 15, no. 10, pp. 4504–4519, 2013.

- [54] M. C. Tropicovsky, J. R. Morris, P. R. Kent, A. R. Lupini, and G. M. Stocks, “Criteria for predicting the formation of single-phase high-entropy alloys,” *Physical Review X*, vol. 5, no. 1, p. 011 041, 2015.
- [55] Y. Lederer, C. Toher, K. S. Vecchio, and S. Curtarolo, “The search for high entropy alloys: A high-throughput ab-initio approach,” *Acta Materialia*, vol. 159, pp. 364–383, 2018.
- [56] Y. Li and W. Guo, “Machine-learning model for predicting phase formations of high-entropy alloys,” *Physical Review Materials*, vol. 3, no. 9, p. 095 005, 2019.
- [57] T. Kostiuchenko, F. Körmann, J. Neugebauer, and A. Shapeev, “Impact of lattice relaxations on phase transitions in a high-entropy alloy studied by machine-learning potentials,” *npj Computational Materials*, vol. 5, no. 1, p. 55, 2019.
- [58] Z. Song et al., “A quick screening of high yield strength points in single-phase co-cr-fe-ni-mo high entropy alloy system based on electronegativity, thermodynamic calculations and machine learning,” *Thermodynamic Calculations and Machine Learning*,
- [59] M.-H. Tsai, R.-C. Tsai, T. Chang, and W.-F. Huang, “Intermetallic phases in high-entropy alloys: Statistical analysis of their prevalence and structural inheritance,” *Metals*, vol. 9, no. 2, p. 247, 2019.
- [60] N. Islam, W. Huang, and H. L. Zhuang, “Machine learning for phase selection in multi-principal element alloys,” *Computational Materials Science*, vol. 150, pp. 230–235, 2018.
- [61] Z. Zhou, Y. Zhou, Q. He, Z. Ding, F. Li, and Y. Yang, “Machine learning guided appraisal and exploration of phase design for high entropy alloys,” *npj Computational Materials*, vol. 5, no. 1, p. 128, 2019.
- [62] S. Dixit, V. Singhal, A. Agarwal, and A. P. Rao, “Multi-label phase-prediction in high-entropy-alloys using artificial-neural-network,” *Materials Letters*, vol. 268, p. 127 606, 2020.
- [63] S. Y. Lee, S. Byeon, H. S. Kim, H. Jin, and S. Lee, “Deep learning-based phase prediction of high-entropy alloys: Optimization, generation, and explanation,” *Materials & Design*, vol. 197, p. 109 260, 2021.

- [64] X. Yang and Y. Zhang, "Prediction of high-entropy stabilized solid-solution in multi-component alloys," *Materials Chemistry and Physics*, vol. 132, no. 2-3, pp. 233–238, 2012.
- [65] L. Xie, P. Brault, A.-L. Thomann, X. Yang, Y. Zhang, and G. Shang, "Molecular dynamics simulation of al-co-cr-cu-fe-ni high entropy alloy thin film growth," *Intermetallics*, vol. 68, pp. 78–86, 2016.
- [66] R. Chen et al., "Composition design of high entropy alloys using the valence electron concentration to balance strength and ductility," *Acta Materialia*, vol. 144, pp. 129–137, 2018.
- [67] K. Deb, "Genetic algorithm in search and optimization: The technique and applications," in *Proceedings of International Workshop on Soft Computing and Intelligent Systems, (ISI, Calcutta, India)*, Proceedings of International Workshop on Soft Computing and Intelligent ..., 1998, pp. 58–87.
- [68] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: Nsga-ii," *IEEE transactions on evolutionary computation*, vol. 6, no. 2, pp. 182–197, 2002.
- [69] A. Kumar, "Encoding schemes in genetic algorithm," *International Journal of Advanced Research in IT and Engineering*, vol. 2, no. 3, pp. 1–7, 2013.
- [70] K. Deb, "An efficient constraint handling method for genetic algorithms," *Computer methods in applied mechanics and engineering*, vol. 186, no. 2-4, pp. 311–338, 2000.
- [71] A. E. Eiben and J. E. Smith, *Introduction to evolutionary computing*. Springer, 2015.
- [72] M. Angelova and T. Pencheva, "Tuning genetic algorithm parameters to improve convergence time," *International Journal of Chemical Engineering*, vol. 2011, no. 1, p. 646 917, 2011.
- [73] R. V. Rao, V. J. Savsani, and D. P. Vakharia, "Teaching-learning-based optimization: A novel method for constrained mechanical design optimization problems," *Computer-aided design*, vol. 43, no. 3, pp. 303–315, 2011.
- [74] A. Tiwari and M. Pradhan, "Applications of tlbo algorithm on various manufacturing processes: A review," *Materials Today: Proceedings*, vol. 4, no. 2, pp. 1644–1652, 2017.

- [75] B. S. Babu, “Tlbo based power system optimization for ac/dc hybrid systems,” in *Journal of Physics: Conference Series*, IOP Publishing, vol. 1916, 2021, p. 012023.
- [76] L. R. Rodrigues and J. P. P. Gomes, “Tlbo with variable weights applied to shop scheduling problems,” *CAAI Transactions on Intelligence Technology*, vol. 4, no. 3, pp. 148–158, 2019.
- [77] D Vidyabharathi and V Mohanraj, “Hyperparameter tuning for deep neural networks based optimization algorithm.,” *Intelligent Automation & Soft Computing*, vol. 36, no. 3, 2023.
- [78] Y Yun, M. Gen, and T. N. Erdene, “Applying ga-pso-tlbo approach to engineering optimization problems,” *Math. Biosci. Eng*, vol. 20, no. 1, pp. 552–571, 2023.
- [79] S. C. Satapathy and A. Naik, “Modified teaching–learning-based optimization algorithm for global numerical optimization—a comparative study,” *Swarm and Evolutionary Computation*, vol. 16, pp. 28–37, 2014.
- [80] R. V. Rao, V. Patel, et al., “An elitist teaching-learning-based optimization algorithm for solving complex constrained optimization problems,” *international journal of industrial engineering computations*, vol. 3, no. 4, pp. 535–560, 2012.
- [81] A. Moazeni, R. Khorsand, and M. Ramezanpour, “Dynamic resource allocation using an adaptive multi-objective teaching-learning based optimization algorithm in cloud,” *IEEE Access*, vol. 11, pp. 23 407–23 419, 2023.
- [82] T. Han et al., “Atlbo-q: Adaptive tlbo-based quantization for mmse mimo detection,” *IEEE Wireless Communications Letters*, 2026.
- [83] M.-Y. Cheng and D. Prayogo, “Fuzzy adaptive teaching–learning-based optimization for global numerical optimization,” *Neural Computing and Applications*, vol. 29, no. 2, pp. 309–327, 2018.
- [84] R. Mishra, N. Sharma, and H. Sharma, “Half-life teaching factor based tlbo algorithm,” in *Advances in Data and Information Sciences: Proceedings of ICDIS 2019*, Springer, 2020, pp. 257–269.
- [85] K. Yu, X. Wang, and Z. Wang, “An improved teaching-learning-based optimization algorithm for numerical and engineering optimization problems,” *Journal of Intelligent Manufacturing*, vol. 27, no. 4, pp. 831–843, 2016.

- [86] B. Gludovatz et al., “Exceptional damage-tolerance of a medium-entropy alloy croni at cryogenic temperatures,” *Nature communications*, vol. 7, no. 1, p. 10602, 2016.
- [87] J. Wei et al., “Machine learning in materials science,” *InfoMat*, vol. 1, no. 3, pp. 338–358, 2019.
- [88] D. King, S. Middleburgh, A. McGregor, and M. Cortie, “Predicting the formation and stability of single phase high-entropy alloys,” *Acta Materialia*, vol. 104, pp. 172–179, 2016.
- [89] S. Guo, “Phase selection rules for cast high entropy alloys: An overview,” *Materials Science and Technology*, vol. 31, no. 10, pp. 1223–1230, 2015.
- [90] Y. Zhang et al., “Phase prediction in high entropy alloys with a rational selection of materials descriptors and machine learning models,” *Acta Materialia*, vol. 185, pp. 528–539, 2020.
- [91] W. Huang, P. Martin, and H. L. Zhuang, “Machine-learning phase prediction of high-entropy alloys,” *Acta Materialia*, vol. 169, pp. 225–236, 2019.
- [92] Z. Pei, J. Yin, J. A. Hawk, D. E. Alman, and M. C. Gao, “Machine-learning informed prediction of high-entropy solid solution formation: Beyond the hume-rothery rules,” *npj Computational Materials*, vol. 6, no. 1, p. 50, 2020.
- [93] K. Kaufmann and K. S. Vecchio, “Searching for high entropy alloys: A machine learning approach,” *Acta Materialia*, vol. 198, pp. 178–222, 2020.
- [94] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, “Learning representations by back-propagating errors,” *nature*, vol. 323, no. 6088, pp. 533–536, 1986.
- [95] B. D. Ripley, *Pattern recognition and neural networks*. Cambridge university press, 2007.
- [96] M. Wistuba, N. Schilling, and L. Schmidt-Thieme, “Hyperparameter optimization machines,” in *2016 IEEE international conference on data science and advanced analytics (DSAA)*, IEEE, 2016, pp. 41–50.
- [97] J. Bergstra and Y. Bengio, “Random search for hyper-parameter optimization,” *The journal of machine learning research*, vol. 13, no. 1, pp. 281–305, 2012.

- [98] R. Bardenet, M. Brendel, B. Kégl, and M. Sebag, “Collaborative hyperparameter tuning,” in *International conference on machine learning*, PMLR, 2013, pp. 199–207.
- [99] J. H. Holland, “Genetic algorithms and adaptation,” in *Adaptive control of ill-defined systems*, Springer, 1984, pp. 317–333.
- [100] H. Alibrahim and S. A. Ludwig, “Hyperparameter optimization: Comparing genetic algorithm against grid search and bayesian optimization,” in *2021 IEEE congress on evolutionary computation (CEC)*, IEEE, 2021, pp. 1551–1559.
- [101] S. Nikbakht, C. Anitescu, and T. Rabczuk, “Optimizing the neural network hyperparameters utilizing genetic algorithm,” *Journal of Zhejiang University-Science A*, vol. 22, no. 6, pp. 407–426, 2021.
- [102] S. Lee, J. Kim, H. Kang, D.-Y. Kang, and J. Park, “Genetic algorithm based deep learning neural network structure and hyperparameter optimization,” *Applied Sciences*, vol. 11, no. 2, p. 744, 2021.
- [103] X. Xiao, M. Yan, S. Basodi, C. Ji, and Y. Pan, “Efficient hyperparameter optimization in deep learning using a variable length genetic algorithm,” *arXiv preprint arXiv:2006.12703*, 2020.
- [104] N. Reunanen, T. Rätty, and T. Lintonen, “Automatic optimization of outlier detection ensembles using a limited number of outlier examples,” *International journal of data science and analytics*, vol. 10, no. 4, pp. 377–394, 2020.
- [105] R. Ferro, G. Cacciamani, and G. Borzone, “Remarks about data reliability in experimental and computational alloy thermochemistry,” *Intermetallics*, vol. 11, no. 11-12, pp. 1081–1094, 2003.
- [106] G. Sheng and C. T. Liu, “Phase stability in high entropy alloys: Formation of solid-solution phase or amorphous phase,” *Progress in Natural Science: Materials International*, vol. 21, no. 6, pp. 433–446, 2011.
- [107] L. Ward, A. Agrawal, A. Choudhary, and C. Wolverton, “A general-purpose machine learning framework for predicting properties of inorganic materials,” *npj Computational Materials*, vol. 2, no. 1, pp. 1–7, 2016.
- [108] B. Meredig and C. Wolverton, “A hybrid computational–experimental approach for automated crystal structure solution,” *Nature materials*, vol. 12, no. 2, pp. 123–127, 2013.

- [109] Y. Lin, J. Zhang, and J. Zhong, “Application of neural networks to predict the elevated temperature flow behavior of a low alloy steel,” *Computational Materials Science*, vol. 43, no. 4, pp. 752–758, 2008.
- [110] C. Liu, “Physical metallurgy and mechanical properties of ductile ordered alloys (fe, co, ni) 3 v,” *International metals reviews*, vol. 29, no. 1, pp. 168–194, 1984.
- [111] J. Zhu, P. Liaw, and C. Liu, “Effect of electron concentration on the phase stability of nbcr2-based laves phase alloys,” *Materials Science and Engineering: A*, vol. 239, pp. 260–264, 1997.
- [112] S. Fang, X. Xiao, L. Xia, W. Li, and Y. Dong, “Relationship between the widths of supercooled liquid regions and bond parameters of mg-based bulk metallic glasses,” *Journal of Non-Crystalline Solids*, vol. 321, no. 1-2, pp. 120–125, 2003.
- [113] A. Takeuchi and A. Inoue, “Calculations of mixing enthalpy and mismatch entropy for ternary amorphous alloys,” *Materials transactions, JIM*, vol. 41, no. 11, pp. 1372–1378, 2000.
- [114] A. Takeuchi and A. Inoue, “Classification of bulk metallic glasses by atomic size difference, heat of mixing and period of constituent elements and its application to characterization of the main alloying element,” *Materials transactions*, vol. 46, no. 12, pp. 2817–2829, 2005.
- [115] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, “Improving neural networks by preventing co-adaptation of feature detectors,” *arXiv preprint arXiv:1207.0580*, 2012.
- [116] V. Nair and G. E. Hinton, “Rectified linear units improve restricted boltzmann machines,” in *Proceedings of the 27th international conference on machine learning (ICML-10)*, 2010, pp. 807–814.
- [117] J. Zou, Y. Han, and S.-S. So, “Overview of artificial neural networks,” *Artificial neural networks: methods and applications*, pp. 14–22, 2009.
- [118] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: A simple way to prevent neural networks from overfitting,” *The journal of machine learning research*, vol. 15, no. 1, pp. 1929–1958, 2014.

- [119] I. Y. Kim and O. de Weck, “Variable chromosome length genetic algorithm for structural topology design optimization,” in *45th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics & Materials Conference*, 2004, p. 1911.
- [120] I. Y. Kim and O. De Weck, “Variable chromosome length genetic algorithm for progressive refinement in topology optimization,” *Structural and Multidisciplinary Optimization*, vol. 29, no. 6, pp. 445–456, 2005.
- [121] L. Cruz-Piris, I. Marsa-Maestre, and M. A. Lopez-Carmona, “A variable-length chromosome genetic algorithm to solve a road traffic coordination multipath problem,” *IEEE Access*, vol. 7, pp. 111 968–111 981, 2019.
- [122] M. Feurer and F. Hutter, “Hyperparameter optimization,” in *Automated machine learning: Methods, systems, challenges*, Springer International Publishing Cham, 2019, pp. 3–33.
- [123] G. Melis, C. Dyer, and P. Blunsom, “On the state of the art of evaluation in neural language models,” *arXiv preprint arXiv:1707.05589*, 2017.
- [124] J. Bergstra, D. Yamins, and D. Cox, “Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures,” in *International conference on machine learning*, PMLR, 2013, pp. 115–123.
- [125] M. Frigge, D. C. Hoaglin, and B. Iglewicz, “Some implementations of the boxplot,” *The American Statistician*, vol. 43, no. 1, pp. 50–54, 1989.
- [126] L. Yu and H. Liu, “Feature selection for high-dimensional data: A fast correlation-based filter solution,” in *Proceedings of the 20th international conference on machine learning (ICML-03)*, 2003, pp. 856–863.
- [127] L. Talavera, “An evaluation of filter and wrapper methods for feature selection in categorical clustering,” in *International Symposium on Intelligent Data Analysis*, Springer, 2005, pp. 440–451.
- [128] R. Muthukrishnan and R. Rohini, “Lasso: A feature selection technique in predictive modeling for machine learning,” in *2016 IEEE international conference on advances in computer applications (ICACA)*, Ieee, 2016, pp. 18–20.
- [129] J. Yang et al., “Effects of transformation-induced plasticity (trip) on tensile property improvement of fe45co30cr10v10ni5-xmnx high-entropy alloys,” *Materials Science and Engineering: A*, vol. 772, p. 138 809, 2020.

- [130] D. G. Kim et al., “Effects of annealing temperature on microstructures and tensile properties of a single fcc phase coccumnni high-entropy alloy,” *Journal of Alloys and Compounds*, vol. 812, p. 152 111, 2020.
- [131] J. Moon, J. M. Park, J. W. Bae, H.-S. Do, B.-J. Lee, and H. S. Kim, “A new strategy for designing immiscible medium-entropy alloys with excellent tensile properties,” *Acta Materialia*, vol. 193, pp. 71–82, 2020.
- [132] H. Kwon et al., “Precipitation-driven metastability engineering of carbon-doped cocrfenimo medium-entropy alloys at cryogenic temperature,” *Scripta Materialia*, vol. 188, pp. 140–145, 2020.
- [133] J. Cai, X. Chu, K. Xu, H. Li, and J. Wei, “Machine learning-driven new material discovery,” *Nanoscale Advances*, vol. 2, no. 8, pp. 3115–3130, 2020.
- [134] J. E. Saal, A. O. Oliynyk, and B. Meredig, “Machine learning in materials discovery: Confirmed predictions and their underlying approaches,” *Annual Review of Materials Research*, vol. 50, pp. 49–69, 2020.
- [135] P. C. Jennings, S. Lysgaard, J. S. Hummelshøj, T. Vegge, and T. Bligaard, “Genetic algorithms for computational materials discovery accelerated by machine learning,” *NPJ Computational Materials*, vol. 5, no. 1, p. 46, 2019.
- [136] M. Sajjan et al., “Quantum machine learning for chemistry and physics,” *Chemical Society Reviews*, vol. 51, no. 15, pp. 6475–6573, 2022.
- [137] B. Kheireddine, B. Zoubida, H. Tarik, and A. Imed, “Application of pso and tlbo algorithms with neural network for optimal design of electrical machines,” *COMPEL-The international journal for computation and mathematics in electrical and electronic engineering*, vol. 37, no. 2, pp. 549–564, 2018.
- [138] K. Y. Gómez Díaz, S. E. De León Aldaco, J. Aguayo Alquicira, M. Ponce-Silva, and V. H. Olivares Peregrino, “Teaching–learning-based optimization algorithm applied in electronic engineering: A survey,” *Electronics*, vol. 11, no. 21, p. 3451, 2022.
- [139] V. Dixit, A. Kumar, N. Sharan, S. Pandey, A. Kumar, and R. S. Singh, “Optimizing intelligent reflecting surface assisted visible light communication networks under blockage and practical constraints using tlbo for iot applications,” *Scientific Reports*, vol. 15, no. 1, p. 27 400, 2025.

- [140] Sumit, R. Shukla, and A. K. Sinha, “Finite element method coupled with tlbo for shape control optimization of piezoelectric bimorph in comsol multi-physics,” *Simulation*, vol. 97, no. 9, pp. 635–644, 2021.
- [141] R. Zhang, S. Zhang, Z. He, J. Jing, and S. Sheng, “Miedema calculator: A thermodynamic platform for predicting formation enthalpies of alloys within framework of miedema’s theory,” *Computer Physics Communications*, vol. 209, pp. 58–69, 2016.
- [142] R. Rao and V. Kalyankar, “Parameters optimization of advanced machining processes using tlbo algorithm,” *EPPM, Singapore*, vol. 20, no. 20, pp. 21–31, 2011.
- [143] R. Eberhart and J. Kennedy, “Particle swarm optimization,” in *Proceedings of the IEEE international conference on neural networks*, Perth, Australia, vol. 4, 1995, pp. 1942–1948.
- [144] Z. Michalewicz, “Gas: Why do they work?” In *Genetic Algorithms+ Data Structures= Evolution Programs*, Springer, 1996, pp. 45–55.
- [145] Z. Chen, Y. Liu, Z. Yang, X. Fu, J. Tan, and X. Yang, “An enhanced teaching-learning-based optimization algorithm with self-adaptive and learning operators and its search bias towards origin,” *Swarm and Evolutionary Computation*, vol. 60, p. 100766, 2021.
- [146] Y. Ma, Y. Li, and L. Yong, “Teaching-learning-based optimization algorithm with stochastic crossover self-learning and blended learning model and its application,” *Mathematics*, vol. 12, no. 10, p. 1596, 2024.
- [147] F. Zou, L. Wang, X. Hei, D. Chen, Q. Jiang, and H. Li, “Bare-bones teaching-learning-based optimization,” *The Scientific World Journal*, vol. 2014, no. 1, p. 136920, 2014.
- [148] D. Chen, F. Zou, J. Wang, and W. Yuan, “Samcctlbo: A multi-class cooperative teaching-learning-based optimization algorithm with simulated annealing,” *Soft Computing*, vol. 20, no. 5, pp. 1921–1943, 2016.
- [149] S. Tuo, L. Yong, F. Deng, Y. Li, Y. Lin, and Q. Lu, “Hstlbo: A hybrid algorithm based on harmony search and teaching-learning-based optimization for complex high-dimensional optimization problems,” *PloS one*, vol. 12, no. 4, e0175114, 2017.

- [150] H. Singh et al., “An integrative tlbo-driven hybrid grey wolf optimizer for the efficient resolution of multi-dimensional, nonlinear engineering problems,” *Scientific Reports*, vol. 15, no. 1, p. 11 205, 2025.
- [151] P. M. de Lima and C. A. Castro, “Binary teaching-learning-based optimization: A new algorithm,” *REVISTA DELOS*, vol. 18, no. 63, e3716–e3716, 2025.
- [152] Y. Kumar and P. K. Singh, “A chaotic teaching learning based optimization algorithm for clustering problems: Y. kumar, p. kumar singh,” *Applied Intelligence*, vol. 49, no. 3, pp. 1036–1062, 2019.
- [153] N. W. Ashcroft and N. D. Mermin, “Solid state,” *Physics (New York: Holt, Rinehart and Winston) Appendix C*, vol. 1, 1976.
- [154] J. Hu et al., “Exploring a general convolutional neural network-based prediction model for critical casting diameter of metallic glasses,” *Journal of Alloys and Compounds*, vol. 947, p. 169 479, 2023.
- [155] Y. Hong, B. Hou, H. Jiang, and J. Zhang, “Machine learning and artificial neural network accelerated computational discoveries in materials science,” *Wiley Interdisciplinary Reviews: Computational Molecular Science*, vol. 10, no. 3, e1450, 2020.
- [156] K. Choudhary et al., “Recent advances and applications of deep learning methods in materials science,” *npj Computational Materials*, vol. 8, no. 1, p. 59, 2022.
- [157] S. Y.-C. Chen, T.-C. Wei, C. Zhang, H. Yu, and S. Yoo, “Quantum convolutional neural networks for high energy physics data analysis,” *Physical Review Research*, vol. 4, no. 1, p. 013 231, 2022.
- [158] Z. Fang, “A high-efficient hybrid physics-informed neural networks based on convolutional neural network,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 10, pp. 5514–5526, 2021.
- [159] S. M. Anwar, M. Majid, A. Qayyum, M. Awais, M. Alnowami, and M. K. Khan, “Medical image analysis using convolutional neural networks: A review,” *Journal of medical systems*, vol. 42, no. 11, p. 226, 2018.
- [160] D. Sarvamangala and R. V. Kulkarni, “Convolutional neural networks in medical image understanding: A survey,” *Evolutionary intelligence*, vol. 15, no. 1, pp. 1–22, 2022.

- [161] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 2002.
- [162] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep learning*. MIT press Cambridge, 2016, vol. 1.
- [163] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [164] J. T. Springenberg, A. Dosovitskiy, T. Brox, and M. Riedmiller, “Striving for simplicity: The all convolutional net,” *arXiv preprint arXiv:1412.6806*, 2014.
- [165] S. Ioffe and C. Szegedy, “Batch normalization: Accelerating deep network training by reducing internal covariate shift,” in *International conference on machine learning*, pmlr, 2015, pp. 448–456.
- [166] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [167] Y. Liu, T. Zhao, W. Ju, and S. Shi, “Materials discovery and design using machine learning,” *Journal of Materiomics*, vol. 3, no. 3, pp. 159–177, 2017.
- [168] Q. Zhang, X. Wang, Y. N. Wu, H. Zhou, and S.-C. Zhu, “Interpretable cnns for object classification,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 43, no. 10, pp. 3416–3431, 2020.
- [169] Y. Wang et al., “A physical constraint-based machine learning model for shale oil production prediction,” *Physics of Fluids*, vol. 36, no. 8, 2024.
- [170] Z. Zhang et al., “Cnn-ipinn: A method study on applying physics-informed neural networks to gear fault diagnosis,” *Measurement Science and Technology*, vol. 36, no. 10, p. 106 101, 2025.