



Faculty of Engineering

**Artificial Neural Network Algorithms to Predict the Compressive
Strength of Fly Ash-Based Geopolymers**

Teo Siaw Hui

**Master of Engineering
2025**

Artificial Neural Network Algorithms to Predict the Compressive Strength of Fly Ash-Based Geopolymers

Teo Siaw Hui

A thesis submitted

In fulfillment of the requirements for the degree of Master of Engineering

(Civil Engineering)

Faculty of Engineering
UNIVERSITI MALAYSIA SARAWAK
2025

DECLARATION

I declare that the work in this thesis was carried out in accordance with the regulations of Universiti Malaysia Sarawak. Except where due acknowledgements have been made, the work is that of the author alone. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.



.....

Signature

Name: Teo Siaw Hui

Matric No.: 21020039

Faculty of Engineering

Universiti Malaysia Sarawak

Date: August 18, 2025

ACKNOWLEDGEMENT

There are no words to adequately thank my supervisor, Associate Professor Dr. Idawati Ismail, for her invaluable patience, inputs, and assistance during my research. Without my co-supervisor, Associate Professor Dr. Tay Kai Meng, who kindly shared his knowledge and experiences, I too could not have embarked on this path. Furthermore, without the kind funding for my study from Ministry of Higher Education from the Fundamental study Grant (FRGS) Cycle 2020, this endeavour would not have been feasible.

I'm also grateful to my research group colleagues, particularly my machine learning counterpart, prospective Ph.D. candidate Mr. Chang Wui Lee, for his selfless helps with coding problems, late-night discussions, and moral support throughout the machine learning simulation. I should also express my gratitude to the university's research laboratory assistants, study participants, and CGS academic support, all of whom inspired me.

Lastly, I would be negligent if I did not mention my family, particularly my loving late father, who passed away the first year I began doing research. My mom, wife, and kids. Their believe in me has kept my spirits and enthusiasm up throughout this journey.

ABSTRACT

This study explores the application of Artificial Neural Networks (ANNs), specifically Self-Organizing Maps (SOM) and Hierarchical Self-Organizing Maps (HSOM), to predict the compressive strength of geopolymer concrete (GPC), activated using sodium hydroxide (NaOH) and sodium silicate (Na_2SiO_3). GPC samples were cast and cured at 80°C . A total of eighteen distinct mix proportions were prepared, and the resulting dataset was used to train the ANN models based on the SOM and HSOM architectures. The study aimed to correlate the effects of the water-to-binder ratio and the molarity of the alkaline solution with the compressive strength of GPC. SOM modeling offers a cost-effective, time-efficient alternative to conventional experimental methods, which are often limited by high material usage, time-consuming procedures, and challenges in capturing complex variable interactions. Model performance was assessed using five-fold cross-validation, yielding consistent results with low Root Mean Square Error (RMSE), Mean Squared Error (MSE), and Mean Absolute Error (MAE), along with a coefficient of determination (R^2) value of 0.7. A Hierarchical SOM (HSOM) approach was also evaluated to enhance clustering and interpretability. Comparative analysis confirmed the effectiveness of both SOM and HSOM in accurately predicting GPC strength. Despite promising outcomes, challenges remain due to limited data diversity and the need to improve the practical interpretability of machine learning outputs. Overall, the findings highlight the potential of SOM, HSOM, and other AI-based techniques as powerful tools for optimizing GPC mix designs and advancing sustainable construction materials.

Keywords: Geopolymer concrete (GPC); compressive strength; Artificial Neural Network (ANN); Self Organizing Map (SOM); Hierarchical SOM (HSOM)

Penggunaan Rangkaian Saraf Tiruan dalam Peramalan Kekuatan Mampatan Geopolimer Yang Berasaskan Abu Terbang

ABSTRAK

Kajian ini meneroka penggunaan Rangkaian Saraf Buatan (Artificial Neural Networks, ANN), khususnya Peta Penyusunan Sendiri (Self-Organizing Maps, SOM) dan Peta Penyusunan Sendiri Berhierarki (Hierarchical SOM, HSOM), untuk meramalkan kekuatan mampatan konkrit geopolimer (GPC) yang diaktifkan menggunakan Natrium Hidroksida (NaOH) dan Natrium Silikat (Na_2SiO_3). Sampel GPC dibentuk dan diawet pada suhu 80°C. Sebanyak lapan belas campuran yang berbeza telah disediakan, dan set data yang terhasil digunakan untuk melatih model ANN berdasarkan algoritma SOM dan HSOM. Objektif kajian ini adalah untuk mengaitkan kesan nisbah air kepada pengikat dan molariti larutan alkali terhadap kekuatan mampatan GPC. Pemodelan SOM menawarkan kaedah alternatif yang menjimatkan kos dan masa berbanding kaedah eksperimen konvensional, yang sering terhad oleh penggunaan bahan yang tinggi, proses ujian yang memakan masa, dan kesukaran dalam menangkap interaksi pemboleh ubah yang kompleks. Prestasi model dinilai menggunakan pengesahan silang lima lipatan (5-fold cross-validation), dengan keputusan yang konsisten serta nilai kesilapan kuasa dua purata (RMSE), kesilapan kuasa dua min (MSE), dan kesilapan mutlak min (MAE) yang amat rendah, bersama dengan nilai pekali penentuan (R^2) sebanyak 0.7. Pendekatan HSOM turut dinilai untuk meningkatkan keupayaan pengelompokan dan kebolehfahaman model. Analisis perbandingan mengesahkan keberkesanan kedua-dua SOM dan HSOM dalam meramalkan kekuatan GPC dengan tepat. Walaupun keputusan yang diperoleh adalah memberangsangkan, namun masih terdapat cabaran seperti keterbatasan data dan keperluan untuk meningkatkan kebolehfahaman model dalam aplikasi praktikal. Secara keseluruhan, penemuan ini

menyerlahkan potensi teknik pembelajaran mesin seperti SOM, HSOM, dan model AI lain sebagai alat yang berkesan untuk pengoptimuman reka bentuk campuran GPC dan pembangunan bahan binaan mampan.

Kata kunci: *Konkrit geopolimer (GPC); kekuatan mampatan; Rangkaian Saraf Buatan (ANN); Peta Penyusunan Sendiri (SOM); Peta Penyusunan Sendiri Berhierarki (Hierarchical SOM, HSOM)*

TABLE OF CONTENTS

	Page
DECLARATION	i
ACKNOWLEDGEMENT	ii
ABSTRACT	iii
ABSTRAK	iv
TABLE OF CONTENTS	vi
LIST OF TABLES	x
LIST OF FIGURES	xi
LIST OF ABBREVIATIONS	xiii
CHAPTER 1 INTRODUCTION	1
1.1 Study Background	1
1.1.1 Geopolymer Concrete	4
1.1.2 Artificial Intelligence	6
1.2 Problem Statement	7
1.3 Research Question	9
1.4 Objectives	9
1.5 Chapter Summary	10

CHAPTER 2 LITERATURE REVIEW	11
2.1 Overview of Geopolymer Technology Study	11
2.1.1 Importance of Fly Ash as a Precursor and Aluminosilicate Source	13
2.1.2 The Role of Machine Learning in Geopolymer Research	15
2.2 Geopolymerization Process and Mechanisms	16
2.2.1 Function of Precursor in Geopolymerization	18
2.2.2 Influence of Alkaline Activators on Geopolymer Properties	20
2.2.3 Curing Conditions and Their Impact on Strength Development	25
2.2.4 Applications of Fly Ash-Based Geopolymers	28
2.3 Trends in Machine Learning Applications for Geopolymer Research	30
2.4 Machine Learning Overview	33
2.4.1 Application of Machine Learning in Geopolymer Science	34
2.4.2 Machine Learning Algorithms in Geopolymer Science	37
2.4.3 Supervised vs. Unsupervised Learning Approaches for Geopolymer Analysis	39
2.4.4 Predicting Geopolymer Properties Using Machine Learning	44
2.5 Adoption of Machine Learning for Concrete Studies	45
2.5.1 Self-Organising Map (SOM)	47
2.5.2 Hierarchical Self-Organizing Map (HSOM)	50
2.5.3 Comparison of SOM and HSOM	51
2.5.4 SOM in Geopolymer Concrete Strength Prediction	55

2.6	Key Parameter Relationships Studied by Machine Learning	57
2.7	Challenges and Future Prospects	58
2.7.1	Limitations of Existing ML Models in Geopolymer Research	59
2.7.2	Advancing Machine Learning Applications in Geopolymer Future Research	60
2.8	Chapter Summary	61
	CHAPTER 3 METHODOLOGY	64
3.1	Overview	64
3.2	Sample Preparation	65
3.2.1	Pre-test Study	65
3.2.2	Main Design Study	66
3.3	Materials	67
3.4	Specimen Preparation	69
3.5	Compressive Strength Test	70
3.6	Machine Learning Study	71
3.6.1	SOM Algorithm	72
3.6.2	HSOM Algorithm	77
3.7	Project timeline	81
3.8	Chapter Summary	81
	CHAPTER 4 RESULTS AND DISCUSSION	82
4.1	Overview	82

4.2	Compressive Strength Results	83
4.2.1	Discussion on the Effects of Molarity	83
4.2.2	Discussion on the Effects of Water/Binder (W/B) Ratio	85
4.2.3	Discussion on the Effects of Activator Ratio	87
4.3	SOM Simulation Results	89
4.3.1	Simulation Result in 5×5 SOM Map	94
4.4	Verification and Validation	96
4.4.1	Ground Truth for Mix Proportion Verification	97
4.4.2	SOM Training Validation	98
4.4.3	Model Validation Through Hierarchical SOM Analysis	102
4.4.4	Results Comparison among SOM & HSOM	108
4.5	Chapter Summary	110
4.6	Conclusion	111
4.7	Recommendations	112
	REFERENCES	113
	APPENDICES	132

LIST OF TABLES

	Page
Table 2.1 Comparison of ML Algorithms for Predicting Compressive Strength	35
Table 2.2 Comparison of SOM and HSOM (Qu et al., 2021)	52
Table 3.1 Mix Proportion for Trial Study Design	66
Table 3.2 Mix Proportion for Main Design Study	67
Table 3.3 Chemical composition of Class F Fly ash (by XRF %)	68
Table 3.4 Chemical composition of Class F Fly ash based on ASTM 618	68
Table 3.5 Mixing Design of Geopolymer Concrete in 1:1 Activator Ratio	69
Table 4.1 Model structure of SOM at five parameter settings	90
Table 4.2 Prototype Weight vectors for 2×2	92
Table 4.3 Prototype Weight vectors for 5×5	92
Table 4.4 The 4 (2×2 map) clusters of SOM model	94
Table 4.5 The Predicted Parameters Derived from the 25 Clusters of the SOM Model	95
Table 4.6 Mix Proportion Given by Clusters of SOM Model	97

LIST OF FIGURES

	Page
Figure 1.1 Chemical Composition of OPC Clinkers	2
Figure 2.1 Structure of polysialates (Liew et al., 2016)	12
Figure 2.2 Polymerization reaction processes (Zhuang et al., 2016)	13
Figure 2.3 Geopolymerization mechanism (Abdila et al., 2022)	17
Figure 2.4 Chemical composition of alkali-activated materials (K. Chen et al., 2021)	20
Figure 2.5 (a) C-S-H gel structure (b) N-A-S-H gel structure (García et al., 2013)	23
Figure 2.6 Relationship between curing temperature and compressive strength (X. Tian et al., 2020)	27
Figure 2.7 Number of Publication for Geopolymer Study using Machine Learning	32
Figure 2.8 GPC Studies Publication based on Paste, Mortars, and Concrete form	33
Figure 2.9 The Layered Architecture of Artificial Neural Network (Rusna et al., 2022)	40
Figure 2.10 Grid representation of the neurons in SOM (Ullah et al., 2019)	48
Figure 2.11 Illustrations of Rectangular, Cylindrical, and Toroidal grids (Park et al., 2015)	48
Figure 2.12 Hierarchical Self-Organizing Map Structure(Qu et al., 2021)	51
Figure 2.13 Model Predictions vs Experimental Results Comparison	57
Figure 3.1 100*100*100mm Cast Icon Mold, Single and Tri-slots	65
Figure 3.2 Procedure of the Experimental Works	70
Figure 3.3 Example of SOM with 4 Input Vectors and 25 Output Clusters	72
Figure 3.4 The Learning of SOM Algorithm Process	72
Figure 3.5 Flowchart of a Sequential SOM (Winston et al., 2020)	76
Figure 3.6 SOM Input and Output Layers (Yorek et al., 2016)	77

Figure 3.7	The Learning of Hierarchical SOM Algorithm Process	79
Figure 3.8	Flowchart of Sequential HSOM (Qu et al., 2021)	79
Figure 3.9	HSOM Architecture Diagram	80
Figure 3.10	Gantt Chart for the Research Timeline	81
Figure 4.1	Compressive Strength as an Effects of Molarity	84
Figure 4.2	Compressive Strength as an Effects of Water to Binder (W/B) Ratio	86
Figure 4.3	Compressive Strength as an Effects of Activator Ratio	87
Figure 4.4	Neighbouring Connection in the 5x5 SOM Map	94
Figure 4.5	Ground Truth Verification Results	98
Figure 4.6	SOM Validation Training Results	99
Figure 4.7	HSOM Validation Training Results	103
Figure 4.8	HSOM results into different sub-levels (a)	104
Figure 4.9	HSOM results into different sub-levels (b)	105
Figure 4.10	HSOM results into different sub-levels (c)	106
Figure 4.11	HSOM results into different sub-levels (d)	107
Figure 4.12	Predicted Compressive Strength from SOM	108
Figure 4.13	Predicted Compressive Strength from HSOM	109

LIST OF ABBREVIATIONS

GPC	Geopolymer Concrete
OPC	Ordinary Portland Concrete
WBR	Water Binder Ratio
GGBS	Ground Granulated Blast Slag
ANN	Artificial Neural Network
SOM	Self-Organizing Map
HSOM	Hierarchical Self-Organizing Map
SVM	Support Vector Machines
RF	Random Forest
PCA	Principal Component Analysis
LDA	Linear Discriminant Analysis
SNE	Stochastic Neighbour Embedding
RL	Reinforcement Learning
DQN	Deep Q Networks
RMSE	Root Mean Square Error
MSE	Mean Square Error
MAE	Mean Absolute Error
SHAP	Shapley Additive Explanations
LIME	Local Interpretable Model-agnostic Explanations

CHAPTER 1

INTRODUCTION

1.1 Study Background

Ordinary Portland Cement (OPC) is the foundation of modern construction materials and is popular owing to its adaptability and dependence on environmental conditions. OPC exhibits the adaptability and dependability required for construction projects across a wide range of applications. In OPC, adaptability refers to the ability to accommodate a wide range of construction needs, including various mix designs, environmental challenges, and specific project requirements. OPC can be customized to satisfy precise strength, durability, and setting time requirements, making it appropriate for a variety of structural and non-structural applications. Their versatility enables engineers and builders to tailor concrete mixtures to specific project requirements, thereby resulting in optimal performance and lifetime. Dependability in OPC refers to the ability of a system to consistently supply desired performance characteristics. The OPC is subjected to extensive quality control methods during the manufacturing process to ensure consistent composition and quality from batch to batch. This uniformity ensures predictable behavior during construction, allowing for precise control over concrete placement, curing, and eventual structural soundness. As shown in Figure 1.1, the meticulous grinding of clinker, which is a mixture of calcined limestone and clay, along with the inclusion of gypsum, dependable OPC formulations provide engineers, contractors, and stakeholders with certainty that the erected infrastructure will fulfil design specifications and endure predicted loads and environmental conditions throughout its service life. The versatility and reliability of Ordinary Portland Cement

(OPC) have made it an essential material in construction projects worldwide, providing the stability and durability required to create a robust and enduring built environment.



Figure 1.1: Chemical Composition of OPC Clinkers

Owing to the calcium, silicon, aluminum, and iron oxides that make up its chemical composition, OPC has a strong compressive strength, which makes it ideally suited for a wide range of structures, from modest sidewalks to towering skyscrapers. Moreover, the fast-setting time of OPC concrete has accelerated construction or project scheduling and often guarantees on-time project completion, which has given an additional time at the remaining schedule for various kinds of testing on the environmental conditions. In the construction sector, this natural grey-colored cement is a mainstay, providing engineers and builders with reliable materials for resilient concrete structures, precast segments, and mortars. Operating under strict manufacturing guidelines, the OPC follows exact formulas to ensure uniformity and excellence in all its uses. The OPC is a perfect representation of contemporary buildings, offering a strong base on which the built environment may thrive.

Although OPC provides numerous benefits for construction, it also presents notable drawbacks that warrant consideration, including a large carbon footprint, which stems from the enormous energy consumption and CO₂ emissions involved in its manufacturing process. Impact on the environment is a key concern. Calcining limestone, a key component of cement production, requires high-temperature kiln procedures that are energy-intensive and emit considerable amounts of greenhouse gases, primarily carbon dioxide (CO₂). This has resulted in environmental degradation and accelerated climate change. OPC is responsible for large carbon dioxide emissions globally, with a 7% contribution annually and increasing year after year (Burduhos et al., 2018; Matsimbe et al., 2022). In some applications, OPC's high alkalinity of OPC may cause complications. Because of its corrosive properties owing to its high alkalinity, reinforcing steel buried within concrete may eventually deteriorate over time. Thus, the structural integrity of concrete components can be jeopardized. In addition, the manufacture of OPC is dependent on limited or unrenovable natural resources, such as clay and limestone, which raises global concerns regarding resource depletion and long-term sustainability in the construction industry. Some architectural designs may find the grey hue of OPC to be less visually appealing compared to other materials or finishes, which calls for the use of additional materials or different types of cement. Despite OPC's extensive use of OPC, these disadvantages highlight the need to investigate substitute cementitious materials and implement sustainable procedures to reduce the environmental impact of OPC use and manufacturing. Environmental influence is a primary concern in Sustainable Development Goals (SDG). The SDG is a collection of 17 interconnected global objectives agreed upon by United Nations Member States in 2015 as part of the 2030 Agenda for Sustainable Development. These goals seek to solve diverse social, economic, and environmental

concerns confronting humanity with the goal of creating a more sustainable and fair future for all, particularly for future generations.

1.1.1 Geopolymer Concrete

In recent years, a new concrete material called “soil silicate concretes” or “soil cement” was developed in the 1950s, and it has been commonly used as geopolymer concrete (GPC) in recent years. GPC represents a ground-breaking innovation in the construction industry, offering a sustainable alternative to the traditional Portland cement-based concrete. Unlike OPC, which relies on calcium silicate-based binders, geopolymer concrete utilizes aluminosilicate materials, such as fly ash, slag, or metakaolin, and is activated by alkali solutions obtained from landfills. GPC presents a compelling array of advantages over OPC, positioning it as a promising alternative that addresses both the environmental and performance considerations. The most significant advantage is its substantially reduced carbon footprint (CF)(Arioz et al., 2012; Torres et al., 2017). GPC production relies on industrial by-products such as fly ash and slag, minimizing the need for energy-intensive processes associated with limestone quarrying and clinker production, thus significantly lowering carbon dioxide emissions. This aligns with the growing emphasis on sustainability in the construction industry in the SDG to mitigate climate change issues. Additionally, GPC exhibits enhanced durability characteristics compared to OPC concrete. Its resistance to chemical corrosion, abrasion, and fire is notably superior, making it well-suited for structures subjected to aggressive environments or higher temperatures. Moreover, GPC often achieves higher compressive strength at an early age, allowing for faster construction schedules and potential cost savings. Its lower alkalinity reduces the risk of alkali-silica reaction (ASR), a common cause of concrete deterioration, thereby enhancing the longevity

and performance of structures. Another advantage stems from the utilization of industrial byproducts in GPC production, which effectively diverts these materials from landfills and promotes resource efficiency. This not only reduces waste, but also conserves natural resources. GPC can also offer improved thermal insulation properties compared to OPC concrete, contributing to energy-efficient building designs and reducing long-term operational costs. Its resistance to sulfate attacks makes it particularly suitable for applications in environments with high sulfate content, such as coastal regions or wastewater treatment facilities. Overall, the advantages of GPC over OPC extend beyond environmental sustainability to encompass enhanced durability, performance, resource efficiency, and suitability for diverse construction applications, positioning it as a compelling choice for evolving requirements of the construction industry.

Although GPC has several encouraging benefits, there are some disadvantages that should be considered. Its comparatively greater start-up cost compared with OPC concrete reveals a major obstruction. The procurement of specialized materials and alkali activators required for GPC manufacturing can result in higher initial costs, which may discourage some construction projects from utilizing GPC, especially those with limited funding. Furthermore, differences in materials and curing conditions may affect the alkali activation process of GPC, necessitating careful quality control and experience during construction to guarantee consistent performance. As a result, further research and optimization are needed to fully understand the long-term performance and durability of GPC under a variety of environmental conditions, despite breakthroughs in research and development. Their compatibility with current construction methods and infrastructure may raise issues; thus, thorough assessment and testing are required before they can be widely used. Regional differences in the availability of appropriate precursors for GPC manufacturing, such as fly

ash and slag, can present logistical difficulties in obtaining ingredients for specific projects. Maximizing the potential of GPC as a high-performance and sustainable substitute for OPC concrete in the building sector requires addressing these constraints. Sufficient investigation, creativity, and standardization are required to overcome these obstacles and broaden the use of GPC in other building scenarios.

1.1.2 Artificial Intelligence

Artificial intelligence (AI) has revolutionized many sectors, and its application in predicting geopolymer characteristics is no exception. An Artificial Neural Network (ANN) is a machine learning algorithm inspired by the structure and function of biological neural networks, specifically the human brain (Dongare et al., 2008; Holmbom et al., 2011a). ANN are designed to process and analyze data, recognize patterns, and make predictions based on given data, which includes effective analysis and extraction of valuable insights from large and complex datasets. Thus, it is widely used to understand the complex relationships between variables. They can learn from past patterns/relationships and provide accurate solutions in various fields such as scientific forecasting, stock market prediction, and engineering simulations (Abiodun et al., 2019). Geopolymers, which are inorganic polymers derived from aluminosilicate precursors, have received considerable attention because of their potential as environmentally acceptable alternatives to traditional cement. Predicting geopolymer characteristics, including concrete strength and durability, is critical for optimizing their use in construction. Artificial intelligence tools, such as machine learning algorithms and neural networks, can quickly evaluate massive amounts of data from tests and simulations to predict geopolymer properties precisely. AI models are often trained to reveal hidden patterns and correlations that regulate geopolymer concrete formation and

performance characteristics by incorporating data on the precursor ingredients, curing conditions, and environmental factors. Motivated by the use of an artificial neural network (ANN) to analyze GPC properties (Kuppusamy et al., 2022), this study seeks to extend the learning of ANN using clustering methods to reveal hidden messages or information of GPC from a collection of GPC samples.

1.2 Problem Statement

Geopolymer concrete is an eco-friendly and sustainable material that utilizes industrial by-products such as fly ash or slag to produce a binder and reduce the carbon footprint of the construction industry. However, despite their potential benefits, geopolymer concrete faces several obstacles that hinder its widespread adoption. One key issue is the lack of standardized mix design and curing procedures for geopolymer concrete. In contrast to traditional concrete, geopolymer concrete relies on various chemical reactions and curing methods. The absence of clear guidelines and protocols for mix design and curing can lead to inconsistent performance and limited confidence in the strength, durability, and long-term behavior of the material.

A significant research problem in predicting the compressive strength of fly ash-based geopolymer concrete arises from the complex and nonlinear interactions among its mix parameters and curing conditions. Factors such as fly ash composition, activator dosage, curing temperature, and curing time all influence the geopolymerization process and ultimately affect strength development. The variability in fly ash sources and the sensitivity of the geopolymerization reactions contribute to unpredictable strength outcomes. Existing empirical models and experimental methods often fail to capture these complex relationships accurately, making prediction of compressive strength challenging.

Another challenge is the limited understanding of the influence of various factors on the properties of geopolymer concrete. Parameters such as the proportion of raw material types, activator dosage, curing temperature, and curing duration play a crucial role in influencing the geopolymerization process, setting time, strength development, and overall performance of geopolymer concrete. Understanding the relationship between curing conditions and strength development is vital for establishing reliable and reproducible strength characteristics of geopolymer concrete.

An artificial neural network (ANN) is a machine learning algorithm inspired by the structure and function of biological neural networks, specifically the human brain. ANN are designed to process and analyze data, recognize patterns, and make predictions based on given data, which includes effectively analyzing and extracting valuable insights from large and complex datasets. Thus, it is widely used to understand the complex relationships between variables. They can learn from past patterns/relationships and provide accurate solutions in various fields such as scientific forecasting, stock market prediction, and engineering simulations.

In summary, this study aimed to use ANN capabilities to investigate the influence of the raw material composition of GPC. By comprehensively understanding these parameters, this study seeks to optimize the strength performance of GPC and establish it as a viable and competitive alternative to conventional concrete in the construction industry. This study involved compressive-strength data collection via multiple condition tests, data training, and validation using ANN models.

1.3 Research Question

- i. What are the factors that affecting the compressive strength of geopolymers using Class F fly ash?
- ii. How can machine-learning approaches be used to predict geopolymer compressive strength?
- iii. How accurate is the Artificial Neural Network (ANN) in predicting fly ash-based geopolymer compressive strength compared to experimental data?

1.4 Objectives

- i. Investigate the influence of key mix parameters, such as water-to-binder ratio, activator molarity, and curing conditions on the compressive strength of fly ash-based geopolymer concrete.
- ii. Correlate the relationship between water-to-binder ratio and alkaline activator molarity with the compressive strength development of geopolymer concrete.
- iii. Develop and evaluate an Artificial Neural Network (ANN) model for predicting the compressive strength of geopolymer concrete, and assess its accuracy by comparing predicted results with experimental data.

By accomplishing these objectives, this study aims to advance the understanding of GPC, optimize its properties, and promote its adoption as a sustainable and viable alternative to conventional OPC in the construction sector.

1.5 Chapter Summary

The purpose of this study aimed to reduce the time and expense usually involved with traditional testing techniques by developing a predictive method for geopolymer concrete (GPC), primarily focusing on compressive strength. Finding the important variables that affect compressive strength and comprehending how they interact were the main goals of the study. The complexity and nonlinearity of these linkages prompted the adoption of an artificial intelligence (AI) based strategy.

To identify patterns and correlations in the input parameters pertaining to GPC mix design, an Artificial Neural Network (ANN) model was created and trained. The model's predictive power for compressive strength was then assessed by contrasting its outputs with experimental findings. The ANN's performance was verified using a different test set to guarantee accuracy and dependability, showcasing its potential as a strong instrument for improving decision-making in sustainable construction methods and optimizing GPC mix design.

CHAPTER 2

LITERATURE REVIEW

2.1 Overview of Geopolymer Technology Study

As an environmentally friendly substitute for conventional Portland cement, geopolymer technology has drawn a lot of interest. Inorganic polymeric materials known as geopolymers are created when alkaline activators react with aluminosilicate precursors. An aluminosilicate gel network created during polymerization gives geopolymers their strength, in contrast to traditional cement, which depends on calcium silicate hydrate (C-S-H) gel (Davidovits, 1994). Superior mechanical qualities, reduced carbon emissions, and increased resistance to chemical and thermal degradation are the outcomes of this unique chemistry (Provis et al., 2009). The concept of eco-friendly geopolymers was first introduced by Davidovits in 1976, using alkaline aluminosilicate binders. The transformation of kaolinite into tridimensional tecto-aluminosilicates, which is utilized in the polycondensation of organic resins at low temperatures, follows a process analogous to the thermosetting method of these materials. This process yielded a nanocomposite with characteristics similar to those of an artificial stone (Burduhos et al., 2018). Geopolymers are inorganic polymeric materials with unique properties and functionalities that complement those of traditional organic polymers, thereby expanding their applications in industries, such as electronics, aerospace, healthcare, and environmental remediation. They can be obtained by mixing a dry solid (aluminosilicate) with an alkaline solution and other components. The major constituent is the source material that rich in silicone (Si) and aluminum (Al), they could be natural minerals such as fly ashes, slags, etc. known as “waste” materials. The liquid is usually composed of sodium (hydroxide or silicate) or potassium (hydroxide or silicate), which are

the soluble alkali metals. Polycondensation reactions that form geopolymers occur when inorganic aluminosilicate-based materials react with an alkaline solution. The geopolymer structure can be divided into three basic forms depending on the silicone and aluminum Si/Al ratios. The three basic units, poly (sialate), poly (sialate siloxo), and poly (sialate disiloxo), are shown in Figure 2.1.

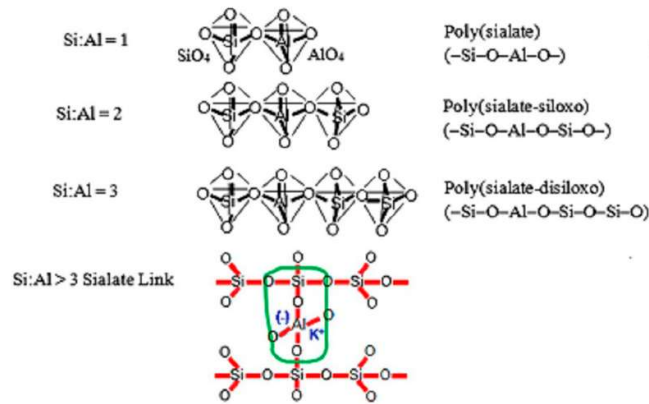


Figure 2.1: Structure of polysialates (Liew et al., 2016)

Generally, geopolymers are synthetic inorganic polymers, and a three-dimensional polymeric chain structure is developed during chemical reactions under alkaline conditions. The composition of the raw materials and the concentrations of the alkaline solutions control the microstructure and mechanical properties of the geopolymers. The significance of geopolymer binders in replacing OPC binders is based on the fact that large amounts of waste materials containing aluminosilicates are obtained from industries and agriculture, resulting in problems with their disposal.

Geopolymers are typically synthesized by mixing source materials with aluminosilicate and alkaline solutions. The source materials were kaolinite, clay, zeolite, fly ash, silica fume, slag, POFA, rice husk ash, and red mud. The most common alkaline liquid

used in geopolymerization is a combination of NaOH/KOH and sodium silicate. This dissolution step releases reactive species, including silicate and aluminate ions, which undergo polymerization reactions to form amorphous three-dimensional networks. When any of the above source materials (e.g., fly ash (FA) in solid form) are mixed with alkali solutions at appropriate concentrations and sodium silicate, condensation reactions occur, where they link to form Si-O-Si and Al-O-Si bonds, creating a cross-linked network structure. Polymerization is typically carried out at elevated temperatures, usually around 60-90°C, to accelerate the reaction kinetics and promote the formation of a dense and durable geopolymer matrix. The polymerization reaction processes for fly ash are shown in Figure 2.2.

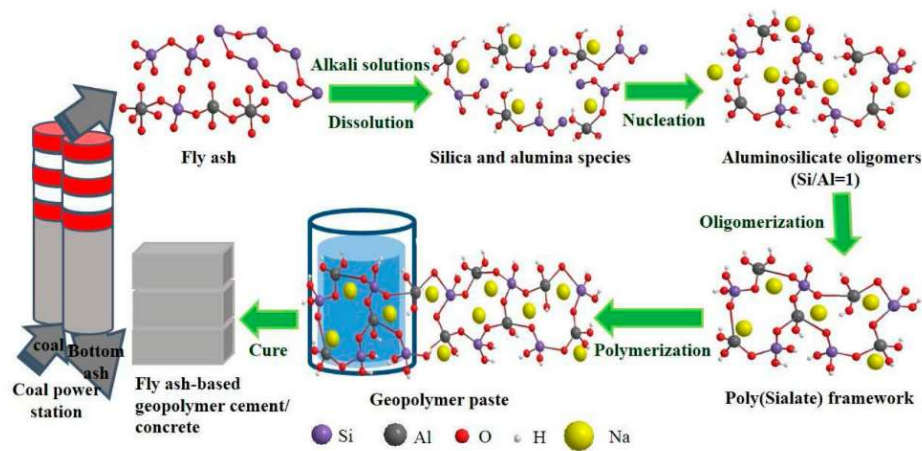


Figure 2.2: Polymerization reaction processes (Zhuang et al., 2016)

2.1.1 Importance of Fly Ash as a Precursor and Aluminosilicate Source

The high concentration of reactive silica and alumina in fly ash, a byproduct of burning coal in power plants, makes it one of the most popular precursors in geopolymer technology. These elements are crucial for the geopolymerization process (Fernández et al.

2003). Class F fly ash is especially beneficial since it improves mechanical, chemical, and long-term durability due to its low calcium concentration. Fly ash's dissolving rate and reactivity are greatly influenced by its fineness, chemical makeup, and mineralogical makeup; these factors ultimately decide how well the resultant geopolymer performs.

Fly ash has several advantages when used in geopolymer concrete (GPC), which makes it a very beneficial ingredient for environmentally friendly building. Improved mechanical strength is one of its main benefits as the development of a thick aluminosilicate matrix promotes strong compressive strength and structural integrity. Furthermore, by decreasing porosity, fly ash improves durability, offering remarkable fire resistance and improved resilience to extreme environmental conditions including sulfate and acid assaults (Davidovits, 2011).

From an environmental perspective, fly ash promotes sustainability by reducing carbon emissions and addressing waste management challenges, as its use diverts it from landfills and repurposes it as a valuable resource (Duxson et al., 2007). Fly ash is also a cost-effective element that enhances workability and long-term strength development in geopolymer concrete, reducing manufacturing costs and making it a sustainable and eco-friendly substitute for conventional cement-based products.

The mineralogical makeup of fly ash has a major role in determining its reactivity. While crystalline phases like mullite and quartz can restrict reactivity, a significant fraction of amorphous silica and alumina facilitates successful geopolymerization (Van Jaarsveld et al., 2002). Achieving desired geopolymer performance requires optimizing these qualities through the selection of precursors, curing conditions, and mix design (Yip et al., 2005). Furthermore, fly ash's chemical makeup including its fineness and reactivity influences

important metrics like durability and compressive strength and is crucial to the geopolymerization process and the integrity of the finished matrix.

2.1.2 The Role of Machine Learning in Geopolymer Research

In geopolymer research, machine learning (ML) has become a game-changing technique that provides data-driven answers to challenging problems in process optimization, property prediction, and material design (Ngo et al., 2025). In geopolymer synthesis, conventional experimental methods are frequently laborious, resource-intensive, and constrained in their capacity to investigate the wide range of parameter space pertaining to precursor compositions, alkaline activators, and curing conditions. On the other hand, machine learning speeds up the research and development process by allowing academics to examine big information, find hidden patterns, and create predictive models.

Predicting material qualities, such compressive strength, based on input factors like precursor composition, alkaline activator ratios, curing conditions, and aging time is one of the main uses of machine learning in geopolymer research. These interactions have been very well modeled by methods like Self-Organizing Maps (SOM), Artificial Neural Networks (ANN), and other machine learning techniques (Ahmari et al., 2013; Awoyera et al., 2020). In addition to eliminating the need for lengthy experimental trials, these predictive models make it possible to optimize geopolymer compositions more precisely and efficiently.

Beyond predicting properties, machine learning is essential to improving geopolymer materials' sustainability. ML algorithms can evaluate the lifespan performance and environmental effect of geopolymers, including information on their energy usage, carbon footprint, and long-term durability. Furthermore, thorough microstructural analysis is made

possible by the combination of machine learning with sophisticated characterisation methods including spectroscopy and image processing. This makes it possible for scientists to investigate phase formations, mechanisms of degradation, and the geopolymer matrix's development over time (Khale et al., 2007; Provis, 2018).

New applications in engineering and construction are being made possible by the collaboration between machine learning and geopolymer research. Researchers may create geopolymers with specialized qualities for certain uses, such fire-resistant coatings, high-strength concrete, and eco-friendly construction materials, by utilizing machine learning driven insights. This improves geopolymers' performance and encourages their use as a sustainable substitute for conventional cement-based materials.

In conclusion, by offering strong instruments for sustainability evaluation, optimization, and predictive modeling, machine learning is transforming the study of geopolymers. A more robust and sustainable constructed environment might result from the creation and use of geopolymers, which could be made possible by the ongoing advancement of machine learning methods.

2.2 Geopolymerization Process and Mechanisms

An intricate chemical process called geopolymerization turns aluminosilicate precursors like fly ash, metakaolin, or slag into a strong and environmentally friendly binder. This procedure creates a three-dimensional aluminosilicate network by dissolving silica and alumina from the precursor in a highly alkaline solution, then poly-condensing these dissolved species (Davidovits, 1994). Excellent mechanical, chemical, and thermal stability qualities make the resultant geopolymer matrix a viable substitute for conventional Portland cement.

The three main steps of the geopolymerization process are (1) dissolution, in which the aluminosilicate structure of the precursor is broken down by the alkaline solution, releasing reactive silica and alumina into the solution; (2) gelation, in which the dissolved species reorganize into a gel-like phase; and (3) polycondensation, in which the gel solidifies into a stiff, cross-linked aluminosilicate network (Duxson et al., 2007). The kind and concentration of the alkaline activator, the curing conditions, and the chemical makeup of the precursor all affect the kinetics and scope of these phases. Figure 2.3 below showing the Geopolymerization mechanism for better illustration purposes.

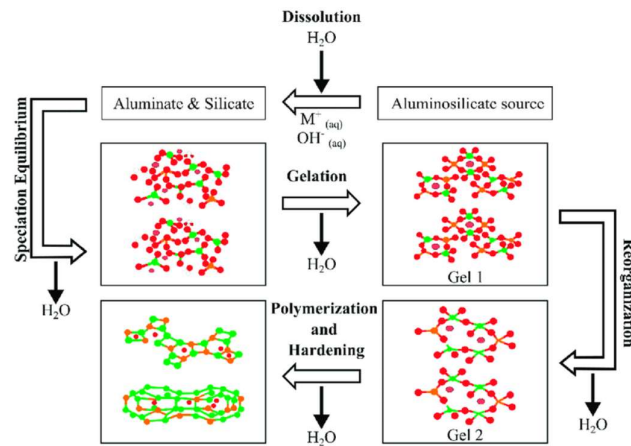


Figure 2.3: Geopolymerization mechanism (Abdila et al., 2022)

Optimizing the performance of geopolymer materials requires a thorough understanding of the geopolymerization process and its underlying principles. Researchers may modify the characteristics of geopolymers for particular uses, such as ecologically friendly coatings or high-strength building materials, by regulating these variables. The basic ideas of geopolymerization, the function of important reactants, and the variables influencing the creation and characteristics of geopolymers are all covered in this part.

2.2.1 Function of Precursor in Geopolymerization

The precursor is an essential part of the geopolymerization process since it is the main source of alumina (AlO_3) and silica (SiO_2), which are needed to construct the aluminosilicate network. The chemical makeup, mineralogical structure, and physical characteristics of the precursor which is usually made from natural minerals or industrial by-products like fly ash, metakaolin, ground granulated blast furnace slag (GGBS), or rice husk ash have a big impact on the geopolymerization process and the properties of the finished product. The chemical makeup is very important since the precursor has to have enough reactive silica and alumina to take part in the reaction. The ratio of SiO_2 to Al_2O_3 is crucial in establishing the geopolymer's mechanical characteristics and degree of polymerization (Khale et al., 2007). For instance, fly ash that contains a lot of amorphous silica and alumina is very reactive and makes a great precursor. Similarly, metakaolin, which is made by calcining kaolinite, is also very reactive because of its amorphous structure. Fly ash's crystalline phases, such as mullite or quartz, are less reactive (Xu et al., 2000).

The reactivity of the precursor is also influenced by its mineralogical structure; amorphous phases dissolve more easily in alkaline solutions than crystalline phases do (Criado et al., 2007; Garcia et al., 2015). Particle size distribution and fineness are important physical characteristics as well since finer particles have a higher surface area, which improves interaction with the alkaline activator and speeds up the geopolymerization process. According to studies, adding finely ground fly ash or slag to geopolymer concrete enhances its workability and strength development. The structure of the precursor is broken down by the alkaline solution during the dissolving step of geopolymerization, releasing silica and alumina ions into the solution (Yip et al., 2005). After that, these ions proceed through polycondensation to create the geopolymer matrix, a three-dimensional

aluminosilicate network (Torres et al., 2017). The precursor's properties directly influence the kinetics of dissolution, gelation, and polycondensation, ultimately determining the geopolymer's mechanical strength, durability, and microstructure.

In addition to increasing the sustainability of geopolymers, the use of industrial byproducts as precursors, such as fly ash and slag, lowers the carbon footprint of building materials. The construction sector may drastically reduce CO₂ emissions and encourage more environmentally friendly building techniques by using geopolymers made from these precursors for conventional cement. The uses of geopolymers are numerous and include trash encapsulation, fire-resistant coatings, and high-strength building materials (Provis et al., 2014; Zhang et al., 2014). For instance, Class F fly ash is perfect for alkali-activated materials (AAM) and improves long-term durability because of its high aluminosilicate concentration and low calcium content, whereas metakaolin is commonly utilized because of its high alumina content and reactivity. With a greater CaO concentration, class C fly ash speeds up the development of early strength but may prevent the formation of geopolymer gel. A by-product of the manufacturing of iron, ground granulated blast furnace slag (GGBFS), is activated by alkali to produce calcium aluminosilicate hydrate (C-A-S-H) gels. Because of its high calcium content, tobermorite and calcium carbonate (CaCO₃) are formed, which improves the compressive strength and chemical bond structure of the geopolymer (Aziz et al., 2020).

Fly ash, a byproduct of burning coal, has a high aluminosilicate concentration and is used as a major precursor in geopolymerization. Its composition controls the development of final strength and reaction kinetics, especially the SiO₂/AlO₃ ratio. Class C fly ash has more CaO and speeds up early strength increase, but it can also obstruct the production of

geopolymer gel. In contrast, Class F fly ash has less calcium and promotes improved long-term durability (Van Jaarsveld et al., 2002).

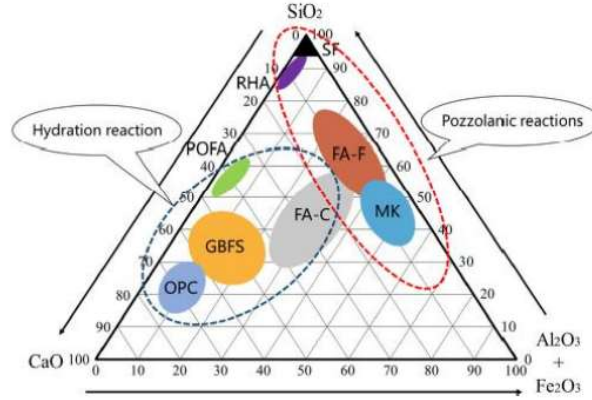


Figure 2.4: Chemical composition of alkali-activated materials (K. Chen et al., 2021)

To sum up, the precursor is crucial for the process of geopolymerization since it supplies the raw elements required for the chemical reaction and affects the final characteristics of the geopolymer. To create high-performance geopolymers suited for particular uses, precursors must be chosen and optimized according to their chemical makeup, mineralogical structure, and physical characteristics. Geopolymers, which use industrial wastes like fly ash and slag, provide a sustainable substitute for conventional cement, lowering environmental impact and encouraging more environmentally friendly building methods.

2.2.2 Influence of Alkaline Activators on Geopolymer Properties

In the geopolymerization process, alkaline activators play a vital role because they dissolve aluminosilicate precursors and promote the development of a strong aluminosilicate network. The kind, amount, and proportion of the activator to the precursor have a big impact

on the mechanical characteristics, workability, setting time, and durability of the geopolymer. With unique benefits and drawbacks, sodium hydroxide (NaOH), sodium silicate (Na_2SiO_3), and sodium carbonate (Na_2CO_3) are often used activators. NaOH is used extensively because it may speed up the geopolymerization process and make fly ash more soluble. Higher NaOH concentrations, however, have the potential to compromise the integrity of the material by causing excessive shrinkage and quick setting, even if they can increase compressive strength (Provis et al., 2009). By adding soluble silica to the process, sodium silicate which is frequently employed in conjunction with NaOH promotes gel formation and improves structural stability. Increased compressive strength, less permeability, and quicker polymerization rates are the outcomes of this (Duxson et al., 2007). However, especially at higher Na_2SiO_3 to NaOH ratios, its high viscosity might decrease the setting time and limit workability (Khattab et al., 2023). Sodium carbonate, on the other hand, functions as a gentler activator and provides slower reaction kinetics, which enhance durability and workability. In order to regulate the setting process and lessen problems like efflorescence and microcracking, it is commonly used with NaOH (Khale et al., 2007).

One important element influencing the viscosity, setting time, and mechanical characteristics of the geopolymer is the ratio of Na_2SiO_3 to NaOH. A cross-linked aluminosilicate structure is encouraged by a greater silicate concentration, usually 1.5 to 2.5, which improves durability and compressive strength. On the other hand, too much NaOH can result in efflorescence and microcracking, while too high a silicate content might decrease workability (Chavan et al., 2017). Potassium-based substitutes such potassium hydroxide (KOH) and potassium silicate have been investigated as alternatives to sodium-based activators. These activators provide enhanced alkalinity and better solubility of

silicate species, making them especially useful for precursors with a high silica concentration. For Class F fly ash, sodium-based activators continue to be more effective, even if potassium-based activators offer a good substitute in some situations (Burduhos et al., 2018).

In order to lessen the negative environmental effects of geopolymer synthesis, recent developments have concentrated on waste-derived and sustainable activators. Alternatives to conventional chemical activators have been studied, including incinerator ash, recycled glass, and solid potassium carbonate. These environmentally friendly choices affect the production of gel phases such as calcium aluminosilicate hydrate (C-A-S-H) and sodium aluminosilicate hydrate (N-A-S-H), contingent on the calcium content of the precursor (Luukkonen et al., 2019). Utilizing activators made from waste promotes the recycling of industrial byproducts and lessens dependency on traditional chemicals, which is consistent with the ideas of the circular economy. As an illustration of the significance of activator selection in maximizing geopolymer performance, metakaolin-based geopolymers activated by sodium waterglass have shown exceptional compressive strength (Tchakouté et al., 2017).

In a nutshell the characteristics of geopolymers are greatly influenced by the selection and mix of alkaline activators. Although sodium-based activators are still the most often utilized, waste-derived and potassium-based alternatives provide encouraging avenues for the creation of sustainable geopolymers. It takes careful activator type, concentration, and ratio optimization to strike the ideal balance between mechanical performance, workability, and environmental sustainability. The geopolymer industry may move closer to more environmentally friendly and effective building materials by utilizing sustainable activators and adjusting their application to certain precursors.

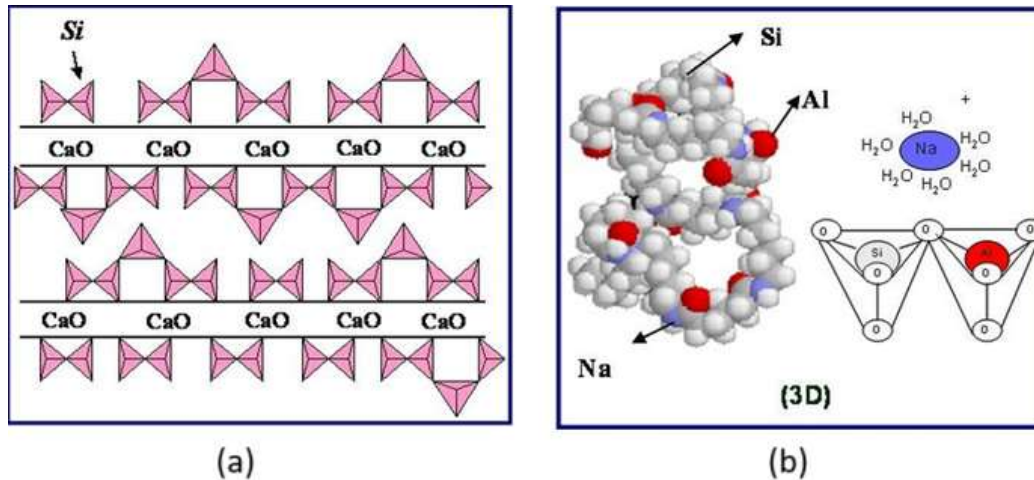


Figure 2.5: (a) C-S-H gel structure (b) N-A-S-H gel structure (García et al., 2013)

2.2.2.1 Types of Alkaline Activators (Sodium-Based vs. Potassium-Based)

The selection of potassium- or sodium-based activators has a major impact on the geopolymerization process, influencing the final geopolymer's durability, setting time, and mechanical characteristics. Because they are readily available and reasonably priced, sodium-based activators such sodium hydroxide (NaOH) and sodium silicate (Na_2SiO_3) are frequently utilized (Brough et al., 2002). They encourage the development of dense microstructures, which raises stability over time and compressive strength. Workability and structural integrity may be impacted by sodium-based systems' tendency to require longer setting periods and more shrinkage.

On the other hand, potassium-based activators, such as potassium hydroxide (KOH) and potassium silicate (K_2SiO_3), have faster reaction kinetics, which leads to greater early strength development and faster setting times. Potassium's higher ionic radius enhances ion mobility, which speeds up polymerization and affects the structural development of the

geopolymer. Potassium-based systems, however, can be more sensitive to temperature changes and need further tweaks to avoid certain problems such drying shrinkage.

The ideal balance between workability, setting time, and durability determines which activators potassium- or sodium-based should be used. Combining the two activators can sometimes maximize performance by utilizing their individual advantages to provide better mechanical characteristics and long-term stability (Lizcano et al., 2012). This hybrid strategy overcomes the drawbacks of separate activators and enables customized geopolymer formulations that satisfy particular application needs.

2.2.2.2 Effects of Activator Molarity and Ratio on Geopolymer Properties

The effects of activator molarity and ratio are including compressive strength, setting time, and microstructure. The ratio of sodium hydroxide (NaOH) to sodium silicate (Na_2SiO_3) and the molarity of the alkaline activator are important variables that affect the compressive strength and setting time of geopolymers. The concentration of hydroxide ions available to dissolve aluminosilicate precursors is determined by the molarity of the alkaline solution, especially NaOH. Increased molarity promotes precursor dissolution, increases compressive strength, and quickens the geopolymerization reaction. However, due to shrinkage and thermal stress, an overly high molarity might cause quick setting, decreased workability, and an increased risk of microcracking (Chandra et al., 2017). In order to balance reactivity and workability, molarity optimization is crucial.

Another important factor influencing the characteristics of geopolymers is the ratio of Na_2SiO_3 to NaOH. By providing soluble silica, sodium silicate aids in the development of a dense network of aluminosilicate. By fortifying the geopolymer matrix through better cross-linking, a greater silicate content typically between 1.5 and 2.5 improves compressive

strength and durability (Pavithra et al., 2016). But raising the silicate ratio also makes the slurry more viscous, which can speed up setting and decrease workability. On the other hand, inadequate polymerization, delayed strength development, and efflorescence can be caused by lower silicate ratios or by using too much NaOH.

The mechanical performance and setting behavior of geopolymers are directly impacted by the interaction between molarity and activator ratio. A greater Na_2SiO_3 to NaOH ratio, for example, speeds up setting but may need to be adjusted to preserve workability. On the other hand, lower ratios or lower molarity might prolong the setting period, which could postpone the development of strength but may be useful in some situations. According to studies, by adjusting these characteristics, geopolymers with certain qualities, such strong early strength, prolonged workability, or improved durability, may be created (Abdullah et al., 2011).

In conclusion, two important variables that control the compressive strength and setting time of geopolymers are the molarity of the alkaline activator and the ratio of Na_2SiO_3 to NaOH. Researchers and engineers may improve the geopolymerization process to obtain desired performance characteristics by carefully balancing these factors. This ensures that the material is suitable for certain industrial and construction applications.

2.2.3 Curing Conditions and Their Impact on Strength Development

The creation of microstructures, reaction kinetics, and the ultimate mechanical characteristics of geopolymers are all greatly impacted by the curing conditions too, which are crucial to the geopolymerization process. It is well known that high curing temperatures, usually between 60°C and 90°C , accelerate the geopolymerization reaction and cause strength to grow quickly. A thick silicate hydrate (C-S-H) gel is formed at these

temperatures due to the improved dissolution of aluminosilicate precursors, which is a major factor in the material's strength (Huseien et al., 2016). According to studies, fly ash-based geopolymers can have their compressive strength increased by up to 70% by heat-curing them for three to four hours at 60 to 90°C. However, heat stress and moisture loss at temperatures above 90°C might result in microcracking and decreased strength (Lim et al., 2018). Because it guarantees effective geopolymerization without negative consequences, the ideal curing temperature range for the majority of geopolymers is therefore usually accepted to be between 60 and 90°C. The curing time is also crucial; fly ash-based geopolymers are frequently cured for 6–24 hours at 60°C, 80°C, or even 100°C. Longer curing durations and higher temperatures can promote the development of early strength, but they must be properly regulated to prevent thermal damage or excessive shrinkage. Given that ambient circumstances frequently lead to partial polymerization of precursors and activator solutions, heat-cured geopolymers generally have better compressive strengths than their ambient-cured equivalents. This is a problem for applications involving in-situ construction, as ambient curing can result in longer setting periods and reduced material strength (Yip et al., 2005).

Continuous microstructure densification and polymerization affect the long-term strength development of geopolymers. Particles of unreacted fly ash may serve as nucleation sites for the development of new gel phases, resulting in long-term strength increase. However, exact control over material compositions and curing conditions is necessary to provide exceptional long-term performance. For example, ambient curing may be made possible by the use of high-reactivity precursors and customized activator formulations, which eliminates the requirement for high temperatures (Hadi et al., 2019; Khalil et al., 2020). The significance of curing temperature is highlighted by research of Tian et al.,

(2020), which showed that the fastest geopolymerization reaction and the maximum compressive strength were obtained by curing fly ash and Chinese copper tailings-based geopolymers at 80°C for 48 hours. The study found that when the temperature increased from 25°C to 80°C, the compressive strength gradually increased by around 25 MPa. To illustrate the negative consequences of very high curing temperatures, the strength dropped by around 12 MPa when the temperature was raised to 120°C (see Figure 2.6).

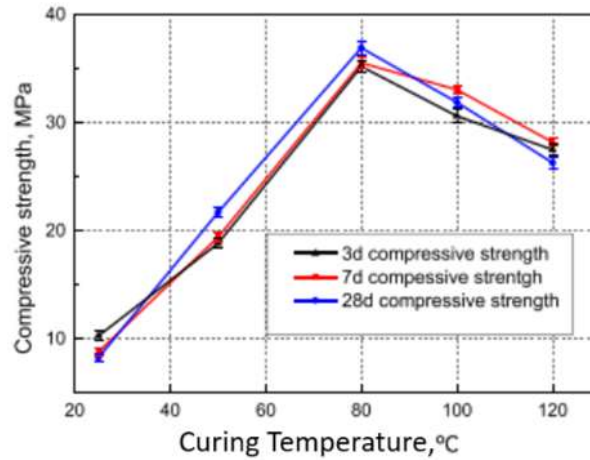


Figure 2.6: Relationship between curing temperature and compressive strength (X. Tian et al., 2020)

To summarize, the strength and durability of geopolymers are greatly influenced by the curing conditions, especially temperature and length. High-reactivity materials and well-formulated products can accomplish ambient curing, even if heat curing at 60 to 90°C is frequently chosen for quick strength growth. Optimizing geopolymer performance requires striking a balance between these variables, particularly in in-situ building applications where the ambient conditions must be taken into account.

2.2.4 Applications of Fly Ash-Based Geopolymers

The remarkable properties of geopolymers, such as their high mechanical strength, chemical resistance, and thermal stability, make them perfect for a range of industrial and environmental applications. In particular, fly ash-based geopolymers have drawn a lot of interest since they may use industrial waste materials and perform better than conventional materials. These characteristics allow them to be used in a variety of contexts, from cutting-edge industrial applications to environmentally friendly building. Geopolymers are generally acknowledged in the building sector as a more ecologically friendly substitute for traditional Portland cement concrete. Geopolymer concrete is a major contributor to lowering the carbon footprint of the building sector since its manufacture produces up to 80% fewer carbon emissions than that of regular Portland cement. Because of its outstanding resistance to chemical assaults, high compressive strength, and longevity, fly ash-based geopolymer concrete is perfect for environmentally friendly infrastructure projects including roads, bridges, and marine constructions. Incorporating industrial waste products such as fly ash and slag not only lessens reliance on landfills but also helps save resources. Further increasing its attraction for long-lasting infrastructure is the fact that geopolymer concrete exhibits higher resistance to sulfate attack and alkali-silica reaction, which are frequent problems with ordinary concrete (Davidovits, 2011; Provis et al., 2009b) .

Beyond construction sector, geopolymers are ideal for high-temperature and refractory applications also due to their exceptional thermal stability. Geopolymers based on fly ash are perfect for use in furnace linings, kiln linings, and foundry applications since they can tolerate temperatures of up to 1200°C without suffering appreciable deterioration. Geopolymers have superior resilience to thermal stress and retain their structural integrity at high temperatures, in contrast to conventional refractories. They are therefore a better option

for industrial settings where materials are exposed to high temperatures. Additionally, geopolymers are adaptable to certain mechanical and thermal needs, providing a flexible option for cutting-edge industrial applications (Duxson et al., 2007). Additionally, geopolymers are essential for waste management and environmental conservation. They effectively immobilize hazardous and radioactive waste and stop dangerous compounds from leaking into the environment because of their stable, chemically linked structure. Heavy metals, nuclear waste, and other hazardous compounds have been effectively encapsulated using fly ash-based geopolymers, offering a secure and environmentally friendly long-term waste storage option. When it comes to solving environmental issues related to industrial waste and polluted locations, this application is quite helpful. According to studies, geopolymers are a dependable material for environmental remediation since they can immobilize certain heavy metals with efficiencies of over 95% (K. Chen et al., 2021; Zhuang et al., 2016).

Geopolymers are a great option for applications needing increased safety and protection too because of their fire-resistant qualities. Geopolymers are incredibly fire resistant and can tolerate temperatures beyond 1000°C without suffering major structural damage, in contrast to materials based on Portland cement, which may break down quickly in hot conditions. Their application in the creation of fire-resistant panels, coatings, and structural elements for residential, commercial, and industrial buildings is a result of this characteristic. In addition to increasing building safety, geopolymer-based fireproof materials help lower losses and damage from fire. They are a useful material for improving fire safety in construction because of their capacity to retain their integrity in the face of intense heat (Zhao et al., 2021). In addition to these proven applications, a number of new ones are being investigated for fly ash-based geopolymers. Because of their quick setting

time and great strength, they may be used in additive manufacturing techniques, such as 3D printing for building. Additionally, geopolymers are being researched for application as adsorbents in water purification, corrosion-resistant coatings, and lightweight composites. Innovation in materials science and engineering is made possible by their adaptability and adjustable qualities (Mehta et al., 2017; Panda et al., 2018).

In short, the unique features of fly ash-based geopolymers make them suitable for a wide range of applications, including sustainable building, high-temperature refractories, environmental remediation, and fire-resistant materials. Geopolymers' capacity to use industrial byproducts, together with their excellent performance and environmental advantages, positions them as a critical material for producing robust and sustainable solutions across a variety of sectors. As research and technology develop, geopolymers' potential uses are likely to grow, reinforcing their position in tackling global concerns such as sustainability, safety, and resource efficiency.

2.3 Trends in Machine Learning Applications for Geopolymer Research

The use of machine learning (ML) in geopolymer research has grown significantly over the last ten years (2014–2024), especially in fields like compressive strength prediction, mix design optimization, and material behavior analysis. Studies combining machine learning (ML) with geopolymer concrete are steadily growing, according to a bibliometric examination of academic papers from databases such as Scopus, Web of Science, ScienceDirect, Springer Nature, and IEEE. The increased interest in using artificial intelligence (AI) to create sustainable building materials is reflected in this spike in research effort. Large experimental datasets, improvements in machine learning algorithms, and the

requirement for precise prediction models to take the place of conventional empirical approaches are some of the main causes propelling this movement.

As shown in Figure 2.7, publication patterns show that ML applications in geopolymer research significantly increased after 2018, which corresponded with the wider acceptance of AI in materials science. Self-organizing maps (SOMs), support vector machines (SVMs), and artificial neural networks (ANNs) are some of the most well-known machine learning approaches that are frequently used to forecast durability and strength characteristics. In order to enable more effective and data-driven approaches to geopolymer design, these techniques have proven their capacity to reveal intricate correlations between input factors (such as mix proportions, curing conditions) and output attributes (such as compressive strength, setting time).

There are still a number of gaps in the literature despite these advancements. For example, there is presently little use of sophisticated machine learning approaches in geopolymer research, such as deep learning and hybrid models. These gaps show that further research is necessary, especially in order to create more complex prediction frameworks. By putting forward an ANN-SOM hybrid model to forecast the compressive strength of geopolymer concrete, this work fills one such gap and advances the continuous endeavors to use AI for sustainable material innovation.

Using reputable academic databases such as Scopus, Web of Science (WoS), Science Direct, IEEE Xplore, and Springer Nature Link, a systematic literature analysis was carried out to examine the use of machine learning (ML) in geopolymer research from 2014 to 2024. Because of their broad coverage of peer-reviewed journal articles, conference proceedings, and research papers in the fields of materials science, artificial intelligence, and civil

engineering, these databases were chosen. Several keywords were included in the search strategy, including "data-driven modeling in geopolymer research," "artificial intelligence in geopolymer materials," "machine learning in geopolymer concrete," and "compressive strength prediction using ML." The search results were filtered using Boolean operators (AND, OR), guaranteeing a thorough but targeted dataset.

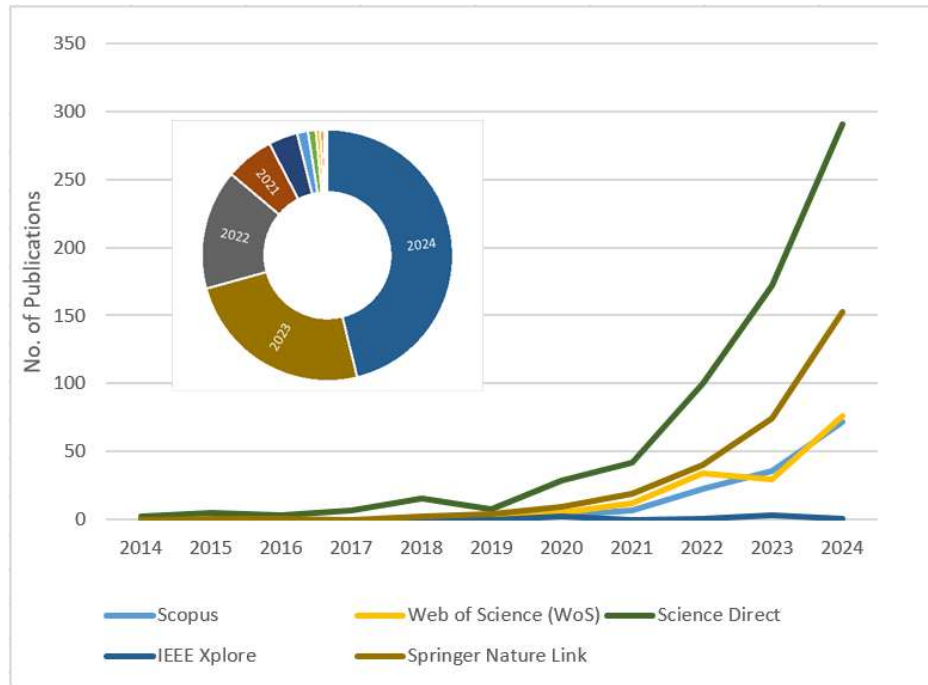


Figure 2.7: Number of Publication for Geopolymer Study using Machine Learning

Several inclusion and exclusion criteria were used to guarantee the information gathered was accurate and pertinent. Peer-reviewed conference papers, book chapters, and journal articles published between January 2014 and February 2024 were the only ones taken into account. Excluded studies were those that only used conventional statistical techniques or experimental geopolymer research devoid of machine learning applications. Moreover, redundant articles from other databases were eliminated to avoid duplication.

Besides, several studies have looked at geopolymers in a variety of forms, such as paste, mortar, and concrete, each of which has unique qualities and uses. The main areas of study for geopolymer paste include binder chemistry, microstructural development, and reaction processes. Workability, bonding properties, and strength development with fine particles are all examined in mortar investigations. The most researched type of concrete, geopolymer concrete, assesses sustainability, durability, and large-scale structural performance for building applications. Comparative evaluations of these forms aid in mix design optimization and geopolymer formulation customization for particular engineering requirements.

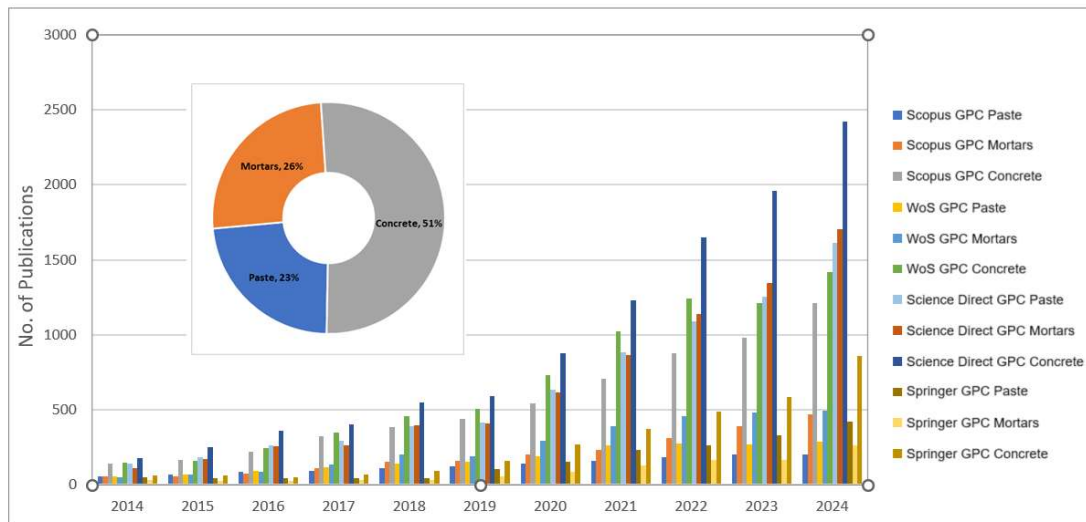


Figure 2.8: GPC Studies Publication based on Paste, Mortars, and Concrete form

2.4 Machine Learning Overview

Machine learning (ML) is a subfield of artificial intelligence (AI) that focuses on developing algorithms that enable computers to learn from and make predictions or judgments based on data. Instead of being expressly built for a specific activity, machine

learning systems improve over time as they are exposed to more data. Machine learning is frequently employed in a variety of applications, including concrete studies.

2.4.1 Application of Machine Learning in Geopolymer Science

In geopolymer science, machine learning (ML) has become a game-changing technique that provides creative ways to anticipate attributes, improve material design, and speed up research and development. Known for their exceptional mechanical qualities and sustainability, geopolymers are a complex system that depends on a variety of factors, such as the mix design, curing circumstances, alkaline activator type, and precursor composition. Understanding these linkages using traditional experimental methods frequently requires a significant investment of time and resources. However, machine learning makes it possible for academics to examine enormous datasets, spot trends, and create prediction models that expedite the optimization procedure.

Predicting properties is one of the main uses of machine learning in geopolymer research. Based on input factors like precursor type, activator concentration, and curing conditions, researchers may use experimental data to train algorithms that predict important qualities like compressive strength, setting time, and durability. In modeling these interactions, methods like Random Forest (RF), Support Vector Machines (SVMs), and Artificial Neural Networks (ANNs) have shown impressive accuracy. For instance, fly ash-based geopolymers' compressive strength may be accurately predicted using ANNs, negating the need for lengthy experimental testing(R. Kumar et al., 2024). Similarly, by determining the best mix of activators and precursors to provide desired performance characteristics, ML models may improve mix designs.

Recent studies have shown that ANN, in particular, demonstrates strong predictive power due to its ability to model complex nonlinear relationships. For example, Mahameid et al. (2024) achieved an R^2 of 0.932 when using ANN to predict the compressive strength of fly ash class-F and GGBS geopolymer source concrete, reinforcing the relevance of ANN for similar materials. SVM has also been applied effectively, particularly in smaller or well-structured datasets. Zhang et al. (2021) developed a hybrid SVM model optimized with a cuckoo search algorithm, which achieved high accuracy in predicting the compressive strength of fly ash-based geopolymer concrete. While Decision Trees (DT) are praised for their simplicity and interpretability, their performance can be limited. Ahmad et al. (2022) found that DT models reached an R^2 of approximately 0.90. In contrast, ensemble methods like Random Forest significantly improve performance by aggregating multiple trees. Both Zhang et al. (2021) and Parhi et al. (2023) demonstrated that RF achieved R^2 values up to 0.95, highlighting its ability to handle complex and nonlinear interactions within GPC datasets. Table 2.1 below shows comparison of machine learning algorithm for predicting compressive strength of fly ash based geopolymer concrete in various studies.

Table 2.1: Comparison of ML Algorithms for Predicting Compressive Strength

Algorithm	Strengths	Weaknesses	Suitability for Fly Ash-Based GPC
ANN	Captures complex nonlinearities; adapts to large datasets; widely used	Requires large dataset; risk of overfitting; black-box nature	High ($R^2 = 0.93$ reported)

SVM	High accuracy for small/medium datasets; avoids overfitting	Less effective with large or noisy data; hard to tune	Moderate (optimized SVM shown effective)
Decision Tree	Easy to interpret; low computational cost	Can be unstable; prone to overfitting	Moderate ($R^2=0.90$)
Random Forest	Handles overfitting better than DT; works well with missing data	Less interpretable; required tuning	High ($R^2 = 0.95$)

Microstructural analysis is an important area of application for machine learning too. Phase creation, gel structure, and degradation mechanisms can be better understood by combining data from methods like scanning electron microscopy (SEM) and X-ray diffraction (XRD) with advanced machine learning algorithms. For example, ML models have been used to examine the connection between aluminosilicate gel production and curing conditions, allowing researchers to modify the microstructure for increased strength and durability.

In evaluating sustainability, machine learning is also essential. ML algorithms are capable of assessing the environmental effect of geopolymer manufacturing, including carbon emissions, energy consumption, and resource use, through the analysis of lifecycle data. As a result, more environmentally friendly substitutes for conventional cement-based products are being created. Additionally, ML can optimize the use of industrial byproducts like fly ash and slag, further enhancing the sustainability of geopolymers.

ML models can forecast how different curing temperatures, activator ratios, and precursor compositions will affect the geopolymerization process in process optimization. This results in consistent and superior geopolymer products by enabling real-time modifications and better control over the manufacturing process. For instance, ML has been utilized to balance compressive strength and workability by optimizing the Na_2SiO_3 to NaOH ratio (Feng et al., 2020).

Notwithstanding its promise, issues with data scarcity, model interpretability, and the requirement for high-quality datasets hinder the use of machine learning in geopolymer research. Nevertheless, these constraints are being addressed by continuous improvements in machine learning algorithms as well as the growing amount of experimental data that is available. It is anticipated that future studies will concentrate on creating hybrid models, combining machine learning with sophisticated characterisation methods, and investigating the application of deep learning to more intricate geopolymer systems.

To sum up, machine learning is transforming geopolymer research by offering strong instruments for sustainability evaluation, process optimization, and predictive modeling. As machine learning techniques advance, they might open up new avenues for the creation and use of geopolymers, resulting in a built environment that is more robust and sustainable (Chaabene et al., 2020).

2.4.2 Machine Learning Algorithms in Geopolymer Science

In geopolymer science, machine learning (ML) algorithms have become indispensable tools that allow researchers to anticipate material characteristics, evaluate complicated information, and improve formulations with surprising efficiency. These algorithms may be roughly divided into four categories: deep learning, reinforcement

learning, unsupervised learning, and supervised learning. Each of these categories offers special skills for tackling certain geopolymer research difficulties. For property prediction and mix design optimization, supervised learning techniques like Random Forest (RF) and Artificial Neural Networks (ANNs) are frequently utilized. For example, ANNs have been used to forecast durability and compressive strength based on the composition of the precursor and the curing conditions, eliminating the need for lengthy experimental trials (A. Ahmad et al., 2021a). In order to classify geopolymer samples and determine important influencing elements, unsupervised learning methods like as Principal Component Analysis (PCA) and K-Means Clustering assist in finding patterns and correlations in unlabeled data. Through iterative feedback, reinforcement learning, despite its less widespread application, has the ability to optimize activator ratios and curing conditions. Using multi-layered neural networks, deep learning, a branch of machine learning, analyzes complicated datasets like microstructural pictures from SEM or XRD to provide details about phase formation and long-term degradation mechanisms.

Beyond predicting properties, machine learning techniques are also used for microstructural study and sustainability evaluation. Researchers may better comprehend the geopolymerization process and customize materials for particular applications by combining machine learning (ML) with sophisticated characterisation techniques. Furthermore, ML models can assess how the manufacture of geopolymers affects the environment, which helps to advance the creation of more environmentally friendly building materials. Despite their potential, challenges such as a shortage of high-quality datasets, interpretability of the models, and data scarcity still exist. It is anticipated that future studies will concentrate on creating hybrid models, incorporating machine learning into real-time monitoring systems, and investigating deep learning for increasingly intricate geopolymer systems (Jamal et al.,

2025). To sum up, machine learning algorithms are transforming geopolymer science by providing strong instruments to speed up research, maximize efficiency, and advance sustainable building materials.

2.4.3 Supervised vs. Unsupervised Learning Approaches for Geopolymer Analysis

Supervised learning and unsupervised learning are two essential machine learning techniques with different applications in geopolymer research. Supervised learning is perfect for property prediction and optimization because it employs labeled datasets to anticipate certain outcomes, such compressive strength or setup time. To simulate the correlations between input factors (such as precursor composition and activator concentration) and output qualities, methods such as gradient boosting techniques, support vector machines (SVMs), and artificial neural networks (ANNs) are frequently used. On the other hand, unsupervised learning helps with exploratory research and reveals hidden links in geopolymer systems by analyzing unlabeled data to find patterns or clusters. Techniques like clustering algorithms and Self-Organizing Maps (SOMs) are used to decrease dimensionality or categorize data, offering insights into important influencing elements. When combined, these methods offer complementary insights that help researchers better understand complicated behaviors, optimize formulations, and improve the performance and design of geopolymer materials for sustainable applications.

2.4.3.1 Supervised Machine Learning

Artificial Neural Networks (ANNs) are a key component of supervised machine learning, designed to represent complicated interactions between inputs and outputs. Supervised learning focuses on training models with labeled data, where each input corresponds to a known outcome. ANNs succeed in this sector because of their multi-

layered design, which allows them to understand sophisticated data patterns(Ling et al., 2021). ANNs create accurate predictions on unseen data by iteratively modifying the weights and biases via a method known as backpropagation. The adaptability and ability to approximate nonlinear functions is what makes ANNs so effective. This makes them ideal for applications such as classification, regression, and forecasting. For example, when predicting compressive strength in concrete mix design, an ANN may examine constituent quantities and curing conditions to produce accurate forecasts. That is why the use of ANNs to predict the compressive strength of geopolymer concrete has become popular in civil engineering (Shahmansouri et al., 2021). Each layer of the ANN converts the input data into increasingly abstract representations, incorporating both linear and nonlinear correlations that standard models might overlook. Thus, ANNs demonstrate supervised learning by using labeled datasets to generalize patterns and produce accurate predictions (Dao et al., 2019; Gupta et al., 2022). Their versatility and predictive capability have made them valuable in a variety of sectors, including material science, finance, and healthcare, where comprehending complicated input-output interactions is essential.

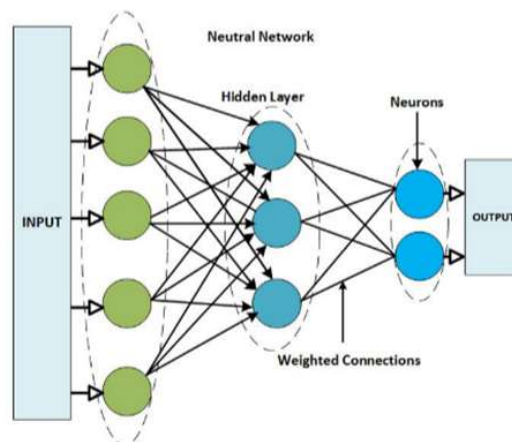


Figure 2.9: The Layered Architecture of Artificial Neural Network (Rusna et al., 2022)

Supervised learning is a variant of machine learning that trains the model using labelled data. This implies that each training example consists of an input and the proper output, with the algorithm's purpose being to learn the mapping from input to output. Once the dataset is trained, the model may make predictions on previously unknown data. These datasets are intended to "supervise" or train algorithms for data classification and accurate forecasting of outcomes. Supervised learning is commonly used for problems requiring prediction or classification, and it is classified into two subcategories: regression and classification (Nasteski, 2017).

Classification and regression are the two main task types that are included in supervised learning. Algorithms that classify test data into predetermined groups like separating apples from oranges or putting spam emails in a different folder from the inbox are used to solve classification challenges. Support vector machines (SVMs), decision trees, random forests, and linear classifiers are examples of common classification techniques. Each has certain advantages when it comes to processing various kinds of data. Regression, on the other hand, is a supervised learning method that is perfect for forecasting numerical values as it can be used to determine the causal links between dependent and independent variables. Regression models, for example, may predict a company's sales income based on market trends and previous data (I. C. Yeh, 2007). Linear regression, ridge regression, support vector regression, Lasso regression, polynomial regression, Bayesian linear regression, decision tree regression, and random forest regression are all commonly used regression techniques. Regression and classification work well together to solve a variety of supervised learning predicted issues.

2.4.3.2 Unsupervised Machine Learning

Unsupervised learning is a kind of machine learning in which the model learns from unlabeled data, meaning the input dataset does not contain predefined output labels or target values. These algorithms identify hidden patterns in data without requiring human intervention (thus the term "unsupervised"). In contrast to supervised learning, which uses labeled samples to assist the model, unsupervised learning algorithms find patterns or structures in the data on their own. This makes it especially helpful when investigating fresh datasets or uncovering undiscovered insights without being aware of the results beforehand. The primary applications of unsupervised learning usually include clustering, association, dimensionality reduction, and pattern recognition (Dike et al., 2018; Qu et al., 2021).

Clustering, association, and dimensionality reduction are some of the fundamental methods used in unsupervised learning, each with a specific function in the analysis of unlabeled data. A popular method for identifying patterns and structures in a dataset is clustering, which groups data points according to their basic commonalities. For applications like customer segmentation, anomaly detection, and picture segmentation, algorithms like K-means, hierarchical clustering, and DBSCAN are crucial because they can autonomously find groups of extremely similar data points. For example, hierarchical clustering creates a tree-like structure based on data-point distances, whereas K-means divides data into K clusters by maximizing cluster centroids. In data-driven industries, these techniques aid in decision-making and offer crucial insights (Yorek et al., 2016). Another type of unsupervised learning is association, which is frequently used in recommendation engines and market basket analysis to find links between variables in a dataset using rules. High-dimensional data presents issues that are addressed by dimensionality reduction

techniques including Principal Component Analysis (PCA), t-distributed stochastic neighbour embedding (t-SNE), and Linear Discriminant Analysis (LDA). These techniques improve generalization performance, model interpretability, and computing economy by minimizing the number of features while maintaining important information. To provide the best results for the machine learning job at hand, choosing the right approach necessitates careful thought in order to strike a balance between dimensionality reduction and the preservation of important data characteristics (Kumar et al., 2017).

2.4.3.3 Reinforcement Learning

Reinforcement Learning (RL) is a dynamic type of machine learning where an agent learns to make decisions by interacting with its environment. As opposed to supervised learning, which depends on labeled data, reinforcement learning (RL) is dependent on trial and error. The agent acts, gets feedback in the form of rewards or penalties, and modifies its behavior over time to maximize cumulative rewards. Based on behavioural psychology, this learning process involves the agent avoiding negative events (penalties) and learning from favourable ones (rewards). The key components of RL are the agent, the environment, actions the agent can take, and the reward signal that provides feedback on the success of an action.

A major problem in reinforcement learning is striking a balance between exploitation (selecting known acts that yield large rewards) and exploration (trying novel actions). The agent's goal is to discover the best course of action, or a tactic that optimizes long-term gains. Value functions, which gauge the relative merits of a given situation or course of action in terms of potential rewards, are frequently used to do this. In order to determine the optimal

policy through ongoing interaction with the environment, algorithms such as Q-Learning and Deep Q Networks (DQN) are frequently adopted (Gu et al., 2022).

Applications for RL are numerous including robotics, where it trains robots to perform complex tasks like walking or picking objects; gaming, where it has produced superhuman performance in games like AlphaGo; and autonomous vehicles, where it assists self-driving cars in making decisions in real time based on obstacles and traffic. In engineering, RL may improve processes like concrete mix design, modifying parameters to reach optimal performance, while in healthcare, it is utilized to optimize treatment planning. All things considered, RL is a flexible and effective tool for resolving challenging decision-making issues in a range of sectors because to its capacity to adjust and learn from feedback.

2.4.4 Predicting Geopolymer Properties Using Machine Learning

In geopolymer research, machine learning (ML) has emerged as a game-changing method that makes it possible to forecast material characteristics with great efficiency and precision. Researchers can model intricate relationships between input parameters (like precursor composition, activator concentration, and curing conditions) and output properties (like compressive strength, setting time, and durability) by utilizing algorithms like Random Forest (RF), Support Vector Machines (SVMs), and Artificial Neural Networks (ANNs). By eliminating the need for lengthy experimental trials, these predictive models speed up the process of optimizing geopolymer formulations (Ahmari et al., 2013).

One of machine learning's main advantages is its capacity to examine big, complicated datasets and reveal hidden associations that are hard to find using more conventional techniques. To provide useful information for material design, machine learning algorithms, for example, can forecast how changes in the activator ratio or curing

temperature affect the mechanical and microstructure characteristics of geopolymers (Chu et al., 2021). ML also facilitates real-time monitoring and process optimization through integration with sophisticated characterization methods like X-ray diffraction (XRD) and scanning electron microscopy (SEM), allowing researchers to make data-driven modifications throughout the geopolymerization process (Jagadesh et al., 2023).

In addition to predicting properties, machine learning (ML) makes sustainability evaluations easier by analyzing the environmental effects of various formulations. This contributes to the creation of more environmentally friendly substitutes for conventional cement-based products, supporting international sustainability objectives (Z. Zhang et al., 2014). Data scarcity, model interpretability, and the requirement for high-quality datasets are still issues, nevertheless. According to Garcia et al. (2015), future studies should concentrate on creating hybrid models, merging machine learning with multi-scale modeling techniques, and investigating deep learning for increasingly intricate geopolymer systems.

In summary, machine learning is transforming the capacity to forecast the characteristics of geopolymers by providing effective instruments to improve formulations, boost efficiency, and advance sustainability. As machine learning techniques advance, they might open up new avenues for the creation and use of geopolymers, resulting in a built environment that is more robust and sustainable.

2.5 Adoption of Machine Learning for Concrete Studies

The use of machine learning (ML) has led to considerable advances in the field of concrete science, notably in the study of geopolymers. Machine learning approaches have shown to be extremely useful in predicting material characteristics, improving mix designs, and comprehending microstructure-property correlations. Traditional machine learning

methods, such as Random Forests (RF), Support Vector Machines (SVMs), and Artificial Neural Networks (ANNs), have demonstrated their ability to uncover complex relationships between material composition and performance outcomes by analyzing large datasets of mix designs and experimental measurements (Young et al., 2019). These approaches have decreased the need for trial-and-error experimentation, enabling real-time quality monitoring, and aided the creation of more sustainable and efficient concrete mixes. Furthermore, machine learning has helped to improve curing procedures, anticipate concrete behavior under certain situations, and improve overall building material quality.

In the context of geopolymers, ML has been effectively used to anticipate and improve mix designs, considerably decreasing the requirement for long experimental trials (Tapeh et al., 2023). For example, ML models have been used to anticipate the performance of geopolymer combinations, allowing researchers to find the best formulas more rapidly. Self-Organizing Maps (SOMs) are one technique used to display and analyze complicated datasets, offering significant insights into the links between input variables and output attributes. These approaches have been very beneficial in studying the microstructure-property correlations of geopolymers, which are essential for designing sustainable building materials. ML's capacity to analyze enormous datasets and simulate complicated, non-linear interactions has made it a must-have tool for concrete research (Chaabene et al., 2020).

The adoption of ML in concrete research provides various benefits over traditional experimental procedures. First, it increases efficiency by lowering the time and expense of trial-and-error experimentation via data-driven forecasts and optimizations. Second, ML improves the accuracy of property predictions by capturing complex relationships in data. Third, it promotes sustainability by improving mix designs and reducing material waste. Finally, ML techniques are scalable, making them appropriate for industrial-scale

applications. As the field of concrete research evolves, the importance of machine learning is projected to expand even further. These technologies have the potential to stimulate innovation in the creation of sustainable building materials, particularly geopolymers, by offering data-driven insights while decreasing reliance on expensive and time-consuming experimental approaches (Ahmad et al., 2021b). Future research might focus on combining ML with sophisticated characterisation techniques like microscopy and spectroscopy to improve our understanding of material behavior and performance.

2.5.1 Self-Organising Map (SOM)

A self-organising map (SOM), also known as a Kohonen map, is a popular type of neural network-based clustering model that performs unsupervised learning tasks to cluster similar data samples into clusters (groups). An SOM is effective for visualizing and representing high-dimensional data with representation in a low-dimensional map, which is commonly a 2D or 3D grid space (Teuvo Kohonen, 1990). An SOM is distinguished by its self-organizing feature to organize nodes according to learning with topological relationships among the nodes and data samples, which enables them to self-organize based on the grid distance. SOMs modify their weights to represent the distribution of data samples, which are typically positioned at the center of a cloud of data samples. During learning, neuron weights are first initialized with random values or vectors, which are subsequently updated iteratively using a competitive learning algorithm that promotes the resemblance of neighbouring neurons in terms of the input data they represent (Asan et al., 2012; Chang et al., 2017).

An SOM is composed of neurons arranged on a regular, usually one- or two-dimensional, grid. Higher-dimensional grids are also feasible; however, they are rarely

employed because of the difficulties associated with their display. The SOM toolbox supports map grids of any dimension, although visualization is only available for one- and two-dimensional maps. The structure of the map is determined by the neighbourhood relation that connects each neuron to its neighbour. In a two-dimensional scenario, the map neurons can be placed on a rectangular or hexagonal lattice, as illustrated in Figure 2.10. If the sides of the map are joined, the overall shape of the map becomes a cylinder or toroid, as shown in Figure 2.11.

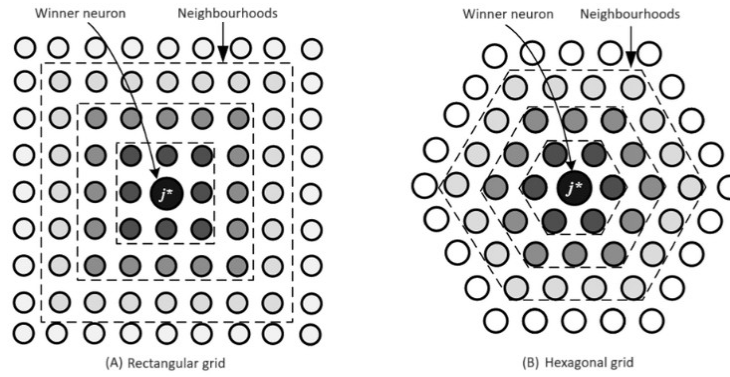


Figure 2.10: Grid representation of the neurons in SOM (Ullah et al., 2019)

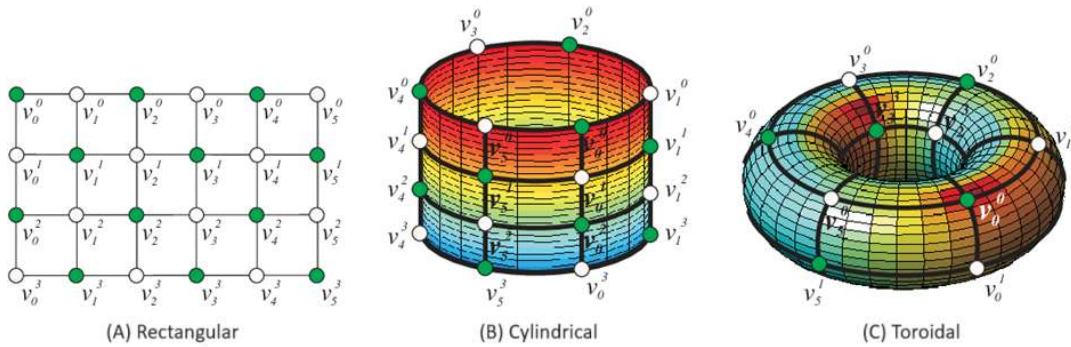


Figure 2.11: Illustrations of Rectangular, Cylindrical, and Toroidal grids (Park et al., 2015)

Each SOM neuron i is coupled with a d -dimensional prototype (also known as weight, reference, codebook, or model) vector $m_i = [m_{i1} \ m_{i2} \dots \ m_{id}]$, where d denotes the input vector dimension. Each neuron has two locations: one in the input space (the prototype vector) and one in the output space (the map grid). Thus, SOM is a vector projection method that specifies a nonlinear projection from the input space to a lower-dimensional output space. In contrast, during the training, the prototype vectors moved according to the probability density of the input data. Thus, SOM is a vector quantization algorithm. Simply put, SOM is an algorithm that combines two tasks: vector quantization and projection.

The neurons that are close to neuron i are their immediate neighbours or N_{i1} . Figure 2.10 shows neighbourhoods of various sizes in the rectangular and hexagonal lattices. The innermost polygon represents the one-neighbourhood, the second the two-neighbourhood, and the largest three-neighbourhood. The neighbourhood function regulates the closeness of the neurons. The bubble is the simplest neighbourhood function: it is constant across the winner unit's neighbourhood but zero elsewhere. The resulting granularity of the mapping is determined by the neighbourhood function and the number of neurons. The larger the area in which the neighbourhood function has high values, the more rigid is the map. The larger the map, the more adaptable it becomes. This interplay affects the accuracy and generalization capability of SOM (Kumar et al., 2016).

In the conventional SOM, the number of neurons and their topological relationships are fixed from the start. There are four difficulties to resolve: number of neurons, map grid dimensions, map lattice, and shape. The number of neurons should typically be as large as possible, with the neighbourhood size determining the smoothness and generalization of the mapping. Even when the number of neurons surpasses the number of input vectors, the mapping does not suffer significantly as long as the neighbourhood size is chosen correctly.

However, as the size of the map increases, for example, to tens of thousands of neurons, the training phase becomes computationally impractical in most applications.

For possible applications, the map grid should have the same shape as the data manifold. Consequently, toroid and cylindrical shapes are recommended only when the data are known to be circular. For the default sheet-shaped map, it is advised that the side length along one dimension should be longer than the others for the map to properly orient itself. The eigenvalues of the training dataset can be used as guidelines for determining the map grid side lengths. A hexagonal lattice is commonly preferred because it ensures that all six neighbours of a neuron are at the same distance. This smoothened the maps and made them visually appealing (Holmbom et al., 2011b).

2.5.2 Hierarchical Self-Organizing Map (HSOM)

The hierarchical Self-Organizing Map (HSOM) is an advanced extension of the traditional Self-Organizing Map (SOM), designed to handle large and complex datasets by organizing them hierarchically. Unlike flat SOM, which provides a single-layered view of clustered data, HSOM employs a multilayered structure that refines clustering results progressively across multiple levels. At the top layer, the HSOM provides a broad overview of the data, while subsequent layers explore deeper into sub-clusters, enabling multi-resolution analysis. This approach enhances scalability by breaking datasets into smaller, more manageable subsets and improves interpretability by preserving the topological relationships inherent in the data.

HSOM has been applied in various fields, such as data mining, bioinformatics, text clustering, market segmentation, and image processing, where it effectively reveals both high-level patterns and detailed substructures. Despite its computational complexity, the

hierarchical nature of the HSOM makes it a powerful tool for exploring and visualizing intricate datasets, offering a comprehensive view of the relationships within the data. An HSOM typically resembles a tree-like structure with each node (representing a cluster) connected to an SOM or a set of sub-nodes in the next layer. This hierarchical visualization helps to comprehensively understand data distributions and relationships (Henriques et al., 2012). Figure 2.12 shows the HSOM multi-layer structure where the base consists of main nodes, each sublayer expands into more detailed maps, and the leaf nodes in the final layer represent the most specific data clusters.

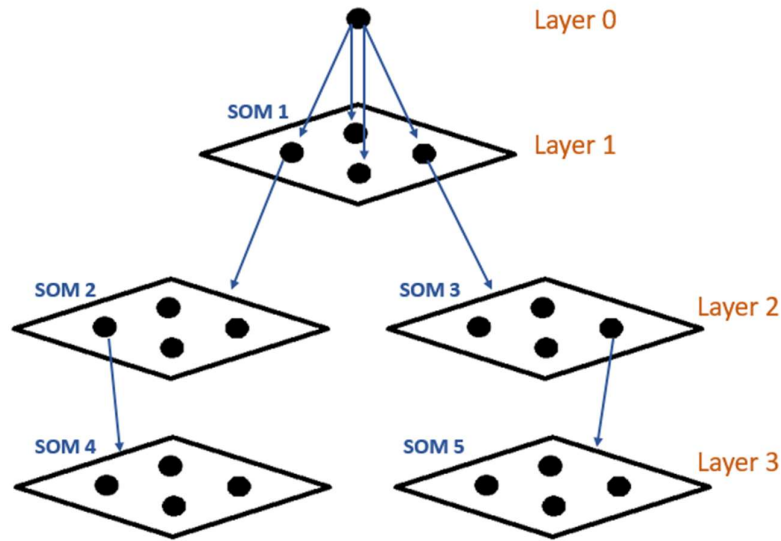


Figure 2.12: Hierarchical Self-Organizing Map Structure(Qu et al., 2021)

2.5.3 Comparison of SOM and HSOM

For better presentation and assessment, Table 2.1 displays a comparison table of Self-Organizing Maps (SOM) and Hierarchical Self-Organizing Maps (HSOM) across several

dimensions. This table facilitates the clear differentiation of the two distinct machine learning techniques.

Table 2.2: Comparison of SOM and HSOM (Qu et al., 2021)

Aspect	Self-Organizing Map (SOM)	Hierarchical Self-Organizing Map (HSOM)
Structure	A single-layer, flat grid of neurons (nodes) arranged in a 2D or 3D lattice.	A multi-layer, tree-like structure with SOMs at each layer, forming a hierarchical arrangement.
Clustering Approach	Produces a single map that represents the data in terms of clusters.	Refines clusters hierarchically; each cluster in the parent map is further analysed in child maps.
Data Representation	Represents data relationships and clusters in one layer, often less suitable for large or complex datasets.	Represents relationships at different levels, enabling the analysis of both macro and micro structures in data.
Scalability	Suitable for small to medium datasets; can struggle with large-scale data.	Handles larger and more complex datasets effectively due to its layered approach.
Topology Preservation	Preserves topological relationships between data points on a single layer.	Preserves topological relationships across multiple layers, maintaining local and global structures.

Training Complexity	Computationally less intensive easier to implements.	Computationally more demanding due to training multiple SOMs across layers.
Visualization	Results in a single map that is easier to visualize and interpret.	Requires visualization tools that can handle hierarchical structures and multiple maps.
Applications	Used in basic clustering, visualization, dimensionality reduction, and pattern recognition tasks.	Applied in domains requiring multi-resolution analysis, such as bioinformatics, text mining, and large-scale data.
Training Data Handling	Processes all data at once in a single training phase.	Processes data in stages, with child maps trained on subsets derived from parent clusters.
Implementation Effort	Relatively easy to implement and requires less tuning.	Requires more effort to implement, including decisions on the number of layers and stopping criteria.
Limitations	Struggles with scalability and does not provide hierarchical insights.	Computationally expensive and potentially complex to interpret for users unfamiliar with hierarchical data structures.

In Summary, SOM is a simpler single-layer approach that is ideal for smaller datasets and straightforward visualizations. It is computationally efficient, but lacks the ability to capture hierarchical relationships or handle large datasets effectively. However, HSOM, with its hierarchical design, excels at multiscale data analysis and is better suited for larger, more complex datasets. It offers richer insights but at the cost of higher computational complexity and implementation effort.

For this study investigation, the SOM/HSOM architecture provides clear advantages over other popular machine learning models including Support Vector Machines (SVM), Random Forest (RF), and deep learning techniques. Though they frequently need substantial feature tweaking and may not perform well on high-dimensional, non-linear data unless adequately pre-processed, SVMs and RFs are strong for classification and regression applications. Although deep learning models may handle intricate interactions, they usually require a lot of data and a lot of processing power, which may not be possible in every experimental situation. On the other hand, SOM and HSOM are especially well-suited for pattern identification and exploratory analysis in materials science since they are excellent at visualizing and clustering high-dimensional datasets without requiring a lot of labelled data. These characteristics justify their selection as the primary analytical tools in this research.

2.5.4 SOM in Geopolymer Concrete Strength Prediction

Many researchers have turned to machine learning techniques as cutting-edge methods for predicting the compressive strength of geopolymer concretes. The vast majority of the research aimed to forecast the compressive strength as the model's output by utilizing the fly ash's characteristics, activators, aggregates, additional water, and curing circumstances as the model inputs. Most machine learning models include fly ash, either as a weight or as a ratio with another parameter; only two studies have included the SiO_2 and Al_2O_3 content of fly ash. However, none of the machine learning models have considered the additional chemical, physical, and mineralogical characteristics of fly ash that have been discovered to be essential input variables during experimental investigations. Therefore, it has been speculated that some machine learning models perform poorly for test data because they do not account for the highly variable nature and source dependency of fly ash (Rathnayaka et al., 2024).

In most models, the effect of an alkaline activator is considered. The concentration of alkali hydroxides affects the dissolution process of the precursor material, resulting in the formation of a concentrated gel phase that finally rearranges into geopolymers. Moreover, the silicate ions produced from the alkali silicate solution determine the type of final product, which in turn determines the quality of the geopolymer. In addition, a sensitivity analysis of the model parameters identified curing time, temperature, and water content as key factors influencing the geopolymer reaction. Higher temperatures increase the reactivity, leaving fewer unreacted particles. This significantly affected the initial and ultimate setting times of the geopolymer concrete. The high-water content in geopolymer concrete can lead to segregation and a reduced compressive strength (Khan et al., 2021).

The combination of chemical and mineralogical features of fly ash may result in a standard prediction model that can be used to estimate the compressive strength of geopolymer concrete, regardless of its origin. Furthermore, it allows the combination of fly ash with various sources of variable quality to enhance the resulting concrete. The NaOH concentration, NaOH content, and Na_2SiO_3 content appear to be crucial input variables for machine-learning models because they serve as activators in geopolymerization (Gomaa et al., 2020). With the exception of one study that did not adopt any measures to assess model performance, every study in the current review used either the coefficient of correlation or the coefficient of determination in conjunction with additional error measures that provided an indication of the overall magnitude of the error, such as RMSE, MSE, or MAE (Bhogayata et al., 2021). To determine any potential practical implications, such as the models' behavior for new fly ash sources outside the database, workability, and initial setting time, it is important to evaluate the model performance using laboratory experiments for verification. This will make it practical to evaluate the performance of the model and continue to optimize it while guaranteeing the viability of the final concrete. Five research model predictions and experimental findings are shown in Figure 2.14. The remaining studies did not employ experimental testing to confirm the model's performance.

In a study by Gunasekara et al. (2021), the coefficient of determination based on model predictions and experimental results was 96.6. In contrast, the models developed by Gomaa et al. (2020) and Van Dao et al. (2019) predicted the compressive strength with coefficients of determination of 94.3 and 94.5, respectively. The coefficients of determination of the models developed by Ahmad et al. (2021) and Bhogayata et al. (2021) were 76.0 and 0.45, respectively. It is advised that all produced models be evaluated using

experimental data from several fly ash sources to identify model performance, enhance accuracy, and guarantee the model's capacity for generalization.

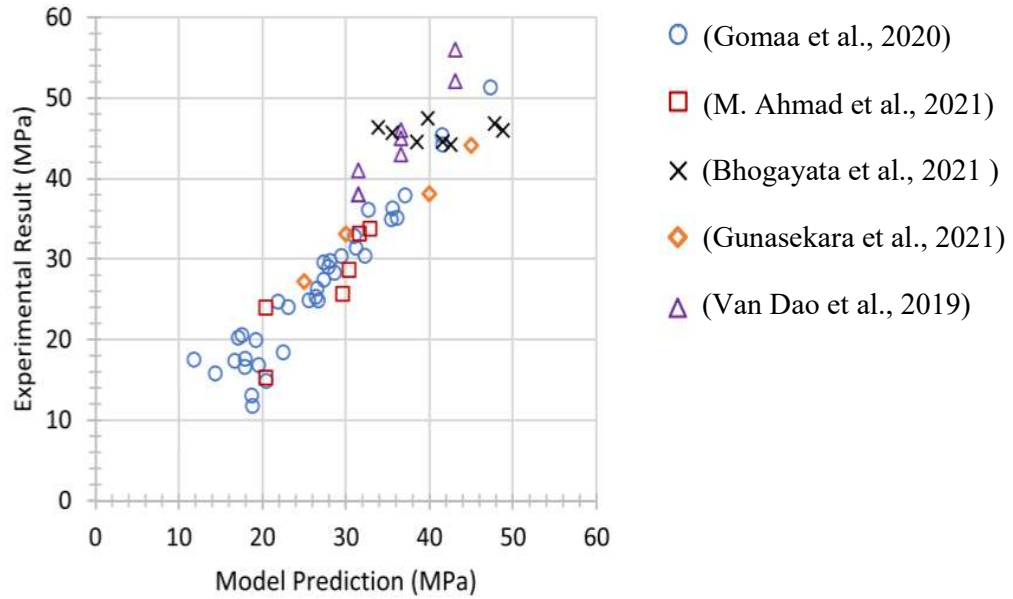


Figure 2.13: Model Predictions vs Experimental Results Comparison

2.6 Key Parameter Relationships Studied by Machine Learning

By examining the connections between different input factors and performance results, machine learning (ML) approaches have been widely used in geopolymers research to predict material qualities. The relationship between precursor content and compressive strength is one of the most studied. The kind of precursor (such as fly ash, slag, or metakaolin) and characteristics like fineness and particle size distribution have been used to predict strength using models like decision trees and artificial neural networks (ANNs) (Kupaei et al., 2014). The impact of activator concentration on strength and setup time is a crucial topic for further study. The effects of factors like sodium silicate and NaOH

concentrations on these qualities have been examined using methods like regression models and support vector machines (SVMs) (Ma et al., 2019).

Studies also utilized self-organizing maps (SOMs) to examine how curing conditions affect mechanical characteristics. Specifically, the effects of temperature, curing time, and relative humidity on strength development have been examined (Lim et al., 2018). Furthermore, workability has been predicted using mix proportions and machine learning. Relationships between the water to binder ratio, superplasticizer dose, precursor reactivity, and results like slump flow and viscosity have been modeled using techniques like statistical process regression (GPR) and polynomial regression. Lastly, the long-term resilience of geopolymers to environmental exposure has been investigated using random forest models, which have been used to analyze performance in situations including sulfate assault, carbonation, and freeze-thaw cycles (Pradhan et al., 2022). Together, these experiments show how versatile machine learning is in revealing intricate correlations and refining geopolymer compositions for improved performance.

2.7 Challenges and Future Prospects

The scarcity of big, high-quality datasets is one of the main obstacles to using machine learning for geopolymer research. The capacity of the model to generalize is diminished by variations in the experimental setup, precursor compositions, and activator concentrations among various investigations. This problem might be lessened with the use of open-access repositories and standardized data gathering procedures (Mohammed et al., 2021). By increasing generalization across various datasets, data augmentation methods and transfer learning may further improve the performance of ML models.

2.7.1 Limitations of Existing ML Models in Geopolymer Research

In geopolymer research, machine learning (ML) models have a number of drawbacks despite their capacity for prediction. One significant obstacle is their incapacity to extrapolate outside of the learned dataset. Because ML models mostly depend on the caliber and volume of training data, they are unable to produce accurate predictions for situations that do not fall within the data range. This is especially difficult in the study of geopolymers, because variations in raw ingredients, curing conditions, or exposure to the environment can cause material behavior to vary greatly.

The empirical character of many machine learning models is another drawback. Without taking into account the basic chemical and physical principles of geopolymerization, the majority of models rely on experimental data. For example, machine learning (ML) can forecast compressive strength based on mix proportions, but it frequently ignores underlying mechanisms such as microstructure formation or alkali activation. This restricts the models' robustness and generalizability, particularly for new formulations or untested situations (Q. Tian et al., 2024).

Hybrid models that combine machine learning with simulations based on physics have been presented as a solution to these problems. Such methods can increase accuracy and dependability by combining data-driven predictions with domain-specific knowledge (Ji et al., 2024). Combining machine learning (ML) with reaction kinetics or thermodynamic modeling, for instance, may improve predictions for situations not included in the training dataset.

The lack of interpretability and explainability in ML models is a recurring problem too. A large number of models, including artificial neural networks (ANNs), function as

"black boxes," offering little information about how they make decisions. In geopolymer research, where knowledge of input-output interactions is essential for material design, their acceptance is hampered by this opacity. To increase interpretability, methods such as SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-agnostic Explanations) are being investigated; nevertheless, their use in geopolymer research is still restricted (Hamilton et al., 2024).

Furthermore, generalizability and scalability are important issues. A lot of machine learning models are trained on little, lab-scale datasets, which could not accurately represent variability in the real world. Disparities between expected and actual performance may result from variations in the suppliers of raw materials, mixing techniques, or curing settings. A potential remedy to improve scalability and applicability is to use transfer learning techniques, which involve adapting models trained on one dataset to new, similar datasets (Huang et al., 2020).

In conclusion, even though ML models have potential for studying geopolymers, problems including scalability, lack of explainability, empirical dependence, and extrapolation limits need to be resolved. Promising approaches to getting over these restrictions include transfer learning, interpretability tools, and hybrid models.

2.7.2 Advancing Machine Learning Applications in Geopolymer Future Research

Enhancing data quality, model interpretability, and forecast accuracy are crucial for advancing machine learning (ML) for geopolymer research. The creation of bigger, high-fidelity geopolymer datasets via standardized databases and experimental partnerships is a top goal in order to provide thorough and trustworthy training data for machine learning models (Chen et al., 2024). More accurate simulations and performance projections can

result from integrating multi-scale modeling techniques that integrate experimental, computational, and machine learning-based predictions. This helps close the gap between macroscopic and microscopic material attributes (Li et al., 2023).

In order to provide real-time advances in geopolymer formulation predictions, future developments should concentrate on creating self-learning adaptive models that can continually update their parameters as new experimental data becomes available (Nazari et al., 2015). By recreating repeated experimental trials in a virtual environment prior to undertaking real-world testing, reinforcement learning might also be used to optimize mix design formulations (Provis et al., 2014). These issues may be resolved to greatly improve ML applications in geopolymer research, which will result in better material qualities, more economical manufacturing, and increased sustainability in infrastructure development and building.

2.8 Chapter Summary

The literature on geopolymer technology was thoroughly reviewed in this chapter, with special attention paid to the function of fly ash as a precursor, the geopolymerization procedure, and the impact of important processing variables on the characteristics of fly ash-based geopolymers. The significance of alkaline activators, curing conditions, and precursor qualities in influencing the mechanical and durability properties of geopolymers was emphasized in the conversation. The chapter also examined how machine learning (ML) are increasingly being used in geopolymer research, highlighting how they may be used to forecast material qualities, improve mix designs, and lessen the need for experimental trial and error. The capacity of key machine learning (ML) approaches to reveal intricate

correlations between input variables and performance results was examined. These techniques included support vector machines (SVMs), hybrid models, self-organizing maps (SOMs), and artificial neural networks (ANNs). Notably, SOM and its advanced variant, hierarchical self-organizing maps (HSOM), share a common foundation and have proven particularly effective for pattern recognition and clustering in high-dimensional material datasets. As such, the SOM-HSOM framework represents a central analytical approach for this research, offering a powerful toolset for discovering nonlinear interactions in geopolymers behavior. In addition, recent advancements were discussed, including the introduction of increasingly complex algorithms like hierarchical self-organizing maps (HSOM), which have the ability to solve challenging research problems in the field of geopolymers.

The SOM/HSOM framework demonstrates significant potential for industrial adoption beyond theoretical applications. By establishing correlations between key compositional variables including activator chemistry, molar ratios, and curing conditions and compressive strength outcomes, these models offer the capability to substantially reduce iterative experimentation requirements. This aligns well with growing industry demand for data-driven mix optimization approaches. However, practical implementation faces several important constraints that must be acknowledged. The models' effectiveness depends heavily on standardized input data, which may prove challenging for smaller producers lacking access to advanced material characterization tools. Additionally, while SOM clusters provide valuable pattern recognition, translating these complex relationships into practical, engineer-friendly guidelines remains an ongoing challenge. Perhaps most significantly, model performance may degrade when applied to fly ash sources showing substantial compositional variability compared to training datasets. Effective mitigation strategies will

likely involve pilot-scale validation studies, development of hybrid models incorporating fundamental physicochemical constraints, and implementation of periodic retraining protocols to maintain accuracy with new material sources. Although high initial implementation costs may present adoption barriers for some operations, documented case studies in comparable material systems consistently demonstrate reductions in R&D-related material waste following successful implementation, suggesting strong potential for return on investment.

Despite the growing application of ML in geopolymers research, current predictive models face several challenges and limitations. A major issue is the variability and limited availability of high-quality datasets, which affects the accuracy and robustness of predictions. Many models, particularly those relying on ANNs, are often treated as “black boxes,” offering little interpretability regarding how input features influence outcomes. Additionally, issues such as overfitting, lack of standardized data preprocessing, and inconsistent validation protocols further hinder model generalizability across different material systems and experimental conditions. These shortcomings suggest that, while ML holds great potential, its practical application in geopolymers research still requires refinement and cautious implementation. These constraints highlight the need for more research, particularly in using cutting-edge machine learning approaches to create sustainable and high-performing geopolymers materials. The findings of this research serve as the foundation for the subsequent chapters, which try to fill these gaps by developing and implementing novel machine learning models for the prediction and enhancement of fly ash-based geopolymers concrete qualities.

CHAPTER 3

METHODOLOGY

3.1 Overview

In this study, mixtures containing fly ash, sodium hydroxide, sodium silicate solution, fine aggregates, coarse aggregates, and water, in which four variables (W/B ratio, molarity, NaOH to Na₂SiO₃ ratio, and curing time) were used as input parameters for modeling.

Alkaline activators are essential for binder hydration and play a crucial role in geopolymerization. Two types of activators were utilized in this study: sodium hydroxide (NaOH) and sodium silicate (Na₂SiO₃). Both activators were procured in flake form, ensuring their availability and consistency for the experiments. The activator for the geopolymer was prepared through standard dry mixing, followed by dissolving the mixture in tap water based on a predetermined water to binder ratio (WBR). The concentration of the alkaline activator was carefully adjusted to maintain its proportionality with the geopolymer solution. In this study, three concentration levels were tested (8, 10, and 12 molarity) to evaluate their effects on the performance of the material. A total of 1,620 samples were prepared in the main design study to collect the compressive strength data.

For pattern analysis, the dataset was segmented into 25 clusters to explore meaningful groupings and hidden trends among the mix designs and their corresponding compressive strength outcomes. A pre-test study involving nine mixed proportions was conducted prior to the main design study, resulting in the preparation of 135 samples. This preliminary study aims to guide the direction of the main design study.

3.2 Sample Preparation

To prepare the geopolymer concrete specimens, a metal mold with dimensions of 100 mm $\times$ 100 mm $\times$ 100 mm was used for compressive strength measurement. All samples were cured in an oven for 24 h at 80 °C. After 24 h of curing, the samples were removed from the mold and left at ambient temperature in sealed bags for hardening until they were ready for testing at 7, 14, and 28 days of age. To achieve reliable repeatability and reproducibility, at least 30 replicates were prepared for each test in the main study design. The results for 7 days and 14 days are for references only, whereas all results for 28 days are gathered mainly for machine learning training purposes.



Figure 3.1: 100*100*100mm Cast Icon Mold, Single and Tri-slots

3.2.1 Pre-test Study

Each pre-test mixing involved the preparation of 15 samples to obtain preliminary results before advancing to the main design study. There are total of 9 pre-test study which prepared in different mix proportions, all test parameters are discussed to identify the main design study direction. This approach ensures the feasibility of the study while simultaneously testing the applicability of machine learning modeling before initiating large-scale sampling.

Table 3.1: Mix Proportion for Trial Study Design

Mix No.	Liquid to binder ratio	Molarity	Sodium Hydroxide to Sodium Silicate ratio	Number of samples	Curing time (days)	Temp 'C
1	0.3	8	1:0.5	15	28	80
2	0.4					
3	0.5					
4	0.3	10				
5	0.4					
6	0.5					
7	0.3	12				
8	0.4					
9	0.5					

3.2.2 Main Design Study

Building on the findings of the pre-test design research, the main design study aimed to establish a clear framework for collecting a substantial dataset for machine learning training. Guided by the insights from the pre-test, the main design study was broadly divided into two key directions:

Sampling set 1:

To test a machine learning algorithm that predicts how the molarity of the activator and the liquid-to-binder ratio affect the strength of GPC.

Sampling set 2:

To evaluate a machine learning method that predicts the effect of the sodium silicate/hydroxide ratio on GPC strength.

Table 3.2: Mix Proportion for Main Design Study

Mix No.	Liquid to binder ratio	Molarity	Sodium Hydroxide to Sodium Silicate ratio	Number of samples	Curing time (days)	Total Sample	Temp 'C
1	0.3	8	1:0	30	7, 14, 28	90	80
2	0.4			30		90	
3	0.5			30		90	
4	0.3	10		30		90	
5	0.4			30		90	
6	0.5			30		90	
7	0.3	12		30		90	
8	0.4			30		90	
9	0.5			30		90	
10	0.3	10	1:0.5	30		90	
11	0.4			30		90	
12	0.5			30		90	
13	0.3		1:1	30		90	
14	0.4			30		90	
15	0.5			30		90	
16	0.3		1:1.5	30		90	
17	0.4			30		90	
18	0.5			30		90	

3.3 Materials

Fly ash from a furnace power plant supplier in Sejingkat, Kuching Sarawak, was used in this study to create the geopolymer concrete examples, Table 3.3 lists the chemical composition that makeup of the material. X-ray fluorescence (XRF) was used at the University of Malaya (UM) to determine the chemical composition. Fly ash from Sejingkat has a chemical composition that meets the ASTM C618 requirements, which categorizes it

as low-calcium fly ash (Class F). The classification was performed according to the criteria listed in Table 3.4.

Table 3.3: Chemical composition of Class F Fly ash (by XRF %)

Component	Mass (% oxide)
MgO	0.5
Al ₂ O ₃	15.7
SiO ₂	53
P ₂ O ₅	1.44
SO ₃	0.144
K ₂ O	4.71
CaO	5.21
TiO ₂	1.57
MnO	0.295
Fe ₂ O ₃	17.1
Rb ₂ O ₂	0.081
SrO	0.372
ZrO	0.000ppm

Table 3.4: Chemical composition of Class F Fly ash based on ASTM 618

Chemical Composition	Percentage, %
Silicon dioxide (SiO ₂) plus aluminum oxide (Al ₂ O ₃) + iron oxide (Fe ₂ O ₃) min %	70.0
Calcium oxide (CaO), %	< 10.0

Fine aggregates with a 4.75 mm sieve were employed, and coarse aggregates ranging from 20 to 40 mm were utilized. Sodium Hydroxide flakes (NaOH), which served as an alkaline activator, were mixed at concentrations of 8, 10, and 12 molarity, while Sodium Silicate powder (Na_2SiO_3) was added in designated ratios in accordance with Sodium Hydroxide.

3.4 Specimen Preparation

The geopolymer concrete was mixed with fine aggregates, coarse aggregates, alkaline reactants (NaOH only or NaOH + Na_2SiO_3), fly ash, and water depending on the required molarity. These solid components were thoroughly mixed either by a mixing machine or by hand, then alkaline reactant flakes were added to the dry mixture, and finally, water was added. Ninety samples were prepared from each mix, with 30 samples crushed at 7, 14, and 28 days separately to measure the compressive strength.

Table 3.5: Mixing Design of Geopolymer Concrete in 1:1 Activator Ratio

Geopolymer	No. of Samples	Fly Ash (Kgs)	Aggregates (Kgs)		NaOH (Kgs)	Na_2SiO_3 (Kgs)
			Fine	Coarse		
Fly Ash Class F concrete	90	35	70	140	4.20	4.20

The mixture was then poured into specimen molds measuring 100 mm × 100 mm × 100 mm and was kept in an oven that had been preheated to a constant temperature of 80°C for 24 h. Subsequently, it was removed, covered with plastic, and left at room temperature for 28 d to determine its compressive strength. The same procedure was repeated until all the test parameters were obtained. The overall experimental procedure is illustrated in Figure 3.2.

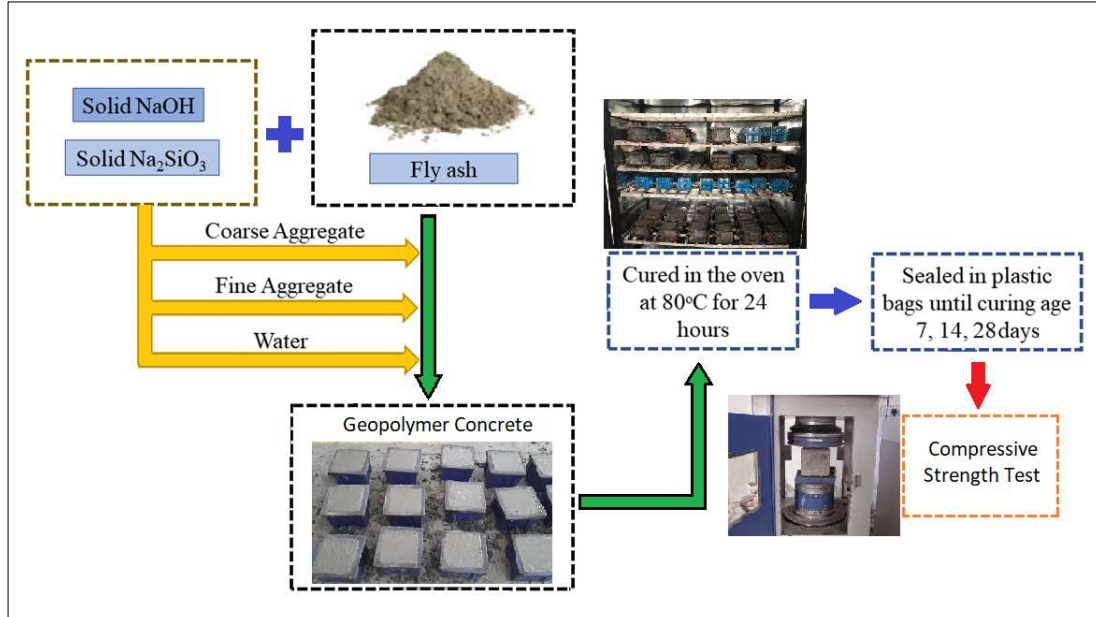


Figure 3.2: Procedure of the Experimental Works

3.5 Compressive Strength Test

Geopolymer concrete samples (100 mm × 100 mm × 100 mm) were tested for compressive strength after curing periods of 7, 14, and 28 d. For each curing age, 30 samples were examined to assess the variation in the compressive strength. All methods of preparation, sampling, and testing were conducted in accordance with BS 1881, Part 116:1983. This compressive strength test was carried out on an automatic 500kN tension/1000kN compression testing machine (UTM-3000 FPR). This equipment was created to address the demand for accurate compression testing of concrete cube samples up to 150 mm and cylinders up to 160 mm × 320 mm. The pace rate applied in the compressive strength tests in this study was 1kN/s. The concrete cube was placed on a compression testing machine that applied a compressive force to the cube until it failed or cracked. The maximum load that the cube could withstand prior to failure was measured, and the

compressive strength of the concrete was derived by dividing the maximum load by the cross-sectional area of the cube. The relevant formulas are as explained in the equation below:

$$\sigma = P/A \quad \text{Equation 3.1}$$

where,

σ = Compressive strength

N/mm² P = Maximum load applied

A = Cross-sectional area at which load is applied, mm

3.6 Machine Learning Study

In this study, two ANN-based learning algorithms, Self-Organizing Map (SOM) and Hierarchical Self-Organizing Map (HSOM), were used for comparative purposes. The SOM is an unsupervised learning method that models data samples from data collection and visualizes the model on a lower-dimensional map size, which reduces the complexity of analyzing and understanding the high-dimensional data space created by the number of parameters. HSOM is an advanced machine learning technique that combines self-organizing maps with hierarchical clustering to analyze and visualize high-dimensional data. Figure 3.3 below showing the SOM neural network architecture connections that explains the number of 4 input vectors and 2D matrix with 5x5 nodes that make up 25 clusters.

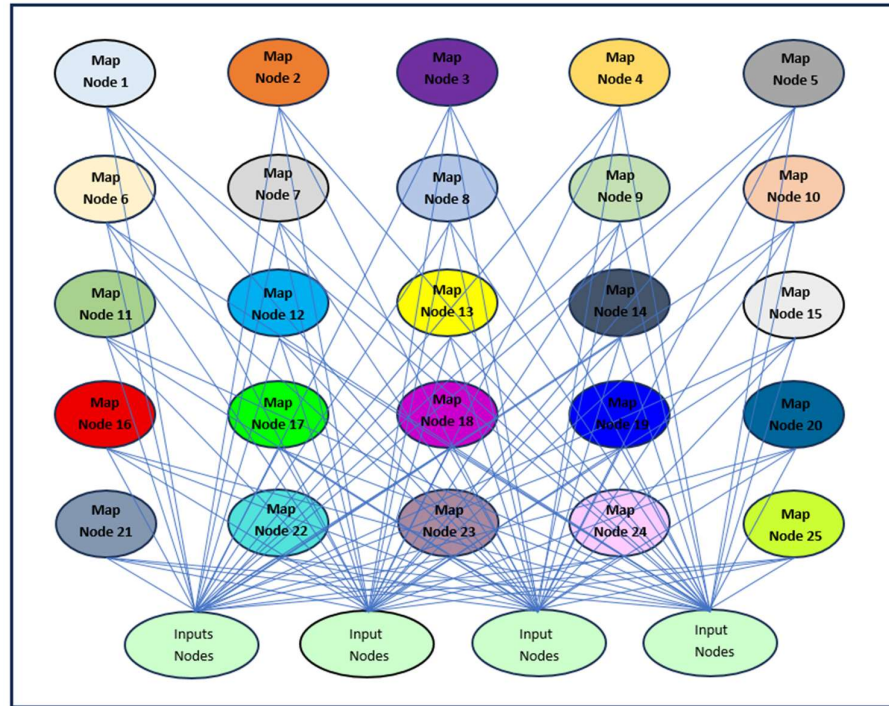


Figure 3.3: Example of SOM with 4 Input Vectors and 25 Output Clusters

3.6.1 SOM Algorithm

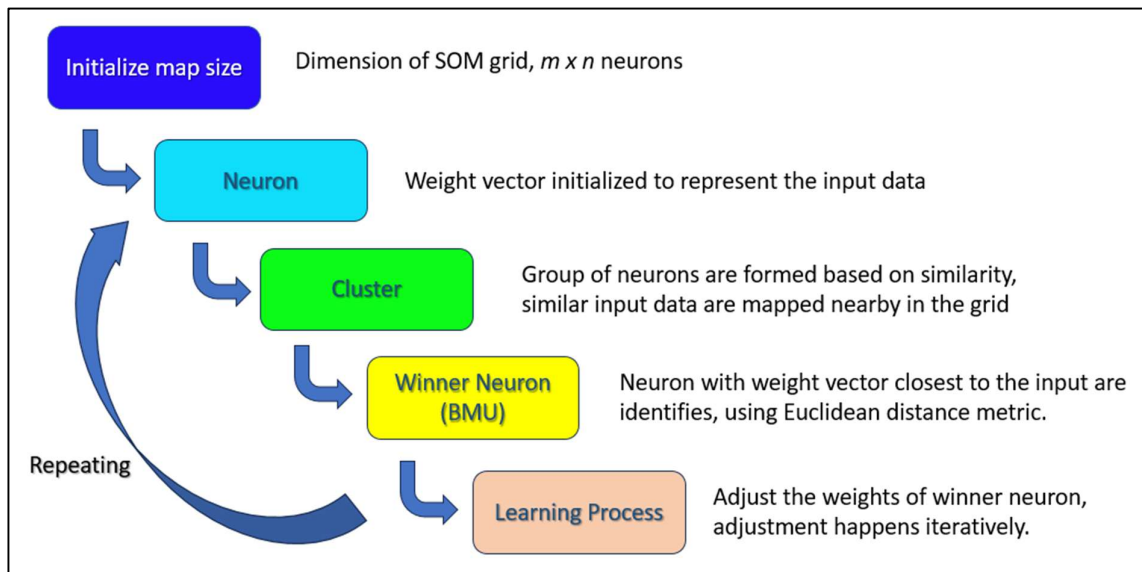


Figure 3.4: The Learning of SOM Algorithm Process

The learning of the SOM algorithm as showing in Figure 3.4 required two predefined parameters: (1) number of nodes or map size, and (2) topological structure (arrangement) of the nodes. The number of nodes was determined based on trial and error for an optimum model, whereas the structure of the nodes was determined as a hexagonal or grid arrangement. Each node is attributed to a prototype weight vector that explains the parameters.

For instance, $W = \begin{bmatrix} w_{1,1}, w_{1,2}, \dots, w_{1,p} \\ \vdots \\ w_{c,1}, w_{c,2}, \dots, w_{c,p} \end{bmatrix}$ indicates that there are c nodes and p is the number of input parameters. All values of W are bounded to an interval of $[0,1]$ to ease similarity matching. Euclidean distance is used in the similarity match measurement, where the winner is determined at the smallest Euclidean distance value. The winner node's prototype weight vector is updated to adapt the input parameters to the winner node, and the prototype weight vectors of the winner's adjacent nodes are updated to self-organize the nodes according to the topological relationship of the SOM map. Artificial neural networks (ANNs), usually called neural networks (NNs), are computing systems inspired by the biological neural networks that constitute animal brains. It can be used to predict the effects of geopolymer parameters using the compressive strength results. Thus, geopolymer concrete samples were prepared with various activator concentrations and alkaline-solution ratios. From the compressive strength results gathered, all the parameters used were analyzed and fed into the SOM. When some inputs are presented to the SOM, the nodes with the highest similarity to the input are considered winners and are updated towards the input. In this manner, an SOM naturally identifies data clusters.

The basic algorithm can be defined as the weight vector in a Self-Organizing Map (SOM), which indicates each neuron's position in the feature space and iteratively evolves

to match the patterns in the input dataset. Initially, the weight vectors are assigned values at random within the range of the dataset or are initialized with random samples of input data.

As explained in the Equation 3.2 below:

$$X = \{x_1, x_2, \dots, x_n\} \quad \text{Equation 3.2}$$

where X is input dataset which n is determined by total how many samples.

$$x_i = \{x_{i1}, x_{i2}, \dots, x_{ip}\} \quad \text{Equation 3.3}$$

x_i refers to one data sample, and p depends on the number of parameters/features included.

$$W = \{w_1, w_2, \dots, w_{(ab)}\} \quad \text{Equation 3.4}$$

Initialize prototype weight vectors; W is defined as the SOM cluster that relies on the map size; if the map size is 2×2 , there will be clusters W1 through W4, and if the map size is 4×4 , there will be clusters W1 to W16.

$$w_j = \{w_{j1}, w_{j2}, \dots, w_{jp}\} \quad \text{Equation 3.5}$$

w_j refers to each cluster weight vector, where $j1$ is the predicted parameter 1 and $j2$ is the predicted parameter 2.

During training, each input vector is matched to the neuron with the closest weight vector, as calculated by a distance measure, such as the Euclidean distance. This neuron, known as the Best Matching Unit (BMU), and its neighbors' weight vectors are modified to bring them closer to the input vector. Equation 3.6 below explains the Euclidean distance while Equation 3.7 explains the BMU:

$$d(x_i, w_j) = \sqrt{(x_{1i} - w_{1j})^2 + (x_{2i} - w_{2j})^2 + \dots + (x_{pi} - w_{pj})^2} \quad \text{Equation 3.6}$$

$$BMU = \arg_j \min\{d(x_i, w_j)\} \quad \text{Equation 3.7}$$

To achieve convergence, the update process is regulated by a learning rate and neighborhood function, both of which steadily decrease. During this iterative process, the weight vectors evolve to reflect the distribution and relationships in the input data, providing as the basic "recipe" for mapping and grouping the data within the SOM as explained in the Equation 3.8 below:

$$w_{j(new)} = w_{j(old)} + h(BMU, j)(x_i - w_{j(old)}) \quad \text{Equation 3.8}$$

where Learning rate function,

$$h(BMU, j) = \alpha \exp\left(\frac{|r_{BMU} - r_j|}{\delta}\right) \quad \text{Equation 3.9}$$

α is the preset learning rate of the SOM, which is between zero and one.

$|r_{BMU} - r_j|$ is the distance between the BMU nodes and their j th neighbor, and δ determines the neighborhood size.

Thus, when $j=BMU$, $h(BMU, BMU)$, learning rate is equal to α . Else, $h(BMU, j)$, learning rate is $< \alpha$. Figure 3.5 below is showing the flowchart of a sequential SOM algorithm for better understanding.

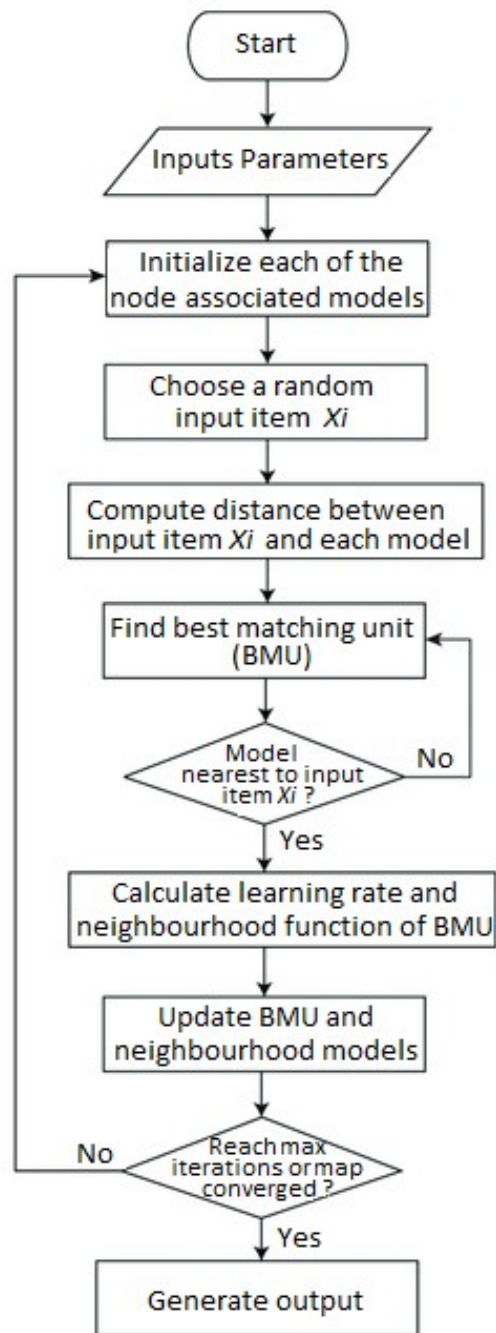


Figure 3.5: Flowchart of a Sequential SOM (Winston et al., 2020)

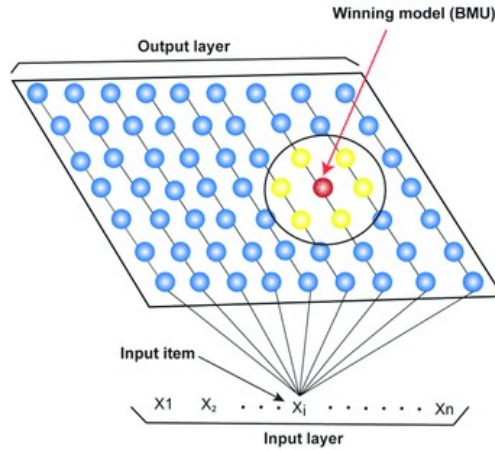


Figure 3.6: SOM Input and Output Layers (Yorek et al., 2016)

In Figure 3.6, the input items (X_1 - X_n) are compared individually to models/nodes. The red model represents the best-match unit, and the surrounding models (yellow, within the larger circle) are all its neighbours. The network neuron whose pattern most closely resembles the input data is known as the Best Matching Unit (BMU). The BMU is determined by comparing the weights of each neuron to the input values; the neuron with the least difference is selected. Following the identification of the BMU, the surrounding neurons, referred known as neighboring units, are also somewhat modified to resemble the input. This keeps comparable inputs near together on the grid and aids the map in identifying patterns in the data.

3.6.2 HSOM Algorithm

The flowchart outlines the process of training a hierarchical Self-Organizing Map (HSOM). It begins by loading the training dataset, represented by $X = \{x_i\}$, where each x_i

is a data point. Usually, the training cycle or better known as epoch are set at a fixed rate 200, this means each cycle of simulation will go for 200 rounds simulation. The first step is to set up the map $R = \{r_j\}$, which consists of c clusters. These clusters are then initialized with prototype weight vectors $W = \{w_1, w_2, \dots, w_c\}$, within a specified range, usually between 0 and 1. For each input data point x_i , the algorithm identifies the best matching unit (BMU), which is cluster r_j whose weight vector is closest to x_i . Once the BMU is identified, the weight vectors W are updated according to Kohonen's learning rule, which adjusts the weight vectors to fit the data better.

Next, the algorithm calculates the sample hit counts, denoted as $S = \{s_1, s_2, \dots, s_c\}$, which represents the number of data points assigned to each cluster. A threshold θ was applied to these hit counts, and if a cluster's hit count s_j exceeded the threshold, the cluster was considered for further refinement. If the threshold is not satisfied, the process ends for that cluster. However, if the cluster has sufficient data points, a subset of the training data from that cluster is selected to train a new submap for finer clustering. This hierarchical refinement process continues iteratively, allowing for a deeper analysis of data at multiple levels. Figure 3.7 below is showing the HSOM process while Figure 3.8 is showing the sequential flowchart of HSOM for better understanding.

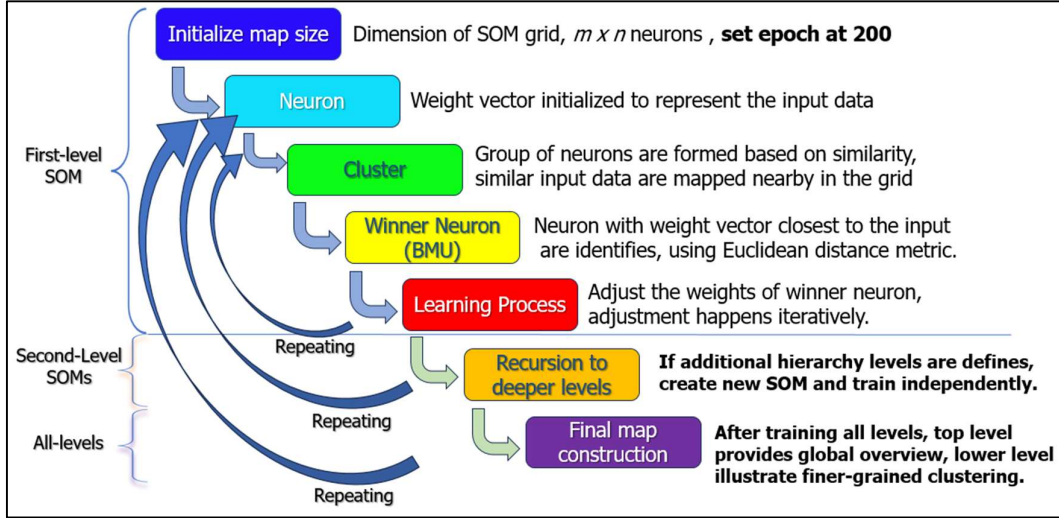


Figure 3.7: The Learning of Hierarchical SOM Algorithm Process

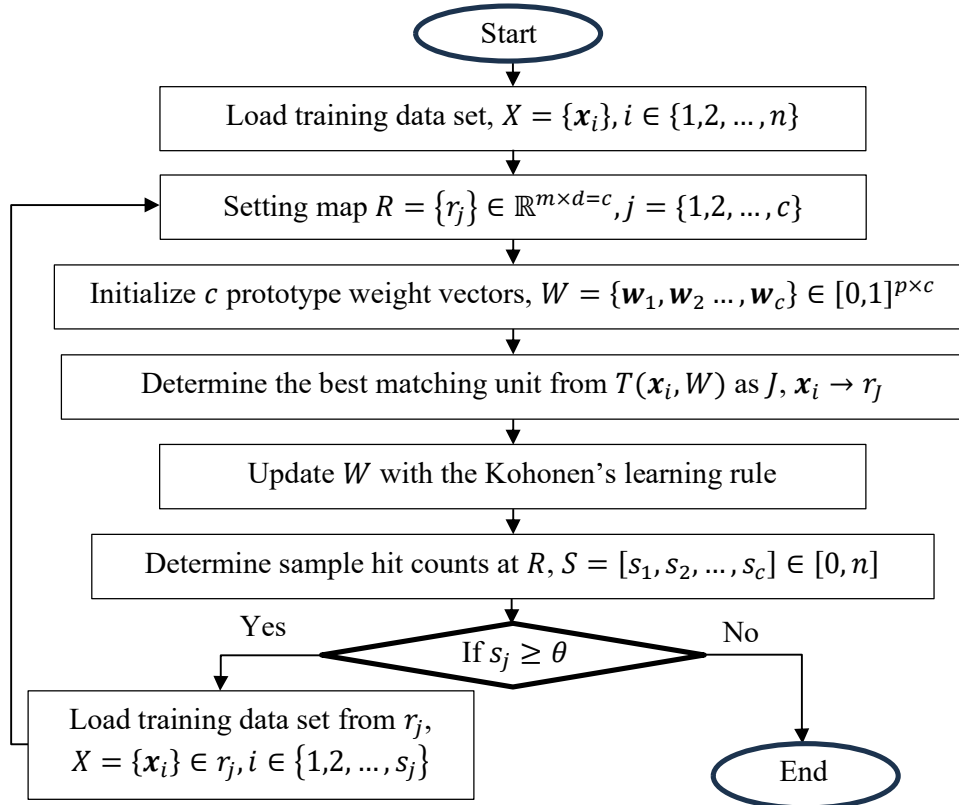


Figure 3.8: Flowchart of Sequential HSOM (Qu et al., 2021)

The process concludes when all clusters either meet the threshold for further refinement or are deemed adequately clustered, leading to the final hierarchical map structure. This multilayered approach allows the HSOM to handle large, complex datasets by refining clusters at multiple levels of resolution, offering both a broad overview and detailed analysis of the data structure. Figure 3.9 below illustrate the actual HSOM architecture and how it functioning.

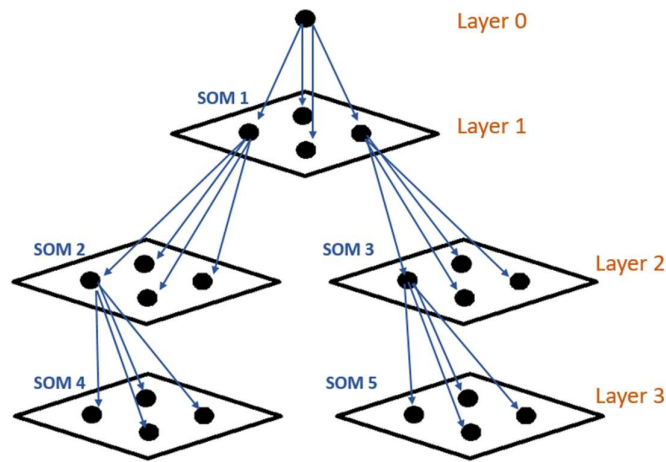


Figure 3.9: HSOM Architecture Diagram

HSOM is a more complex form of the standard SOM that arranges data in several levels. The data is first grouped into large clusters using a top-level SOM. A child SOM at the following level then receives each of these clusters and uses them to carry out more intricate clustering. This procedure can go on for a number of levels, creating a structure like a tree. When working with complicated or sizable datasets, HSOM is helpful because it enables us to examine the data gradually, from broad trends to more specific details, which improves the structure and insight of analysis. The output cluster and layering by HSOM is auto-generated by the algorithm depending on the input's dataset similarity.

3.7 Project timeline

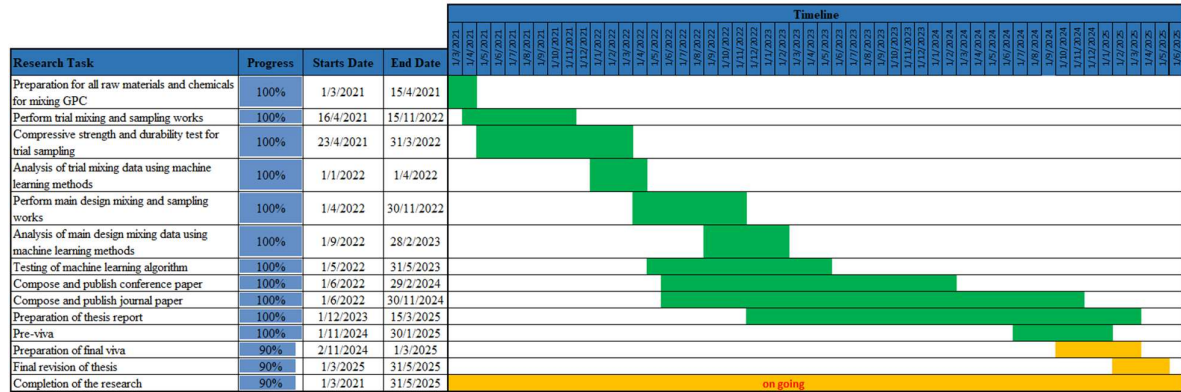


Figure 3.10: Gantt Chart for the Research Timeline

3.8 Chapter Summary

This chapter outlines the study's methodology, emphasizing testing, material selection, sample preparation, and machine learning applications. It starts with a summary of the research methodology before going into depth on how concrete specimens are prepared, including the primary design study and a pre-test study. The methods for preparing the specimen and assessing its compressive strength are covered after a discussion of the components utilized in the experiment, including activators and precursors. Additionally, the chapter examines the use of machine learning for compressive strength prediction, particularly the Self-Organizing Map (SOM) and Hierarchical Self-Organizing Map (HSOM) algorithms. The project timetable concludes by outlining the stages of the investigation, ranging from the use of machine learning models to sample preparation.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Overview

The results and analysis from the use of machine learning approaches to forecast the characteristics of geopolymers are presented in this chapter, with an emphasis on compressive strength and the results of Self-Organizing Map (SOM) simulations. The chapter is set up to validate the models, explain the results in a methodical manner, and offer insights into how successful the suggested approaches are.

Section 4.2 investigates the results of compressive strength, emphasizing the impact of the geopolymer's molarity, water/binder ratio, and activator ratio. The correctness and dependability of the models are evaluated in this part by comparing trial-test results with main design experimental data.

In section 4.3 presents the outcomes of the SOM simulations. This section looks at how SOMs have been applied to the geopolymer dataset to find links and patterns that help us better understand how the material behaves in different scenarios.

The models' validation and verification are the main topics of Section 4.4. It is separated into three sections: The ground truth for mix proportion verification, which makes sure the input data appropriately depicts the physical mixes, is covered in Section 4.4.1. The validation of SOM training is discussed in Section 4.4.2, which assesses how well the SOM captures the complexity of the dataset. The validation of Hierarchical Self-Organizing Map (HSOM) training is finally covered in Section 4.4.3, which evaluates the improved ability of HSOMs to handle more complex data structures.

By means of this thorough examination, Chapter 4 seeks to present an understandable and thorough description of the outcomes, providing insightful conversations that further the field of machine learning-based geopolymer research.

4.2 Compressive Strength Results

This section discusses the compressive strength results of geopolymers obtained from laboratory experiments, focusing on the effects of molarity, water/binder ratio, and activator ratio. The impact of molarity is investigated; greater molarity levels often improve strength and geopolymerization, while excessive molarity might cause lower workability and quick setting. Because too much water might damage the geopolymer matrix, the effect of the water-to-binder ratio is also examined, emphasizing the need of striking a balance between workability and strength. Furthermore, the activator ratio of sodium silicate to sodium hydroxide is investigated, and certain ratios are discovered to maximize the geopolymerization process and improve mechanical performance. This section contributes to a better understanding of the elements influencing geopolymer qualities by comparing these results and offering insightful information on how to optimize mix designs for enhanced geopolymer performance.

4.2.1 Discussion on the Effects of Molarity

The graph boxplot's results for compressive strength show a definite dependence on the alkaline activator's molarity, highlighting its crucial function in the geopolymerization process. The compressive strength is rather low at lower molarity levels; results like 5.46 MPa & 5.88 MPa suggest that the aluminosilicate precursors are not sufficiently activated. This is because hydroxide ions are scarce, which prevents raw materials from dissolving and

the following polycondensation processes that are required to build a strong geopolymer matrix (Duxson et al., 2007; Rohit et al., 2024).

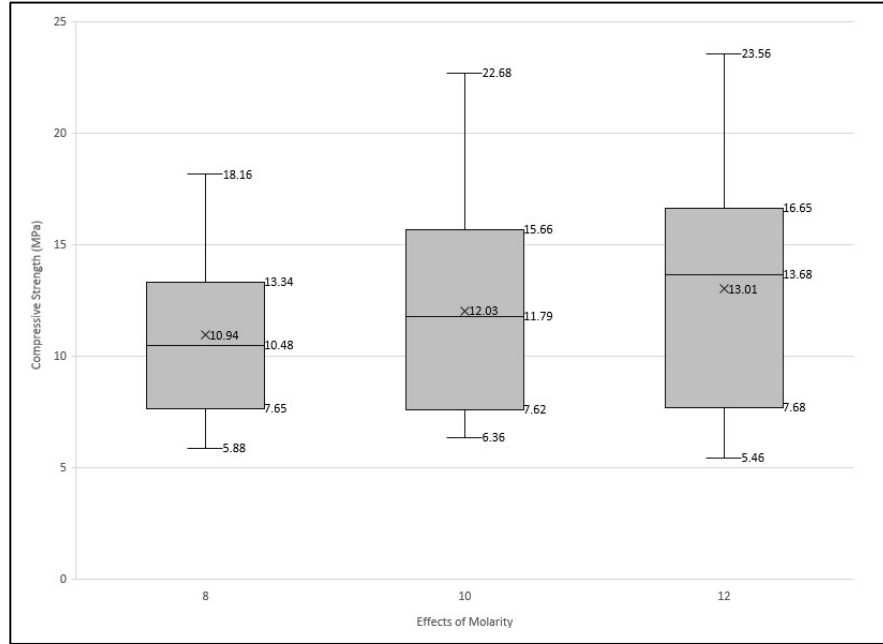


Figure 4.1: Compressive Strength as an Effects of Molarity

A significant rise in compressive strength is shown with increasing molarity, reaching a peak value of 23.56 MPa. According to Provis et al. (2009a), this improvement is in line with the higher concentration of hydroxide ions, which promote the dissolution of silica and alumina species and the development of a denser and more linked geopolymer network. Referring to (Rattanasak et al., 2009), the ideal molarity range seems to provide a compromise between sufficient activation and controllable workability, guaranteeing the formation of a coherent microstructure with exceptional mechanical qualities.

The decrease to 15.66 MPa and lower values at increasing molarities, however, show that the compressive strength starts to decrease beyond this ideal range. The detrimental

consequences of high alkalinity, which can cause quick setting, poor workability, and the development of microcracks within the geopolymer matrix, are probably to blame for this drop (Lloyd et al., 2010). Furthermore, excess sodium carbonate may accumulate as a result of excessive molarity, disrupting the geopolymer network and impairing mechanical performance (Fernandez et al., 2004).

In summary, the findings emphasize how crucial it is to maximize molarity in order to get the required compressive strength in geopolymers. According to Provis et al. (2014), the results are consistent with previous research that highlights the necessity of a balanced alkaline environment for optimizing geopolymer performance. By adjusting mix designs to particular applications, these insights are essential for guaranteeing durability and structural integrity.

4.2.2 Discussion on the Effects of Water/Binder (W/B) Ratio

The compressive strength results presented in the Figure 4.2 reveal a significant influence of the water to binder (W/B) ratio on the mechanical performance of geopolymers. As the W/B ratio increases, a general trend of decreasing compressive strength is observed, consistent with established principles in geopolymer chemistry and concrete technology.

The compressive strength peaks at 26.51 MPa with W/B ratios 0.5, compared to smaller W/B ratio 0.3. The development of a thick and compact geopolymer matrix, where the low water content encourages effective geopolymerization and reduces porosity, is responsible for this high strength (Z. Zhang et al., 2010). A larger concentration of solid

particles is guaranteed by a lower W/B ratio, which promotes stronger interparticle bonds and a more cohesive microstructure (Hardjito, 2004).

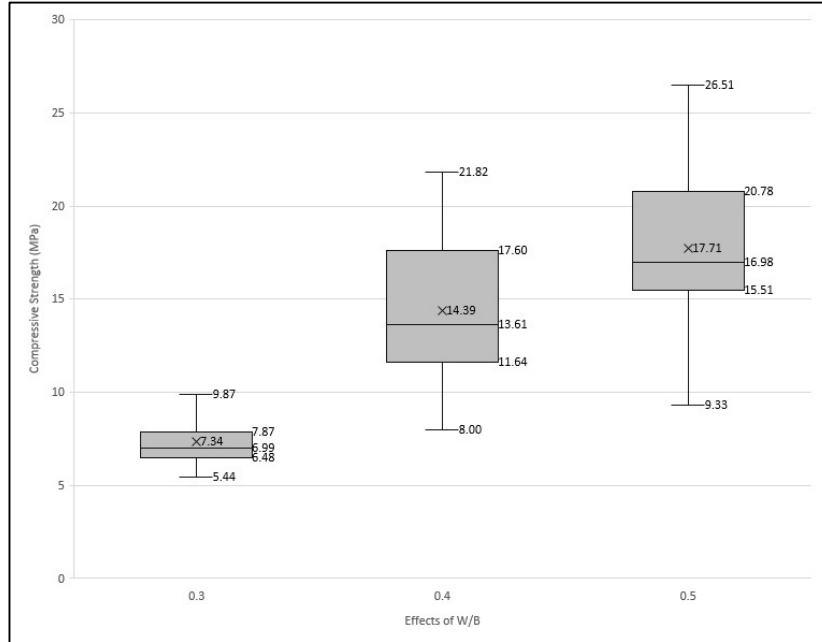


Figure 4.2: Compressive Strength as an Effects of Water to Binder (W/B) Ratio

However, the compressive strength partially decreases when the W/B ratio rises from 0.4 to 0.5 with values falling to 15.51 MPa and 5.44 MPa, respectively. The main cause of this decrease is due to the high-water content, which weakens the interfacial transition zones and raises the porosity of the geopolymer matrix (Kumar et al., 2010). According to Rattanasak et al. (2009), too much water also dilutes the alkaline activator, making it less efficient in dissolving aluminosilicate precursors and impeding the geopolymerization process (Chindaprasirt et al., 2007). Furthermore, segregation and bleeding may result from greater W/B ratios, further jeopardizing the geopolymer's structural integrity (Nematollahi et al., 2014).

The findings are consistent with earlier research, which highlights how crucial it is to maximize the W/B ratio in order to strike a compromise between mechanical performance and workability. Although strength is increased by a lower W/B ratio, workability may be decreased, making the mix challenging to manage. A greater W/B ratio, on the other hand, enhances workability at the price of compressive strength (Rangan et al., 2010). Consequently, choosing the ideal W/B ratio is essential for customizing geopolymer mixtures to meet the demands of particular applications.

4.2.3 Discussion on the Effects of Activator Ratio

According to the experimental results, the activator ratio is a crucial factor affecting the compressive strength of fly ash-based geopolymers. The results show that the activator ratio has a crucial role in influencing the mechanical performance of geopolymer composites, as seen by the wide range of compressive strength values, which vary from 5.46 MPa to 26.51 MPa. Please refer to Figure 4.3 below.

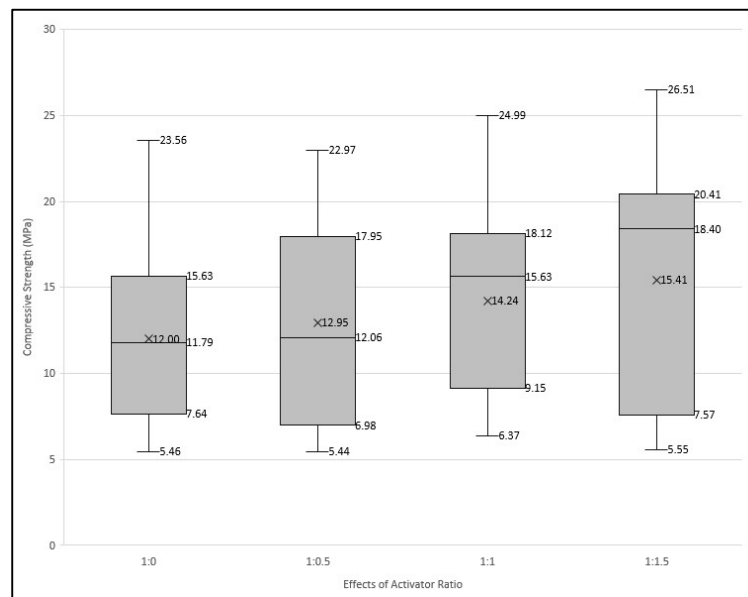


Figure 4.3: Compressive Strength as an Effects of Activator Ratio

The compressive strength peaked at 26.51 MPa with larger activator ratios, which is 1:1.5. This implies that attaining a well-developed geopolymer network requires an ideal activator ratio. A thick and cohesive aluminosilicate matrix may be formed more easily when fly ash particles dissolve more readily due to the higher concentration of alkaline activators at higher ratios (Rangan, 2014). This is consistent with the knowledge that enhanced mechanical characteristics result from the geopolymerization reaction being facilitated by an adequate supply of alkaline ions.

On the other hand, compressive strengths were considerably decreased by smaller activator ratios, especially at ratio 1:0, with some samples showing low compressive strength values (e.g., 5.46 MPa). This phenomenon is explained by inadequate fly ash particle activation, which results in partial geopolymerization and the development of a weak, porous microstructure. The material's structural integrity may be further jeopardized by internal stresses or microcracks, which are indicated by minimal values (Singh et al., 2015). These results demonstrate how the geopolymerization process is negatively impacted by insufficient activator concentrations.

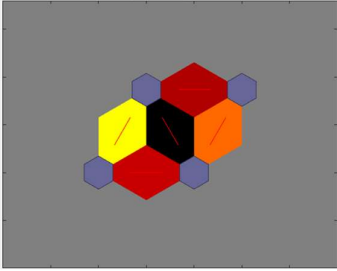
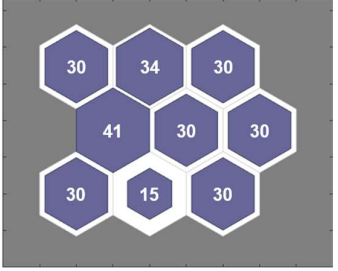
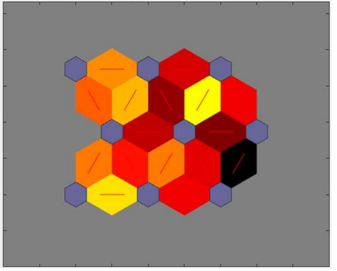
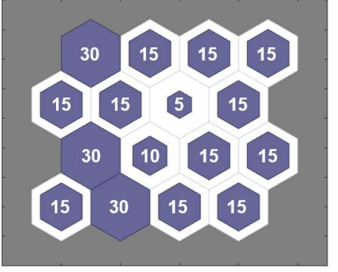
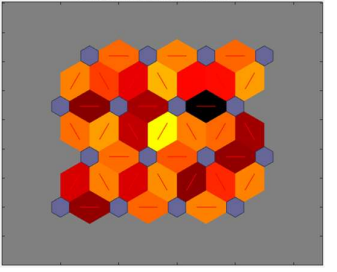
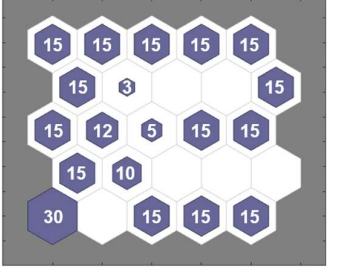
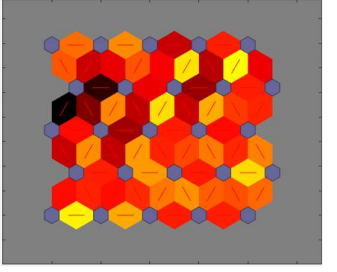
Moderate compressive strengths at value 18.12 MPa were obtained with intermediate activator ratios 1:1, suggesting a balance between material performance and activator concentration. This finding is in line with other research that highlights how crucial it is to maximize the activator ratio in order to strike a balance between mechanical strength and workability (Nath et al., 2014). Higher activator ratios can clearly boost compressive strength, but there is a limit beyond which additional increases could not produce commensurate gains and can even cause other problems like excessive shrinkage or cracking.

To sum everything up, the activator ratio plays a crucial role in the production of fly ash-based geopolymers and directly affects their overall performance and compressive strength. The experimental findings demonstrate that in order to obtain the desired mechanical characteristics, this ratio must be optimized (Ryu et al., 2013). To create more resilient and sustainable geopolymer materials, future studies should concentrate on clarifying the connection between the activator ratio and other factors like curing circumstances and raw material composition.

4.3 SOM Simulation Results

In this analysis, the map size of the Self-Organizing Map (SOM) was set to configurations of 2×2 , 3×3 , 4×4 , and 5×5 nodes. A map size of 2×2 corresponds to four nodes, 3×3 contains nine nodes, 4×4 consists of 16 nodes, and 5×5 includes twenty-five nodes. The structure of the node was set in a hexagonal arrangement, and the results were discussed. The SOM model structures at their respective node numbers and map sizes are listed in Table 1. The first column is the number of nodes, the second column is the topological structure, which explains the number of data sample hits at each hexagonal-shaped node, and the last column is the topological relationship among the nodes, where brighter colors represent highly relevant similarity and darker colors represent otherwise. The nodes are arranged sequentially from the bottom left to the right within each row, progressing in the same order for the subsequent rows from bottom to top. Despite the different numbers of nodes resulting in different visual representations, a higher number of nodes provides more detailed information, which facilitates a better interpretation of the relationships among clusters.

Table 4.1: Model structure of SOM at five parameter settings

No. of nodes	Topological structure	Topological relationship
4 (2x2)	Hits 	SOM Neighbor Weight Distances 
9 (3x3)	Hits 	SOM Neighbor Weight Distances 
16 (4x4)	Hits 	SOM Neighbor Weight Distances 
25 (5x5)	Hits 	SOM Neighbor Weight Distances 

The distribution of data points among the neurons in the SOM is clearly seen by the SOM Hits display. The quantity of data points allocated to each neuron is reflected in its hit

count, which provides information on the dataset's density and clustering tendencies. The information in Table 4.1 illustrates the effectiveness of the model in self-organizing the training data samples into a predetermined number of nodes evenly. For example, a model of four (2×2) nodes has 106 data samples in Cluster 1, 33 data samples in Cluster 2, 74 in Cluster 3, and 57 in Cluster 4. The topological relationship of Cluster 1 was closely related to Cluster 3, but the topological relationship between Clusters 2 and 3 was relatively weak.

To further narrow down the information by changing the model size to model of twenty-five (5×5) nodes, there are clear areas of high and low data density based on the hit numbers. High hit count neurons such as those with values of 15 or 30 indicate regions with a high concentration of data points. These areas most likely reflect the dataset's major patterns or clusters. The neuron with a hit count of 30, for instance, indicates an extremely dense cluster, which might be a big group or a pattern that appears frequently in the data. On the other hand, sparser areas are represented by neurons with lower hit counts, such as those with values of 3, 5, or 10. These regions might be outliers, transitional zones between clusters, or less common patterns. The existence of these neurons draws attention to the dataset's variety and variability, which may be further examined to determine the underlying reasons for these changes.

The topological linkages between clusters may also be inferred from the distribution of hit counts throughout the SOM grid. Cohesive clusters are suggested by neurons that are geographically near to one another and have comparable hit counts. For example, a big homogenous cluster is shown by consecutive neurons with hit counts of 15. Conversely, solitary neurons that have low hit counts (such 3 or 5) might be outliers or unique sub-clusters that don't follow the prevailing trends.

Researchers can rank areas of interest for additional study by determining areas of high and low density. While the sparse areas (neurons with less hit counts of 3 or 5) may require additional inspection to find abnormalities or uncommon patterns, the dense clusters, where neurons with hit counts of 15 or 30 may be investigated to detect common traits or trends. Furthermore, feature selection and model development may be guided by the visualization. While sparse regions could need specific management, such outlier detection or data augmentation, clusters with high hit counts might reveal important patterns that should be included in predictive models.

Table 4.2: Prototype Weight vectors for 2×2

Cluster No.	W_{Molarity}	W_{WBR}	$W_{\text{NaOH(g)}}$	$W_{\text{Silicate(g)}}$	$W_{\text{Water (ml)}}$	$W_{\text{Output (Mpa)}}$
1	0.8707	0.7510	0.8707	0.0000	0.8490	0.4200
2	0.7984	0.6323	0.7984	0.2581	0.6323	0.2966
3	0.8280	0.9774	0.8280	0.0968	0.9774	0.6212
4	0.8333	0.8294	0.8333	0.8529	0.8294	0.6256

Referring to the prototype weights in Table 4.2, it clearly shows that the topological relationship between Clusters 1 and 4 is weak, as the silicate weight for Cluster 1 is zero, while Cluster 4 has a silicate weight of 0.8529. The prototype weights for Clusters 1 and 3 were similar across all parameters, indicating a high degree of similarity between these clusters. In contrast, the topological relationships between Clusters 2 and 3 and between Clusters 3 and 4 were considered weak owing to the significant variation in the silicate weights.

Table 4.3: Prototype Weight vectors for 5×5

Cluster No.	W_{Molarity}	W_{WBR}	$W_{\text{NaOH(g)}}$	$W_{\text{Silicate(g)}}$	$W_{\text{Water (ml)}}$	$W_{\text{Output (Mpa)}}$
1	0.7564	0.6000	0.7564	0.0000	0.6000	0.2787
2	0.7815	0.6178	0.7815	0.1185	0.6178	0.2770

3	0.8333	0.6000	0.8333	0.3333	0.6000	0.2646
4	0.8333	0.6000	0.8333	0.6667	0.6000	0.3013
5	0.8333	0.6000	0.8333	1.0000	0.6000	0.2804
6	0.8333	0.8000	0.8333	0.0000	0.8000	0.4278
7	0.8333	0.8000	0.8333	0.3333	0.8000	0.4216
8	0.8333	0.7037	0.8333	0.5062	0.7037	0.4156
9	0.8333	0.7071	0.8333	0.8333	0.7071	0.4895
10	0.8333	0.7071	0.8333	1.0000	0.7071	0.5181
11	1.0000	0.6000	1.0000	0.0000	1.0000	0.2717
12	0.8333	0.8000	0.8333	0.0000	0.8000	0.6193
13	0.8333	0.8000	0.8333	0.3333	0.8000	0.6392
14	0.8333	0.8000	0.8333	0.6667	0.8000	0.5946
15	0.8333	0.8000	0.8333	1.0000	0.8000	0.7219
16	1.0000	0.8000	1.0000	0.0000	0.8000	0.5184
17	0.6667	0.8000	0.6667	0.0000	0.8000	0.4718
18	0.7623	0.8824	0.7623	0.2647	0.8824	0.5593
19	0.8333	0.9155	0.8333	0.7277	0.9155	0.7128
20	0.8333	1.0000	0.8333	1.0000	1.0000	0.8479
21	1.0000	1.0000	1.0000	0.0000	1.0000	0.6850
22	0.8333	1.0000	0.8333	0.0000	1.0000	0.6401
23	0.6667	1.0000	0.6667	0.0000	1.0000	0.5171
24	0.8333	1.0000	0.8333	0.3333	1.0000	0.7018
25	0.8333	1.0000	0.8333	0.6667	1.0000	0.7174

The implications or information provided by each model can be determined by analyzing their weight representations. Table 4.2 & 4.3 shows the 4 nodes and 25 nodes models at each sample feature in weight vector form, while Table 4.4 shows the model actual information in details. The first column shows the cluster indexes, and the second to sixth columns show the five attributes or features that represent the parameters of the data samples for geopolymers concrete. Referring to Table 4.4, Cluster 1 would produce an expected output of 11.13 MPa with the following materials in given weight accordingly {10.4, 0.38, 4388, 0, 14857}, this corresponds to 10.4 molarity of NaOH, a liquid-binding ratio of 0.38, 4,388 g of sodium hydroxide, 0 g of sodium silicate, and 14,857 ml of water. Each cluster represents unique parameters for geopolymers concrete construction, which is crucial for understanding the structure of given data samples. The model was further analyzed and

labelled with distinct geopolymmer concrete properties at each cluster, enabling the prediction of sample parameters for future use.

Table 4.4: The 4 (2×2 map) clusters of SOM model

Cluster	Molarity	WBR	NaOH(g)	Silicate(g)	Water(ml)	Output (MPa)
1	10.4	0.38	4388	0	14857	11.13
2	9.6	0.32	4024	1626	11065	7.86
3	10	0.49	4173	610	17105	16.47
4	10	0.41	4200	5374	14515	16.59

4.3.1 Simulation Result in 5×5 SOM Map

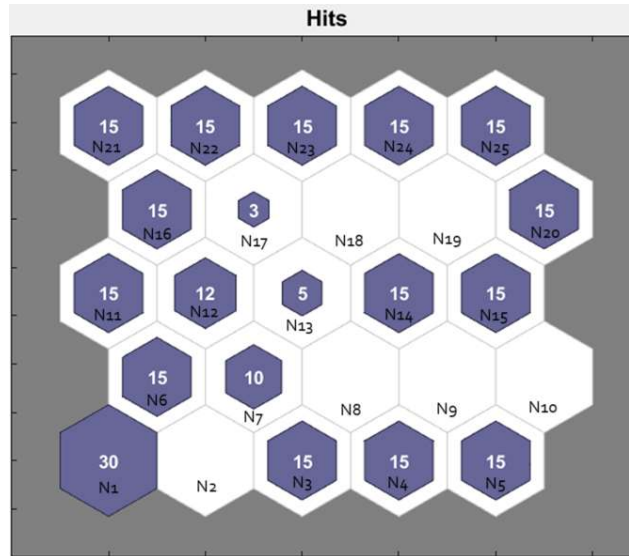


Figure 4.4: Neighbouring Connection in the 5x5 SOM Map

Figure 4.4 illustrates a 5×5 SOM with 25 neurons labeled from N1 to N25. Each neuron represents a group of compressive strengths derived from the various proposed input parameters. Nineteen clusters were populated with sample hits, specifically in neurons N1, N3, N4, N5, N6, N7, N11, N12, N13, N14, N15, N16, N17, N20, N21, N22, N23, N24, and N25. To effectively interpret the hit map, it must be correlated with Table 4.5 below, which was also generated from the SOM simulation.

Table 4.5: The Predicted Parameters Derived from the 25 Clusters of the SOM Model

Cluster	Molarity	WBR	NaOH(g)	Silicate(g)	Water(ml)	Output (MPa)
1	9	0.3	3812	0	10500	7.39
2	9.4	0.3	3939	747	10811	7.34
3	10	0.3	4200	2100	10500	7.01
4	10	0.3	4200	4200	10500	7.99
5	10	0.3	4200	6300	10500	7.43
6	10	0.4	4200	0	14000	11.34
7	10	0.4	4200	2100	14000	11.18
8	10	0.35	4200	3189	12315	11.02
9	10	0.35	4200	5250	12375	12.98
10	10	0.35	4200	6300	12375	13.74
11	12	0.3	5040	0	17500	7.20
12	10	0.4	4200	0	14000	16.42
13	10	0.4	4200	2100	14000	16.95
14	10	0.4	4200	4200	14000	15.76
15	10	0.4	4200	6300	14000	19.14
16	12	0.4	5040	0	14000	13.74
17	8	0.4	3360	0	14000	12.51
18	9.2	0.44	3842	1668	15441	14.83
19	10	0.46	4200	4585	16021	18.90
20	10	0.5	4200	6300	17500	23.48
21	12	0.5	5040	0	17500	18.16
22	10	0.5	4200	0	17500	16.97
23	8	0.5	3360	0	17500	13.71
24	10	0.5	4200	2100	17500	18.61
25	10	0.5	4200	4200	17500	19.02

For instance, neuron N20 contains 15 samples with compressive strengths ranging between 23.01 MPa and 24.45 MPa. This forecasting enables civil practitioners to leverage these clusters to identify the appropriate mix proportion to achieve the desired concrete compressive strength. Meanwhile, neurons N2, N8, N9, N10, N18, and N19, which show zero sample hits, indicate a lack of proposals and can be interpreted as new or unexperimented information. These neurons suggest an entirely new mix proportion for practitioners, potentially yielding comparable compressive strengths, as outlined in Table 4.5 (highlighted in color). This iterative approach, which combines SOM predictions with

experimental validation, can significantly accelerate the optimization process for geopolymer concrete formulations.

Drawing from the SOM predicted data, the following key insights and suggestions can be made.

1. **SOM Efficacy:** SOM has demonstrated its value as an effective method for clustering geopolymer concrete data and visualizing intricate relationships on a two-dimensional plane. This feature enables intuitive interpretation of parameter interactions and facilitates efficient exploration of extensive parameter spaces.
2. **Ideal Mixture Formulation:** SOM analysis indicated an optimal mix design with the following specifications: NaOH Concentration of 10M, Water/binder ratio of 0.5, NaOH quantity of 4200 g, and Na₂SiO₃ of 6300 g. This formulation was anticipated to yield a compressive strength of 23.48 MPa.

4.4 Verification and Validation

The SOM model was verified and validated to ensure reliability and accuracy. Verification focused on confirming that the SOM model was implemented correctly and operated as intended, aligning with the expected theoretical framework. Validation evaluates the ability of the model to accurately cluster data and predict compressive strengths by comparing its outputs with experimental results. This dual process reinforces the confidence in the applicability of the SOM model to geopolymer concrete data analysis and optimization.

4.4.1 Ground Truth for Mix Proportion Verification

The two-mix proportion recommended by 25 clustering out from SOM was chosen for verification in order to test the ground truth. One mix proportion is the exact same formula that has previously been used for training, while the other is a brand-new recipe that SOM has suggested. Model predictions are compared to ground truth to assess how well the model generalizes. This is essential for evaluating the reliability of software systems and determining whether software corresponds to its specifications and requirements. Clusters 18 and 20 were chosen from the suggested SOM clustering list and were mixed according to the mix proportion provided in Table 4.6.

Table 4.6: Mix Proportion Given by Clusters of SOM Model

Cluster	Molarity	WBR	NaOH(g)	Silicate(g)	Water(ml)	Output (MPa)
18	9.2	0.44	3842	1668	15441	14.83
20	10	0.5	4200	6300	17500	23.48

Both mixtures were prepared with 30 samples each using the same mixing method and adhering to all parameters specified by SOM clustering. This approach ensured that the SOM outcomes were consistent and reproducible. Figure 4.5 illustrates the results of the verification tests, which were uplifted and promising. For Cluster 18, the majority of the results were comparable to the baseline value of 14.83 MPa. Similarly, the overall values obtained from Cluster 20 were slightly higher than the baseline result of 23.48 MPa. Both ground truth results obtained are very promising and precise to the simulation results.

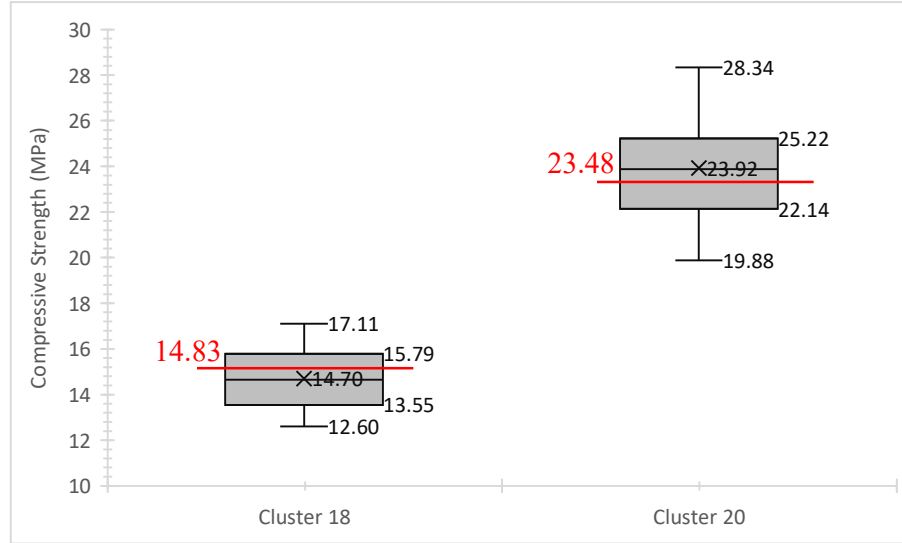


Figure 4.5: Ground Truth Verification Results

4.4.2 SOM Training Validation

A five-fold cross validation procedure was used to verify the repeatability and dependability of the parameters of the SOM generated clusters that were obtained from the clustering study. This method also assessed how well the suggested activator doses and mixing patterns predicted compressive strength. Because it reduces overfitting and offers a reliable measure of model performance, cross-validation is a commonly used technique for evaluating the generalizability of predictive models (Refaeilzadeh et al., 2009). The dataset used in this study was divided into five equal subsets, or "folds." Each fold was used as the test set once, and the model was trained on four folds and tested on the remaining fold. This process was done five times. This guarantees that each and every data point is used for testing and training, resulting in a more accurate assessment of the model's predictive power. A thorough evaluation of the model's accuracy and consistency was provided by the final performance measures, which were calculated as the average of the five iterations.

Figure 4.6 shows that the model performs well in terms of prediction. The anticipated and experimental compressive strength values appear to be closely aligned, with variances usually falling within a realistically acceptable range, as indicated by the low root mean square error (RMSE=2.6862) and mean absolute error (MAE=2.3987). The model's dependability is further reinforced by the mean squared error (MSE=7.2156), which shows a moderate number of squared errors. Furthermore, a significant amount of the underlying data trends is captured by the model, which explains for 70% of the variability in the target variable, according to the coefficient of determination ($R^2=0.7$). While the unexplained variance (30%) may reflect inherent complexities in geopolymers or unmodeled factors, the collective metrics RMSE, MAE, and R^2 affirm the model's utility for preliminary strength prediction in fly ash-based geopolymers.

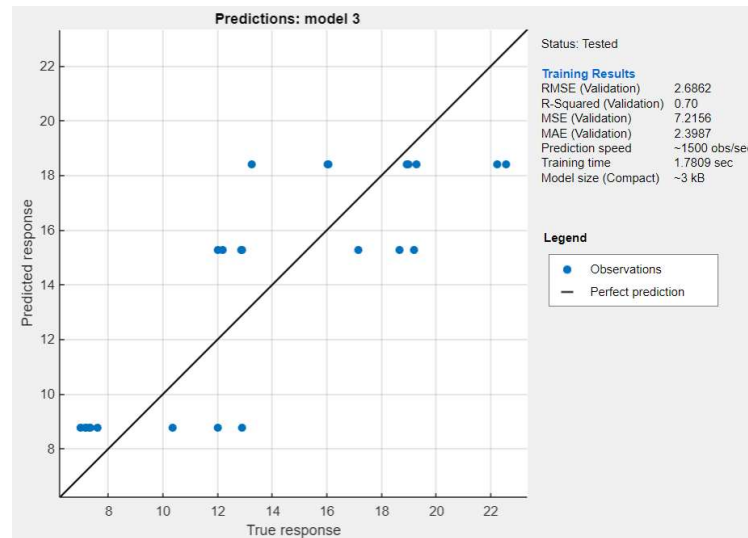


Figure 4.6: SOM Validation Training Results

In order to minimize residual errors for higher-precision applications, future research may investigate feature optimization or hybrid modeling techniques. The model's mean absolute error (MAE) is 2.3987, which translates to a mean absolute percentage error

(MAPE) of 6.5%. This means that, on average, predictions differ by 6.5% from the experimental results. The model's applicability for estimating preliminary strength in geopolymers is made apparent by its low relative error, 2.6862 RMSE, and 0.7 R^2 .

The 'True response' in the result refers to the actual observed or target values (ground truth) from the validation dataset, which are compared against the model's predicted values to compute metrics like RMSE, R-Squared, etc. The legend may have caused confusion, but it just highlights the ideal vs. actual performance of the SOM training validation. This study's performance indicators are in good agreement with those found in previous scholarly publications. Chou et al. (2014), for example, used machine learning approaches to estimate concrete's compressive strength and reported an R^2 value of 0.68 and an RMSE of 3.21. The new model, in contrast, shows better prediction accuracy with a lower RMSE (2.6862) and a higher R^2 (0.7). Similarly, the MAE of 2.3987 obtained in this investigation is lower than the MAE of 2.75 reported by Atici (2011) in their study on concrete strength prediction. These comparisons imply that the suggested model provides improved accuracy and resilience.

When compared to contemporary research, the MSE score of 7.2156 is likewise competitive. For instance, Bilgehan et al. (2010) reported an MSE of 8.45 for their neural network-based concrete strength prediction model. The study's decreased MSE demonstrates how well the SOM clustering strategy and the thorough cross-validation procedure work to improve model performance. Additionally, the model performs within the predicted range for prediction tasks in this domain, as seen by the R^2 value of 0.7, which is within the range of 0.6 to 0.8 often seen in similar research (Yeh, 2007).

However, one must recognize that direct comparisons with other research are sometimes hampered by variations in model designs, preprocessing techniques, and dataset characteristics. For example, performance measurements may be impacted by the use of more sophisticated algorithms or bigger datasets in some studies (Gandomi et al., 2011). Notwithstanding these discrepancies, the study's findings show that the suggested model is dependable and efficient, with performance metrics that are on par with or better than those documented in the body of current research.

The RMSE, MSE, R2, and MAE values demonstrate the model's good performance, which highlights its potential for real-world uses in compressive strength prediction and mixing design optimization. These results are consistent with the increasing amount of research supporting the application of clustering and machine learning methods in material science and engineering (Nguyen et al., 2021). To further improve accuracy and generalizability, future research should investigate other model improvements including adding more characteristics, experimenting with other algorithms, or using ensemble approaches.

Although the remaining unexplained variation is probably due to the intrinsic complexity of geopolymer systems, where it is difficult to properly capture the nonlinear interactions between mixing composition, activator, and curing conditions, this ANN model achieves a respectable prediction accuracy ($R^2 = 0.7$). Scale and variety constraints in the dataset may lead to underfitting, and model performance is further limited by the omission of hard-to-measure microstructural characteristics such as gel porosity and nano-structural development, which are known to have a major impact on strength. Results may potentially be impacted by possible ANN architecture limitations as well as experimental heterogeneity

in sample preparation and testing. Notwithstanding these shortcomings, the model's performance compares well with earlier geopolymer research, for example, equal or surpassing Yeh (2007) which $R^2 = 0.65$, indicating adequate accuracy for initial mix design applications. Precision for structural-grade applications may be increased in the future by using improved characterisation data or using hybrid modeling.

During development, the 5-fold cross-validation strategy was utilized to test the model on different subsets of data that were not used for training in order to confirm its generalizability. Consistent performance measures from this approach indicated a respectable level of resilience to fresh data with comparable compositional ranges. The cross-validation findings and comparison with the literature indicate that the model should generalize well to comparable systems, even though this work did not examine external datasets because of data availability restrictions. Working with industry partners to assess performance using actual production data would be the next stage.

4.4.3 Model Validation Through Hierarchical SOM Analysis

The same dataset was used to simulate data using the HSOM algorithm. In contrast to SOM output, which displays all possible outcomes in a single tabular following the specified nodes and total clusters required, hierarchical SOM simulation generates results at different levels based on the closed relationship pattern. As shown in Figure 4.7, the input data is trained and output dataset is splitting into four distinct levels that combined all patterns that were identical.

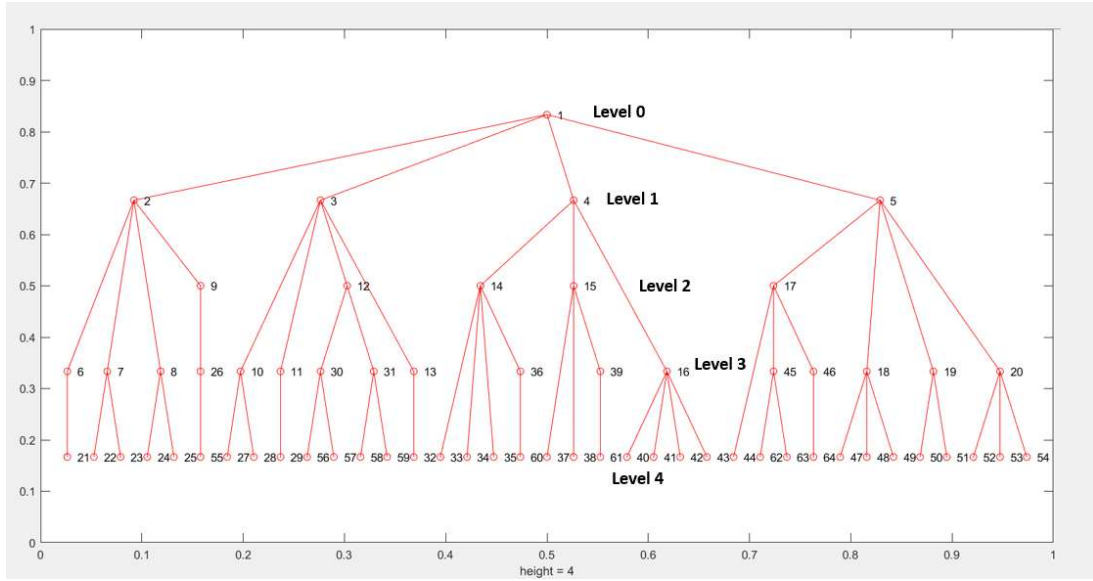


Figure 4.7: HSOM Validation Training Results

Both Figure 4.8 and Figure 4.9 are explaining how the splitting data from level 0 narrow down to the lower-level cluster. Each node in the HSOM graph represents a specific group of data points. The nodes are arranged on a grid, and similar data points are grouped together. As it moves from higher layers (coarse-level) to lower layers (fine-level), the nodes represent more specific clusters or patterns. These figures provide an intuitive way to understand complex data relationships by organizing them into a hierarchical structure. The structure is particularly useful for multi-dimensional datasets, as they help identify patterns at various levels of abstraction. In these concrete studies, HSOM helps to categorize different concrete mix designs based on their properties and performance, allowing for more refined predictions and optimizations.

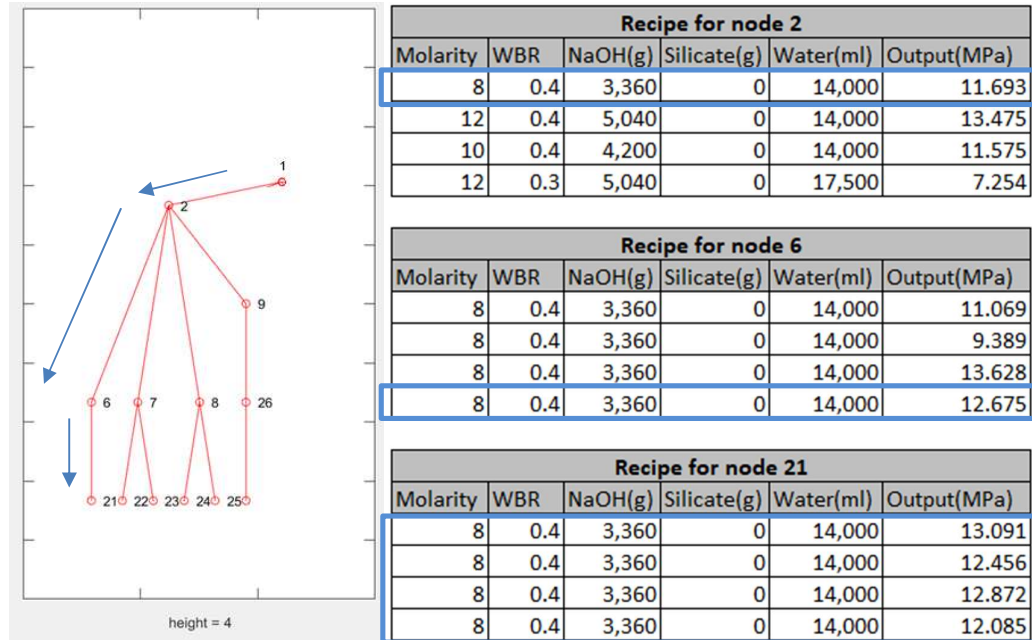


Figure 4.8: HSOM results into different sub-levels (a)

The splitting data from level 0 narrow down to the lower-level cluster is shown in Figure 4.8. Nodes 6, 7, 8, and 9 are the four subclusters for node 2 of level 1, and they continue to display distinct combinations. Because their output ranges are less comparable, nodes 6, 7, and 8 are in level 3, whereas node 9 is in level 2. In order to find commodities in terms of mix proportion, Node 6 is then further subdivided to level 4. The absolute closed combination is displayed by node 21, which is the only node for the lowest sublevel. As a result, there is no way to further refine.

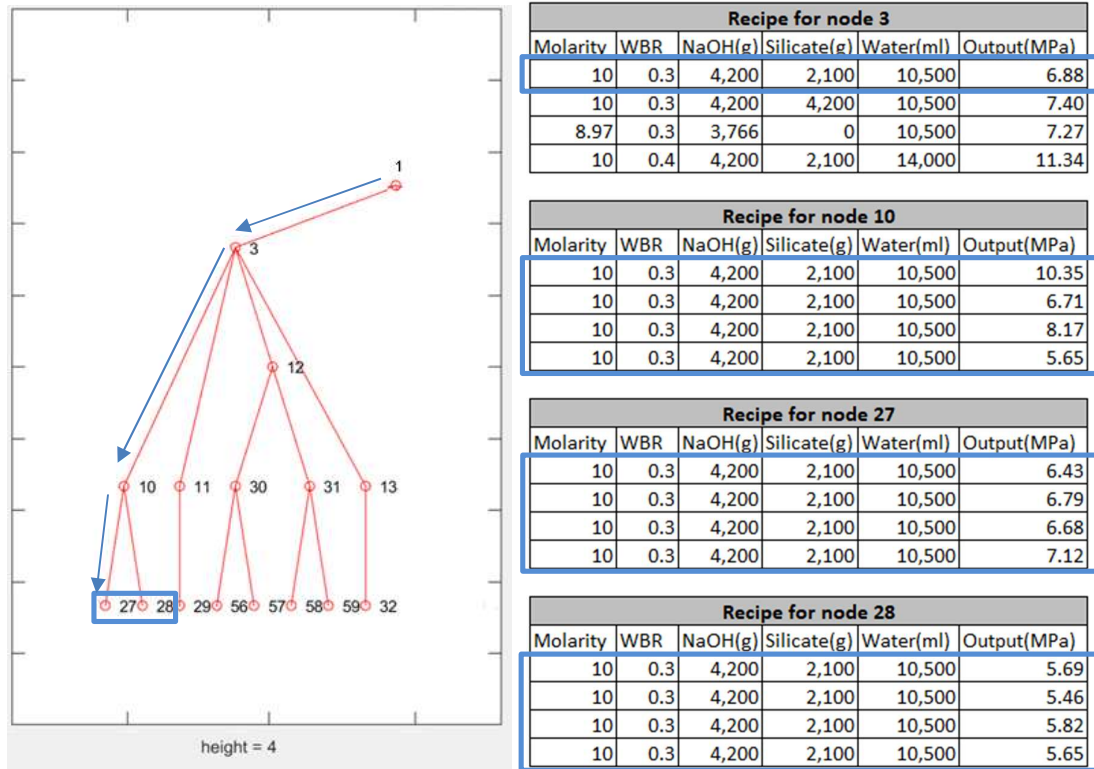


Figure 4.9: HSOM results into different sub-levels (b)

Figure 4.9 shows the data split from level 0 narrow down to the lower level 4, just like Figure 4.8 does. Node 12 is at level 2 alone since it has a newly created recipe from HSOM, while nodes 10, 11, and 13 are at level 3. Node 10 was chosen to show the trends, and at level 4, it further divided into two nodes, node 27 and 28, respectively. Compressive strength at node 27 ranges from 5.46 MPa to 5.82 MPa, whereas node 28 vary from 6.43 MPa to 7.12 MPa.

Figure 4.10 shows the same breaking down procedure. What distinguishes this sub level from the previous two narrowing down processes is that the mix proportions are the same under same line. However, at sub level 4, although having the same mix proportion, the result clusters the nearest four outcomes to nodes 40, 41, 42, and 43 respectively.

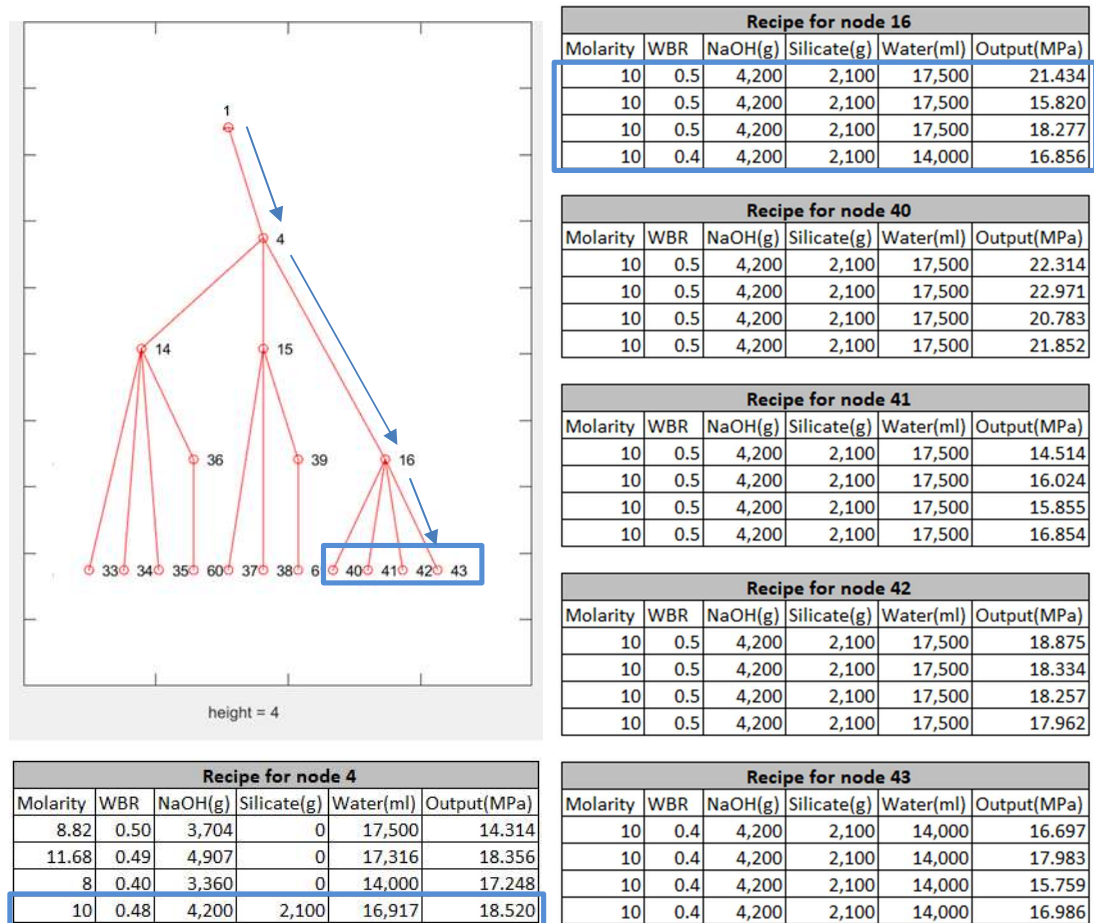


Figure 4.10: HSOM results into different sub-levels (c)

Similar to Figure 4.11 below, the dataset is segregated to node 5 at level 1, then further broken down to node 17 at level 2, and nodes 18, 19, 20 at level 3. Node 19 was picked from level 3 to break down into nodes 50 and 51 of the lowest level 4. These two nodes demonstrate the exact same mix proportion, but at a different range of output that is near to each other.

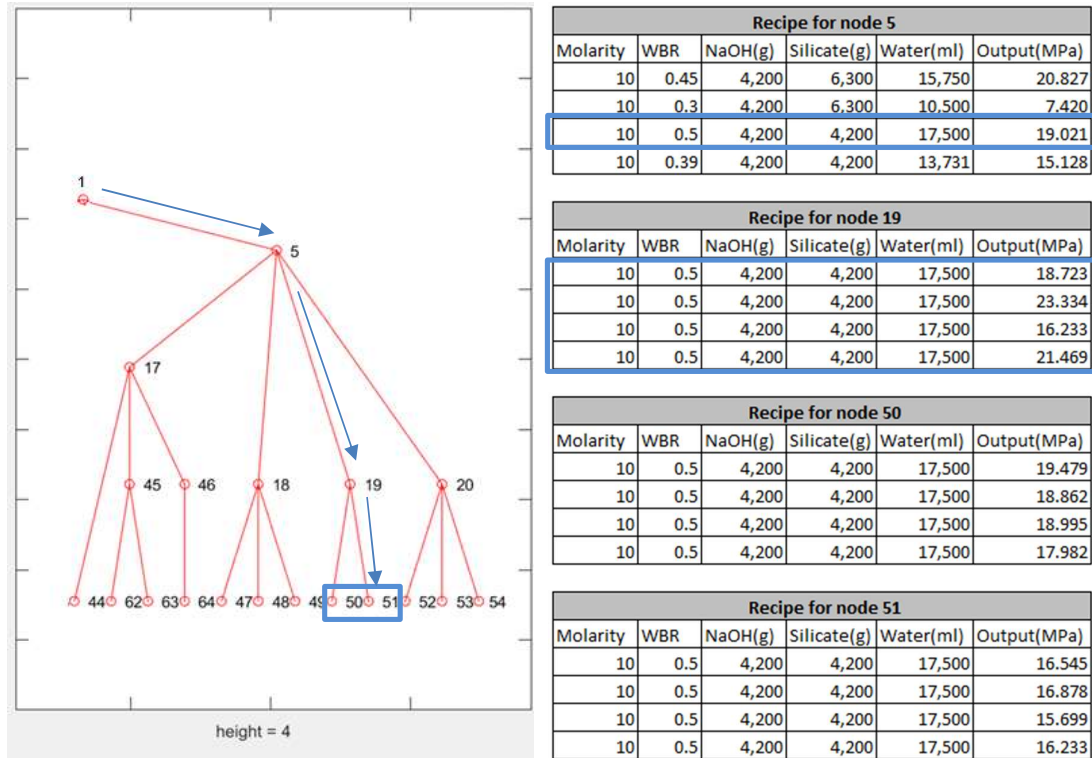


Figure 4.11: HSOM results into different sub-levels (d)

In summary, the HSOM divides the mapping process into multiple hierarchical layers, where each layer represents a subset of the data at increasing levels of granularity. In a higher layer, every neuron (node) functions as a parent and is linked to a smaller SOM (sub-map) in the layer below. The submap learns to arrange data according to the receptive field of its parent. The first layer (top level) is trained on the entire dataset to provide coarse-grained clustering. The subsequent layers are trained only on the data points assigned to their respective parent neurons, creating finer-grained clusters. This hierarchical division continues until a predefined depth is reached or the data cannot be further subdivided meaningfully.

4.4.4 Results Comparison among SOM & HSOM

In this study, the HSOM preserved the topological relationships in the data, ensuring that similar data points remained close in the map across layers. This was validated by comparing the predictive output results of SOM and HSOM, as shown in Figure 4.12 & Figure 4.13.

Cluster	Molarity	WBR	NaOH (g)	Silicate (g)	Water (ml)	Output (MPa)
1	9	0.3	3,812	0	10,500	7.389
2	9.4	0.3	3,939	747	10,811	7.345
3	10	0.3	4,200	2,100	10,500	7.014
4	10	0.3	4,200	4,200	10,500	7.989
5	10	0.3	4,200	6,300	10,500	7.434
6	10	0.4	4,200	0	14,000	11.342
7	10	0.4	4,200	2,100	14,000	11.177
8	10	0.35	4,200	3,189	12,315	11.020
9	10	0.35	4,200	5,250	12,375	12.976
10	10	0.35	4,200	6,300	12,375	13.736
11	12	0.3	5,040	0	17,500	7.204
12	10	0.4	4,200	0	14,000	16.418
13	10	0.4	4,200	2,100	14,000	16.946
14	10	0.4	4,200	4,200	14,000	15.764
15	10	0.4	4,200	6,300	14,000	19.139
16	12	0.4	5,040	0	14,000	13.744
17	8	0.4	3,360	0	14,000	12.509
18	9.2	0.44	3,842	1,668	15,441	14.828
19	10	0.46	4,200	4,585	16,021	18.897
20	10	0.5	4,200	6,300	17,500	22.480
21	12	0.5	5,040	0	17,500	18.160
22	10	0.5	4,200	0	17,500	16.970
23	8	0.5	3,360	0	17,500	13.710
24	10	0.5	4,200	2,100	17,500	18.610
25	10	0.5	4,200	4,200	17,500	19.020

Figure 4.12: Predicted Compressive Strength from SOM

Recipe for node 5					
Molarity	WBR	NaOH(g)	Silicate(g)	Water(ml)	Output(MPa)
10	0.45	4,200	6,300	15,750	20.827
10	0.3	4,200	6,300	10,500	7.420
10	0.5	4,200	4,200	17,500	19.021
10	0.39	4,200	4,200	13,731	15.128

Recipe for node 19					
Molarity	WBR	NaOH(g)	Silicate(g)	Water(ml)	Output(MPa)
10	0.5	4,200	4,200	17,500	18.723
10	0.5	4,200	4,200	17,500	23.334
10	0.5	4,200	4,200	17,500	16.233
10	0.5	4,200	4,200	17,500	21.469

Recipe for node 50					
Molarity	WBR	NaOH(g)	Silicate(g)	Water(ml)	Output(MPa)
10	0.5	4,200	4,200	17,500	19.479
10	0.5	4,200	4,200	17,500	18.862
10	0.5	4,200	4,200	17,500	18.995
10	0.5	4,200	4,200	17,500	17.982

Recipe for node 51					
Molarity	WBR	NaOH(g)	Silicate(g)	Water(ml)	Output(MPa)
10	0.5	4,200	4,200	17,500	16.545
10	0.5	4,200	4,200	17,500	16.878
10	0.5	4,200	4,200	17,500	15.699
10	0.5	4,200	4,200	17,500	16.233

Figure 4.13: Predicted Compressive Strength from HSOM

Referring to Figure 4.12 and Figure 4.13, both SOM and HSOM simulations produced identical mix proportions, yielding a similar compressive strength output of approximately 19 MPa. These matching results reinforce the validity of the ground truth

verification using the same mix proportion. To further confirm the predictive accuracy of the SOM and HSOM model, a verification mix was conducted. The outputs predicted by both SOM and HSOM were then compared against the actual experimental results. A boxplot as shown in Figure 4.14 was generated to visually represent this comparison, displaying the distribution of ground truth compressive strength values alongside the predicted values from both models. The figure highlights how closely each model's predictions align with the experimental results, demonstrating their reliability.

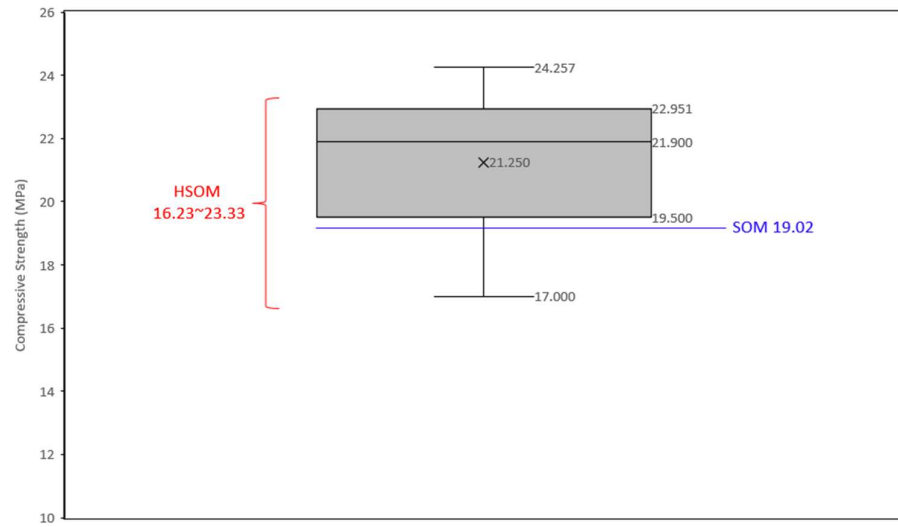


Figure 4.14: Ground Truth Verification vs HSOM & SOM Predictive Results

4.5 Chapter Summary

The ground truth results closely match the values generated from the simulated mix proportions, confirming the reliability of the experimental setup. When comparing the predictive outputs of the SOM and HSOM models, the results appear to be virtually identical. This outcome is highly encouraging, as it reinforces the robustness of the modeling approach and validates the extensive work carried out throughout this study.

4.6 Conclusion

This study demonstrates the effectiveness of Self-Organizing Maps (SOM) and their hierarchical variants (HSOM) in predicting the compressive strength of Geopolymer Concrete (GPC), which is a critical parameter for ensuring structural integrity. By analyzing diverse mix designs and activator dosages, these techniques identified nonlinear relationships and clustered GPC compositions, enabling efficient strength prediction without resource-intensive experimentation. Traditional trial-and-error methods, which are time-consuming and costly, can be circumvented through this data-driven approach, improving both cost efficiency and scalability.

Computational demands vary significantly between SOM and HSOM models. While SOMs exhibit linear scaling with dataset size, HSOMs incur higher costs due to their hierarchical architecture, often requiring GPU acceleration or cloud-based solutions for large-scale datasets, usually larger than 50,000 samples. Key challenges include the curse of dimensionality when incorporating advanced material characterization data and the need for periodic retraining with new material sources. These limitations can be mitigated through strategies such as dimensionality reduction, transfer learning, and edge deployment. For industrial adoption, SOMs offer a cost-effective balance of performance and scalability, whereas HSOMs excel in modeling complex hierarchical relationships but demand greater computational investment. Critically, computational feasibility depends more on data quality and preprocessing than on raw dataset size, underscoring the importance of robust data pipelines.

Both models were thoroughly validated through 5-fold cross-validation, demonstrating reliable generalization with $R^2=0.7$, and low RMSE/MAE. Their

bidirectional predictive capabilities, validated through parallel implementation, underscore their potential to revolutionize mix design paradigms. By minimizing reliance on trial-and-error methods, these approaches provide a transformative framework for optimizing GPC formulations in both research and industrial settings.

4.7 Recommendations

Based on the successful application of SOM and HSOM in predicting the GPC strength, it is recommended that engineers adopt this methodology to streamline the mix design process for Geopolymer Concrete. Future research could explore the integration of additional factors, such as environmental conditions or curing methods, to further refine the predictions. Additionally, extending the dataset to include a broader range of mixed designs and activator dosages could improve the generalizability of the model. Finally, implementing a hybrid model combining SOM with other machine-learning algorithms may enhance predictive accuracy and robustness.

REFERENCES

- Abdila, S. R., Abdullah, M. M. A. B., Ahmad, R., Nergis, D. D. B., Rahim, S. Z. A., Omar, M. F., Sandu, A. V., Vizureanu, P., & Syafwandi. (2022). Potential of Soil Stabilization Using Ground Granulated Blast Furnace Slag (GGBFS) and Fly Ash via Geopolymerization Method: A Review. In *Materials* (Vol. 15, Issue 1). MDPI. <https://doi.org/10.3390/ma15010375>
- Abdullah, M. M. A., Kamarudin, H., Mohammed, H., Khairul Nizar, I., Rafiza, A. R., & Zarina, Y. (2011). The relationship of NaOH molarity, Na₂SiO₃/NaOH ratio, fly ash/alkaline activator ratio, and curing temperature to the strength of fly ash-based geopolymer. *Advanced Materials Research*, 328–330, 1475–1482. <https://doi.org/10.4028/www.scientific.net/AMR.328-330.1475>
- Abiodun, O. I., Jantan, A., Omolara, A. E., Dada, K. V., Umar, A. M., Linus, O. U., Arshad, H., Kazaure, A. A., Gana, U., & Kiru, M. U. (2019). Comprehensive Review of Artificial Neural Network Applications to Pattern Recognition. In *IEEE Access* (Vol. 7, pp. 158820–158846). Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/ACCESS.2019.2945545>
- Ahmad, A., Ahmad, W., Aslam, F., & Joyklad, P. (2022). Compressive strength prediction of fly ash-based geopolymer concrete via advanced machine learning techniques. *Case Studies in Construction Materials*, 16. <https://doi.org/10.1016/j.cscm.2021.e00840>
- Ahmad, A., Ahmad, W., Chaiyasarn, K., Ostrowski, K. A., Aslam, F., Zajdel, P., & Joyklad, P. (2021a). Prediction of geopolymer concrete compressive strength using novel machine learning algorithms. *Polymers*, 13(19).

<https://doi.org/10.3390/polym13193389>

Ahmad, A., Ahmad, W., Chaiyasarn, K., Ostrowski, K. A., Aslam, F., Zajdel, P., & Joyklad, P. (2021b). Prediction of geopolymer concrete compressive strength using novel machine learning algorithms. *Polymers*, *13*(19).

<https://doi.org/10.3390/polym13193389>

Ahmad, M., Rashid, K., Tariq, Z., & Ju, M. (2021). Utilization of a novel artificial intelligence technique (ANFIS) to predict the compressive strength of fly ash-based geopolymer. *Construction and Building Materials*, *301*, 124251.

<https://doi.org/https://doi.org/10.1016/j.conbuildmat.2021.124251>

Ahmari, S., & Zhang, L. (2013). Utilization of cement kiln dust (CKD) to enhance mine tailings-based geopolymer bricks. *Construction and Building Materials*, *40*, 1002–1011. <https://doi.org/https://doi.org/10.1016/j.conbuildmat.2012.11.069>

Ankur Bhogayata and Rinkesh Makwana, S. K. (n.d.). Neural Network for Mixture Design Optimization of Geopolymer Concrete. *ACI Materials Journal*, *118*(4).

<https://doi.org/10.14359/51732711>

Arioz, E., Arioz, O., & Mete Kockar, O. (2012). An Experimental Study on the Mechanical and Microstructural Properties of Geopolymers. *Procedia Engineering*, *42*, 100–105. <https://doi.org/10.1016/J.PROENG.2012.07.399>

Asan, U., & Ercan, S. (2012). *An Introduction to Self-Organizing Maps* (pp. 295–315). https://doi.org/10.2991/978-94-91216-77-0_14

- Atici, U. (2011). Prediction of the strength of mineral admixture concrete using multivariable regression analysis and an artificial neural network. *Expert Systems with Applications*, 38(8), 9609–9618. <https://doi.org/10.1016/j.eswa.2011.01.156>
- Awoyera, P. O., Kirgiz, M. S., Vilorio, A., & Ovallos-Gazabon, D. (2020). Estimating strength properties of geopolymer self-compacting concrete using machine learning techniques. *Journal of Materials Research and Technology*, 9(4), 9016–9028. <https://doi.org/10.1016/j.jmrt.2020.06.008>
- Aziz, I. H., Abdullah, M. M. A. B., Mohd Salleh, M. A. A., Azimi, E. A., Chaiprapa, J., & Sandu, A. V. (2020). Strength development of solely ground granulated blast furnace slag geopolymers. *Construction and Building Materials*, 250, 118720. <https://doi.org/10.1016/J.CONBUILDMAT.2020.118720>
- Ben Chaabene, W., Flah, M., & Nehdi, M. L. (2020). Machine learning prediction of mechanical properties of concrete: Critical review. *Construction and Building Materials*, 260, 119889. <https://doi.org/10.1016/j.conbuildmat.2020.119889>
- Bilgehan, M., & Turgut, P. (2010). The use of neural networks in concrete compressive strength estimation. *Computers and Concrete*, 7(3), 271–283. <https://doi.org/10.12989/cac.2010.7.3.271>
- Brough, A. R., & Atkinson, A. (2002). Sodium silicate-based, alkali-activated slag mortars - Part I. Strength, hydration and microstructure. *Cement and Concrete Research*, 32(6), 865–879. [https://doi.org/10.1016/S0008-8846\(02\)00717-2](https://doi.org/10.1016/S0008-8846(02)00717-2)

- Burduhos, D. D., Abdullah, A. B., Vizureanu, P., & Tahir, F. M. (2018). *Geopolymers and Their Uses: Review*. <https://doi.org/10.1088/1757-899X/374/1/012019>
- Chang, W. L., Pang, L. M., & Tay, K. M. (2017). Application of self-organizing map to failure modes and effects analysis methodology. *Neurocomputing*, 249, 314–320. <https://doi.org/10.1016/j.neucom.2016.04.073>
- Chavan, S., Joshi, G., & Hake, S. (2017). *EFFECT OF ALKALINE ACTIVATOR ON THE STRENGTH OF GEOPOLYMER CONCRETE* (Vol. 3). www.ijariie.com
- Chen, K., Wu, D., Xia, L., Cai, Q., & Zhang, Z. (2021). Geopolymer concrete durability subjected to aggressive environments – A review of influence factors and comparison with ordinary Portland cement. *Construction and Building Materials*, 279, 122496. <https://doi.org/10.1016/J.CONBUILDMAT.2021.122496>
- Chen, Q., Hu, G., & Wu, J. (2024). Prediction of the Unconfined Compressive Strength of a One-Part Geopolymer-Stabilized Soil Using Deep Learning Methods with Combined Real and Synthetic Data. *Buildings*, 14(9). <https://doi.org/10.3390/buildings14092894>
- Chindaprasirt, P., Chareerat, T., & Sirivivatnanon, V. (2007). Workability and strength of coarse high calcium fly ash geopolymer. *Cement and Concrete Composites*, 29(3), 224–229. <https://doi.org/https://doi.org/10.1016/j.cemconcomp.2006.11.002>
- Chou, J.-S., Tsai, C.-F., Pham, A.-D., & Lu, Y.-H. (2014). Machine learning in concrete strength simulations: Multi-nation data analytics. *Construction and Building Materials*, 73, 771–780. <https://doi.org/10.1016/j.conbuildmat.2014.09.054>
- Chu, H. H., Khan, M. A., Javed, M., Zafar, A., Ijaz Khan, M., Alabduljabbar, H., & Qayyum, S. (2021). Sustainable use of fly-ash: Use of gene-expression programming (GEP) and multi-expression programming (MEP) for forecasting the compressive strength

- geopolymer concrete. *Ain Shams Engineering Journal*, 12(4), 3603–3617.
<https://doi.org/10.1016/j.asej.2021.03.018>
- Criado, M., Fernández-Jiménez, A., & Palomo, A. (2007). Alkali activation of fly ash: Effect of the SiO₂/Na₂O ratio: Part I: FTIR study. *Microporous and Mesoporous Materials*, 106(1), 180–191. <https://doi.org/https://doi.org/10.1016/j.micromeso.2007.02.055>
- Dao, D., Ly, H.-B., Trinh, S., Le, T.-T., & Pham, B. (2019). Artificial Intelligence Approaches for Prediction of Compressive Strength of Geopolymer Concrete. *Materials*, 12(6), 983. <https://doi.org/10.3390/ma12060983>
- Davidovits, J. (1994). Properties of Geopolymer Cements. *First International Conference on Alkaline Cements and Concretes*, 131–149.
- Davidovits, J. (2011). *Geopolymer Chemistry and Applications*. Institut Géopolymère.
<https://books.google.com.my/books?id=mJBotwAACAAJ>
- Dike, H. U., Zhou, Y., Deveerasetty, K. K., & Wu, Q. (2018). Unsupervised Learning Based On Artificial Neural Network: A Review. *2018 IEEE International Conference on Cyborg and Bionic Systems, CBS 2018*, 322–327.
<https://doi.org/10.1109/CBS.2018.8612259>
- Dongare, A. D., Kharde, R. R., & Kachare, A. D. (2008). Introduction to Artificial Neural Network. In *Certified International Journal of Engineering and Innovative Technology (IJEIT)* (Vol. 9001, Issue 1).
- Duxson, P., Fernández-Jiménez, A., Provis, J. L., Lukey, G. C., Palomo, A., & Van Deventer, J. S. J. (2007). Geopolymer technology: The current state of the art. *Journal of Materials Science*, 42(9), 2917–2933. <https://doi.org/10.1007/s10853-006-0637-z>

- Feng, D. C., Liu, Z. T., Wang, X. D., Chen, Y., Chang, J. Q., Wei, D. F., & Jiang, Z. M. (2020). Machine learning-based compressive strength prediction for concrete: An adaptive boosting approach. *Construction and Building Materials*, 230, 117000. <https://doi.org/10.1016/j.conbuildmat.2019.117000>
- Fernández-Jiménez, A., & Palomo, A. (2003). Characterisation of fly ashes. Potential reactivity as alkaline cements. *Fuel*, 82, 2259–2265. [https://doi.org/10.1016/S0016-2361\(03\)00194-7](https://doi.org/10.1016/S0016-2361(03)00194-7)
- Fernandez-Jimenez, A., & Palomo, A. (2004). Activation of fly ash: a general view. *ACI Special Publication*, 221, 351–366.
- Gandomi, A., Tabatabaei, S., Moradian, M., Radfar, A., & Alavi, A. (2011). A new prediction model for the load capacity of castellated steel beams. *Journal of Constructional Steel Research*, 67, 1096–1105. <https://doi.org/10.1016/j.jcsr.2011.01.014>
- García-Lodeiro, I., Donatello, S., Fernández-Jiménez, A., & Palomo, A. (2013). Basic principles of hybrid alkaline cements. *Chem Mater Res*, 4, 12–14.
- Garcia-Lodeiro, I., Palomo, A., & Fernández-Jiménez, A. (2015). 2 - An overview of the chemistry of alkali-activated cement-based binders. In F. Pacheco-Torgal, J. A. Labrincha, C. Leonelli, A. Palomo, & P. Chindapasirt (Eds.), *Handbook of Alkali-Activated Cements, Mortars and Concretes* (pp. 19–47). Woodhead Publishing. <https://doi.org/https://doi.org/10.1533/9781782422884.1.19>
- Gomaa, E., Han, T., Elgawady, M., Huang, J., & Kumar, A. (2020). *Prepared for Submission to Cement and Concrete Composites I Machine Learning to Predict Properties of Fresh and Hardened Alkali-activated Concrete I*.

- Gu, S., Yang, L., Du, Y., Chen, G., Walter, F., Wang, J., & Knoll, A. (2022). *A Review of Safe Reinforcement Learning: Methods, Theory and Applications*. <http://arxiv.org/abs/2205.10330>
- Gunasekara, C., Atzarakis, P., Lokuge, W., Law, D. W., & Setunge, S. (2021). Novel analytical method for mix design and performance prediction of high calcium fly ash geopolymer concrete. *Polymers*, 13(6). <https://doi.org/10.3390/polym13060900>
- Gupta, S. (2013). Using Artificial Neural Network to Predict the Compressive Strength of Concrete containing Nano-silica. *Civil Engineering and Architecture*, 1(3), 96–102. <https://doi.org/10.13189/cea.2013.010306>
- Gupta, T., & Rao, M. C. (2022). Prediction of compressive strength of geopolymer concrete using machine learning techniques. *Structural Concrete*, 23(5), 3073–3090. <https://doi.org/10.1002/suco.202100354>
- Hadi, M. N. S., Zhang, H., & Parkinson, S. (2019). Optimum mix design of geopolymer pastes and concretes cured in ambient condition based on compressive strength, setting time and workability. *Journal of Building Engineering*, 23, 301–313. <https://doi.org/10.1016/J.JOBE.2019.02.006>
- Hamilton, R. I., & Papadopoulos, P. N. (2024). Using SHAP Values and Machine Learning to Understand Trends in the Transient Stability Limit. *IEEE Transactions on Power Systems*, 39(1), 1384–1397. <https://doi.org/10.1109/TPWRS.2023.3248941>
- Hardjito, D. (2004). *ACI Materials*. <https://www.researchgate.net/publication/303836414>
- Henriques, R., Lobo, V., & Bacao, F. (2012). Spatial Clustering Using Hierarchical SOM. In *Applications of Self-Organizing Maps*. InTech. <https://doi.org/10.5772/51159>

- Holmbom, A. H., Eklund, T., & Back, B. (2011a). Customer portfolio analysis using the SOM. *International Journal of Business Information Systems*, 8(4), 396–412. <https://doi.org/10.1504/IJBIS.2011.042397>
- Holmbom, A. H., Eklund, T., & Back, B. (2011b). Customer portfolio analysis using the SOM. *International Journal of Business Information Systems*, 8(4), 396–412. <https://doi.org/10.1504/IJBIS.2011.042397>
- Huang, X., Kroening, D., Ruan, W., Sharp, J., Sun, Y., Thamo, E., Wu, M., & Yi, X. (2020). A survey of safety and trustworthiness of deep neural networks: Verification, testing, adversarial attack and defence, and interpretability. *Computer Science Review*, 37, 100270. <https://doi.org/https://doi.org/10.1016/j.cosrev.2020.100270>
- Huseien, G. F., Mirza, J., Ismail, M., & Hussin, M. W. (2016). Influence of different curing temperatures and alkali activators on properties of GBFS geopolymers mortars containing fly ash and palm-oil fuel ash. *Construction and Building Materials*, 125, 1229–1240. <https://doi.org/10.1016/j.conbuildmat.2016.08.153>
- Indra Kumar, D., & Kounte, M. R. (2016). Comparative study of self-organizing map and deep self-organizing map using MATLAB. *International Conference on Communication and Signal Processing, ICCSP 2016*, 1020–1023. <https://doi.org/10.1109/ICCSP.2016.7754303>
- Jagadesh, P., de Prado-Gil, J., Silva-Monteiro, N., & Martínez-García, R. (2023). Assessing the compressive strength of self-compacting concrete with recycled aggregates from mix ratio using machine learning approach. *Journal of Materials Research and Technology*, 24, 1483–1498. <https://doi.org/10.1016/j.jmrt.2023.03.037>

- Jamal, A. S., & Ahmed, A. N. (2025). Estimating compressive strength of high-performance concrete using different machine learning approaches. *Alexandria Engineering Journal*, 114, 256–265. <https://doi.org/10.1016/j.aej.2024.11.084>
- Ji, H., Lyu, Y., Ying, W., Liu, J.-C., & Ye, H. (2024). Machine learning guided iterative mix design of geopolymers concrete. *Journal of Building Engineering*, 91, 109710. <https://doi.org/10.1016/j.jobbe.2024.109710>
- Khale, D., & Chaudhary, R. (2007). Mechanism of Geopolymerization and Factors Influencing Its Development: A Review. *Journal of Materials Science*, 42, 729–746. <https://doi.org/10.1007/s10853-006-0401-4>
- Khalil, M. G., Elgabbas, F., El-Feky, M. S., & El-Shafie, H. (2020). Performance of geopolymers mortar cured under ambient temperature. *Construction and Building Materials*, 242, 118090. <https://doi.org/10.1016/J.CONBUILDMAT.2020.118090>
- Khan, M. A., Memon, S. A., Farooq, F., Javed, M. F., Aslam, F., & Alyousef, R. (2021). Compressive Strength of Fly-Ash-Based Geopolymer Concrete by Gene Expression Programming and Random Forest. *Advances in Civil Engineering*, 2021. <https://doi.org/10.1155/2021/6618407>
- Khattab, S. A., Elshikh, M. M. Y., Elemam, W. E., Elshami, A. A., & Youssf, O. (2023). Effect of Magnetized Water-Based Alkaline Activator on Geopolymer Concrete Mechanical Performance and Durability. *Sustainability (Switzerland)*, 15(23). <https://doi.org/10.3390/su152316315>
- Kumar, R., Althaqafi, E., Patro, S. G. K., Simic, V., Babbar, A., Pamucar, D., Singh, S. K., & Verma, A. (2024). Machine and deep learning methods for concrete strength Prediction: A bibliometric and content analysis review of research trends and future

- directions. In *Applied Soft Computing* (Vol. 164). Elsevier Ltd.
<https://doi.org/10.1016/j.asoc.2024.111956>
- Kumar, S., Kumar, R., & Mehrotra, S. P. (2010). Influence of granulated blast furnace slag on the reaction, structure and properties of fly ash based geopolymer. *Journal of Materials Science*, 45(3), 607–615. <https://doi.org/10.1007/s10853-009-3934-5>
- Kumar, V., Kalitin, D., & Tiwari, P. (2017). Unsupervised learning dimensionality reduction algorithm PCA for face recognition. *2017 International Conference on Computing, Communication and Automation (ICCCA)*, 32–37.
<https://doi.org/10.1109/CCAA.2017.8229826>
- Kupaei, R. H., Alengaram, U. J., & Jumaat, M. Z. (2014). *The Effect of Different Parameters on the Development of Compressive Strength of Oil Palm Shell Geopolymer Concrete*.
<https://doi.org/10.1155/2014/898536>
- Kuppusamy, Y., Jayaseelan, R., Pandulu, G., Kumar, V. S., Murali, G., Dixit, S., & Vatin, N. I. (2022). Artificial Neural Network with a Cross-Validation Technique to Predict the Material Design of Eco-Friendly Engineered Geopolymer Composites. *Materials*, 15(10). <https://doi.org/10.3390/ma15103443>
- Li, L., Ma, Z., & Ming, X. (2023). Multiscale ab-initio modeling and experiment of nano-CaCO₃ and fiber synergy on toughening low-carbon geopolymer composites. *Materials & Design*, 233, 112280.
<https://doi.org/https://doi.org/10.1016/j.matdes.2023.112280>
- Liew, Y.-M., Heah, C.-Y., Mohd Mustafa, A. B., & Kamarudin, H. (2016). Structure and properties of clay-based geopolymer cements: A review. *Progress in Materials Science*, 83, 595–629. <https://doi.org/10.1016/j.pmatsci.2016.08.002>

- Ling, Y., Wang, K., Wang, X., & Li, W. (2021). Prediction of engineering properties of fly ash-based geopolymer using artificial neural networks. *Neural Computing and Applications*, 33(1), 85–105. <https://doi.org/10.1007/s00521-019-04662-3>
- Lizcano, M., Kim, H. S., Basu, S., & Radovic, M. (2012). Mechanical properties of sodium and potassium activated metakaolin-based geopolymers. *Journal of Materials Science*, 47(6), 2607–2616. <https://doi.org/10.1007/s10853-011-6085-4>
- Lloyd, N. A., & Rangan, B. V. (n.d.). *Geopolymer Concrete with Fly Ash*. <http://www.claisse.info/Proceedings.htm>
- Luukkonen, T., Heponiemi, A., Runtti, H., Pesonen, J., Yliniemi, J., & Lassi, U. (2019). Application of alkali-activated materials for water and wastewater treatment: a review. In *Reviews in Environmental Science and Biotechnology* (Vol. 18, Issue 2, pp. 271–297). Springer Netherlands. <https://doi.org/10.1007/s11157-019-09494-0>
- Ma, C., Zhao, B., Guo, S., Long, G., & Xie, Y. (2019). Properties and characterization of green one-part geopolymer activated by composite activators. *Journal of Cleaner Production*, 220, 188–199. <https://doi.org/10.1016/J.JCLEPRO.2019.02.159>
- Mahameid, A., Yasin, A., & Malkawi, A. (2024). Prediction of the Compressive and Tensile Strengths of Geopolymer Concrete Using Artificial Neural Networks. *Civil Engineering and Architecture*, 12, 3193–3206. <https://doi.org/10.13189/cea.2024.120506>
- Matsimbe, J., Dinka, M., Olukanni, D., & Musonda, I. (2022). Geopolymer: A Systematic Review of Methodologies. In *Materials* (Vol. 15, Issue 19). MDPI. <https://doi.org/10.3390/ma15196852>

- Mehta, A., Siddique, R., Singh, B. P., Aggoun, S., Łagód, G., & Barnat-Hunek, D. (2017). Influence of various parameters on strength and absorption properties of fly ash based geopolymer concrete designed by Taguchi method. *Construction and Building Materials*, 150, 817–824. <https://doi.org/10.1016/J.CONBUILDMAT.2017.06.066>
- Mohammed, A. A., Ahmed, H. U., & Mosavi, A. (2021). Survey of mechanical properties of geopolymer concrete: A comprehensive review and data analysis. In *Materials* (Vol. 14, Issue 16). MDPI AG. <https://doi.org/10.3390/ma14164690>
- Nasteski, V. (2017). An overview of the supervised machine learning methods. *HORIZONS.B*, 4, 51–62. <https://doi.org/10.20544/horizons.b.04.1.17.p05>
- Nath, P. K., & Sarker, P. K. (2014). Effect of GGBFS on setting, workability and early strength properties of fly ash geopolymer concrete cured in ambient condition. *Construction and Building Materials*, 66, 163–171. <https://api.semanticscholar.org/CorpusID:137139825>
- Nazari, A., & Sanjayan, J. G. (2015). Chemically graded geopolymer under flexural loading: experimental modelling. *RSC Advances*, 5(60), 48456–48467. <https://doi.org/10.1039/C5RA07174G>
- Nematollahi, B., & Sanjayan, J. (2014). Effect of different superplasticizers and activator combinations on workability and strength of fly ash based geopolymer. *Materials & Design*, 57, 667–672. <https://doi.org/10.1016/j.matdes.2014.01.064>
- Ngo, T.-P., Vu, H.-N., & Bui, Q.-B. (2025). Application of machine learning models for the optimisation of compressive strength and water resistance of geopolymer stabilised compacted earth. *Case Studies in Construction Materials*, 22, e04203. <https://doi.org/10.1016/j.cscm.2025.e04203>

- Nguyen, H., Vu, T., Vo, T. P., & Thai, H.-T. (2021). Efficient machine learning models for prediction of concrete strengths. *Construction and Building Materials*, 266, 120950. <https://doi.org/10.1016/j.conbuildmat.2020.120950>
- Panda, B., & Tan, M. J. (2018). Experimental study on mix proportion and fresh properties of fly ash based geopolymer for 3D concrete printing. *Ceramics International*, 44(9), 10258–10265. <https://doi.org/10.1016/j.ceramint.2018.03.031>
- Parhi, S. K., & Patro, S. K. (2023). Prediction of compressive strength of geopolymer concrete using a hybrid ensemble of grey wolf optimized machine learning estimators. *Journal of Building Engineering*, 71, 106521. <https://doi.org/10.1016/j.jobe.2023.106521>
- Park, J. H., & Ihm, I. (2015). Many-to-many two-disjoint path covers in cylindrical and toroidal grids. *Discrete Applied Mathematics*, 185, 168–191. <https://doi.org/10.1016/j.dam.2014.12.008>
- Pavithra, P., Srinivasula Reddy, M., Dinakar, P., Hanumantha Rao, B., Satpathy, B. K., & Mohanty, A. N. (2016). *Effect of the Na₂SiO₃/NaOH Ratio and NaOH Molarity on the Synthesis of Fly Ash-Based Geopolymer Mortar*. 336–344. <https://doi.org/10.1061/9780784480151.034>
- Pradhan, P., Dwibedy, S., Pradhan, M., Panda, S., & Panigrahi, S. K. (2022). Durability characteristics of geopolymer concrete - Progress and perspectives. *Journal of Building Engineering*, 59, 105100. <https://doi.org/10.1016/J.JOBE.2022.105100>

- Provis, J., & Bernal, S. (2014). Geopolymers and Related Alkali-Activated Materials. *Annual Review of Materials Research*, 44. <https://doi.org/10.1146/annurev-matsci-070813-113515>
- Provis, J. L. (2018). Alkali-activated materials. *Cement and Concrete Research*, 114, 40–48. <https://doi.org/https://doi.org/10.1016/j.cemconres.2017.02.009>
- Provis, J., & Van Deventer, J. (2009a). *Geopolymers: Structures, processing, properties and industrial applications*.
- Provis, J., & Van Deventer, J. (2009b). *Geopolymers: Structures, processing, properties and industrial applications*.
- Qu, X., Yang, L., Guo, K., Ma, L., Sun, M., Ke, M., & Li, M. (2021). A Survey on the Development of Self-Organizing Maps for Unsupervised Intrusion Detection. *Mobile Networks and Applications*, 26(2), 808–829. <https://doi.org/10.1007/s11036-019-01353-0>
- Rangan, B. (2014). Geopolymer concrete for environmental protection. *Indian Concrete Journal*, 88, 41–59.
- Rangan, B., Phd, B., Cpeng, F., Fici, H., & Professor, E. (2010). *Fly Ash-Based Geopolymer Concrete*.
- Rathnayaka, M., Karunasinghe, D., Gunasekara, C., Wijesundara, K., Lokuge, W., & Law, D. W. (2024). Machine learning approaches to predict compressive strength of fly ash-based geopolymer concrete: A comprehensive review. In *Construction and Building Materials* (Vol. 419). Elsevier Ltd. <https://doi.org/10.1016/j.conbuildmat.2024.135519>

- Rattanasak, U., & Chindaprasirt, P. (2009). Influence of NaOH solution on the synthesis of fly ash geopolymer. *Minerals Engineering*, 22(12), 1073–1078. <https://doi.org/https://doi.org/10.1016/j.mineng.2009.03.022>
- Refaeilzadeh, P., Tang, L., & Liu, H. (2009). Cross-Validation. *Encyclopedia of Database Systems*, 532–538, 532–538. https://doi.org/10.1007/978-0-387-39940-9_565
- Rohit, P., Gunneswara Rao, T. D., & Chandrasekhar, M. (2024). Effect of construction demolition waste as fine aggregate and NaOH molarity on strength and fracture parameters of slag based geopolymer mortars. *Journal of Engineering and Applied Science*, 71(1). <https://doi.org/10.1186/s44147-024-00373-2>
- Rusna, K. P., & Kalpana, V. G. (2022). Using Artificial Neural Networks for the Prediction of the Compressive Strength of Geopolymer Fly Ash. In *Technology & Applied Science Research* (Vol. 12, Issue 5). www.etasr.com
- Ryu, G. S., Lee, Y. B., Koh, K. T., & Chung, Y. S. (2013). The mechanical properties of fly ash-based geopolymer concrete with alkaline activators. *Construction and Building Materials*, 47, 409–418. <https://doi.org/10.1016/j.conbuildmat.2013.05.069>
- Shahmansouri, A. A., Yazdani, M., Ghanbari, S., Akbarzadeh Bengar, H., Jafari, A., & Farrokh Ghatte, H. (2021). Artificial neural network model to predict the compressive strength of eco-friendly geopolymer concrete incorporating silica fume and natural zeolite. *Journal of Cleaner Production*, 279, 123697. <https://doi.org/10.1016/j.jclepro.2020.123697>
- Shukor Lim, N. H. A., Samadi, M., Ariffin, N. F., Hussin, M. W., Rafique Bhutta, M. A., Sarbini, N. N., Abd Khalid, N. H., & Aminuddin, E. (2018). Effect of Curing Conditions on Compressive Strength of FA-POFA-based Geopolymer Mortar. *IOP*

- Conference Series: Materials Science and Engineering*, 431(9).
<https://doi.org/10.1088/1757-899X/431/9/092007>
- Singh, B., Ishwarya, G., Gupta, M., & Bhattacharyya, S. K. (2015). Geopolymer concrete: A review of some recent developments. In *Construction and Building Materials*.
<https://doi.org/10.1016/j.conbuildmat.2015.03.036>
- Tapeh, A. T. G., & Naser, M. Z. (2023). Artificial Intelligence, Machine Learning, and Deep Learning in Structural Engineering: A Scientometrics Review of Trends and Best Practices. In *Archives of Computational Methods in Engineering* (Vol. 30, Issue 1, pp. 115–159). Springer Science and Business Media B.V.
<https://doi.org/10.1007/s11831-022-09793-w>
- Tchakouté, H. K., & Rüschler, C. H. (2017). Mechanical and microstructural properties of metakaolin-based geopolymer cements from sodium waterglass and phosphoric acid solution as hardeners: A comparative study. *Applied Clay Science*, 140, 81–87.
<https://doi.org/10.1016/J.CLAY.2017.02.002>
- Teuvo Kohonen. (1990). *The Self-organizing Map*. <https://doi.org/10.1109/5.58325>
- Tian, Q., Su, Z., Fiorentini, N., Zhou, J., Luo, H., Lu, Y., Xu, X., Chen, C., & Huang, J. (2024). Ensemble learning models to predict the compressive strength of geopolymer concrete: a comparative study for geopolymer composition design. *Multiscale and Multidisciplinary Modeling, Experiments and Design*, 7(3), 1793–1806.
<https://doi.org/10.1007/s41939-023-00303-4>
- Tian, X., Xu, W., Song, S., Rao, F., & Xia, L. (2020). Effects of curing temperature on the compressive strength and microstructure of copper tailing-based geopolymers. *Chemosphere*, 253, 126754. <https://doi.org/10.1016/J.CHEMOSPHERE.2020.126754>

- Torres-Carrasco, M., & Puertas, F. (2017). *Alkaline activation of different aluminosilicates as an alternative to Portland cement: alkali activated cements or geopolymers La activación alcalina de diferentes aluminosilicatos como una alternativa al Cemento Portland: cementos activados alcalinamente o geopolímeros*. www.ricuc.cl
- Ullah, I., & Youn, H. (2019). A novel data aggregation scheme based on self-organized map for WSN. *The Journal of Supercomputing*, 75.
<https://doi.org/10.1007/s11227-018-2642-9>
- Van Dao, D., Ly, H. B., Trinh, S. H., Le, T. T., & Pham, B. T. (2019). Artificial intelligence approaches for prediction of compressive strength of geopolymer concrete. *Materials*, 12(6). <https://doi.org/10.3390/ma12060983>
- van Jaarsveld, J. G. S., van Deventer, J. S. J., & Lukey, G. C. (2002). The effect of composition and temperature on the properties of fly ash- and kaolinite-based geopolymers. *Chemical Engineering Journal*, 89(1), 63–73.
[https://doi.org/10.1016/S1385-8947\(02\)00025-6](https://doi.org/10.1016/S1385-8947(02)00025-6)
- Winston, J. J., Turker, G. F., Kose, U., & Hemanth, D. J. (2020). Novel optimization based hybrid self-organizing map classifiers for iris image recognition. *International Journal of Computational Intelligence Systems*, 13(1), 1048–1058.
<https://doi.org/10.2991/ijcis.d.200721.001>
- Xu, H., & Van Deventer, J. S. J. (2000). The geopolymerisation of alumino-silicate minerals. *International Journal of Mineral Processing*, 59(3), 247–266.
[https://doi.org/10.1016/S0301-7516\(99\)00074-5](https://doi.org/10.1016/S0301-7516(99)00074-5)

- Yeh, I. C. (2007). Modeling slump flow of concrete using second-order regressions and artificial neural networks. *Cement and Concrete Composites*, 29(6), 474–480. <https://doi.org/10.1016/j.cemconcomp.2007.02.001>
- Yip, C. K., Lukey, G. C., & van Deventer, J. S. J. (2005). The coexistence of geopolymeric gel and calcium silicate hydrate at the early stage of alkaline activation. *Cement and Concrete Research*, 35(9), 1688–1697. <https://doi.org/https://doi.org/10.1016/j.cemconres.2004.10.042>
- Yorek, N., Ugulu, I., & Aydin, H. (2016). Using Self-organizing neural network map combined with ward's clustering algorithm for visualization of students' cognitive structural models about aliveness concept. *Computational Intelligence and Neuroscience*, 2016. <https://doi.org/10.1155/2016/2476256>
- Young, B. A., Hall, A., Pilon, L., Gupta, P., & Sant, G. (2019). Can the compressive strength of concrete be estimated from knowledge of the mixture proportions?: New insights from statistical analysis and machine learning methods. *Cement and Concrete Research*, 115(September), 379–388. <https://doi.org/10.1016/j.cemconres.2018.09.006>
- Zhang, D., Sun, F., & Liu, T. (2021). Prediction of Compressive Strength of Geopolymer Concrete Based on Support Vector Machine and Modified Cuckoo Algorithm. *Advances in Materials Science and Engineering*, 2021, 1–14. <https://doi.org/10.1155/2021/4286810>
- Zhang, Z., Provis, J. L., Reid, A., & Wang, H. (2014). Geopolymer foam concrete: An emerging material for sustainable construction. *Construction and Building Materials*, 56, 113–127. <https://doi.org/10.1016/j.conbuildmat.2014.01.081>

- Zhang, Z., Yao, X., & Zhu, H. (2010). Potential application of geopolymers as protection coatings for marine concrete: I. Basic properties. *Applied Clay Science*, 49(1), 1–6.
<https://doi.org/10.1016/j.clay.2010.01.014>
- Zhao, J., Tong, L., Li, B., Chen, T., Wang, C., Yang, G., & Zheng, Y. (2021). Eco-friendly geopolymer materials: A review of performance improvement, potential application and sustainability assessment. *Journal of Cleaner Production*, 307, 127085.
<https://doi.org/10.1016/J.JCLEPRO.2021.127085>
- Zhuang, X. Y., Chen, L., Komarneni, S., Zhou, C. H., Tong, D. S., Yang, H. M., Yu, W. H., & Wang, H. (2016). Fly ash-based geopolymer: clean production, properties and applications. *Journal of Cleaner Production*, 125, 253–267.
<https://doi.org/10.1016/j.jclepro.2016.03.019>

APPENDICES

Appendix A: Journal Publications

Teo Siaw Hui, Chang Wui Lee, Annisa Jamali, Mohamad Nazim bin Jambli, Tay Kai Meng, Nur Fadhilah Mohamad Tahan, Idawati Ismail (2025), Predicting Fly ash Geopolymer Compressive Strength from Artificial Neural Network. Semarak Ilmu Journal of Advanced Research in Applied Mechanics, publishing, ISSN: 2289-7895