الجمهورية الجزائرية الديمقراطية الشعبية

People's Democratic Republic of Algeria

وزارة التعليم العالي والبحث العلمي

Ministry of Higher Education and Scientific Research

جامعة العربي التبسي – تبسة

University of Larbi Tebessi – Tebessa –

Faculty of Science and Technology

Department of Civil Engineering

THESIS

Presented for the attainment of the **Academic Master's degree**

In: Civil Engineering

Specialty: Structures

By: Maatoub Siham

Subject:

Machine Learning for predicting Compressive and Flexural strength of Fiber-reinforced Concrete

Presented and publicly defended on the 10/06/2025, before the jury composed of:

Mr.Harkati Elhadi	professor	President
Mr. BOURSAS. Farid	MCA	Rapporteur
Mr.Abderrahmani Saif	MCA	Examiner 1

Class of: 2024/2025



قال تعالى : {وَ قُلْ رَبِّ زِدْنِي عِلْماً}

سورة [طه] :الاية 114

ACKNOWLEDGEMENTS

I praise Allah in the first place for the successful completion of this task and who gave me the health and the ability to succeed in my entire academic journey during 18 years of education.

I also appreciate my deep thanks giving and gratitude to **Dr.Boursas Farid** for his great guidance, persistent assistance and consistent assistance.

I appreciate his advice from the beginning to the end, which has allowed me to gather a full comprehension of this research. It would be much harder to complete without his advice and aid. I also warmly appreciate you for reading and fixing my errors for the time invested.

I thank you, members of the jury: **Pr.Harkati Elhadi** and **Dr. Abderrahmeni Saif** for taking part in the assessment.

I am also grateful to my family for their unwavering support, patience and motivation during the challenging times.

A big thank you to my friends and classemates for their help, motivation and all the moments we shared that made this journey more enjoyable.

Finally, I appreciate anybody who helps directly or indirectly to my final year project.

I would want to express my appreciation of the comments and recommendations that have been vital to the accomplishment of this study.

Dedication

With all feelings of respect and all the experience of gratitude. I dedicate my graduation and my joy:

To the sweat and I taught me to succeed, it does not come patience and insistence.

To the light that illuminated my path and the lamp whose light never goes out.

To the one who gave his most precious things and from whom I derived my strength and self-esteem.

My dear father 'Sebti'

To the one who made heaven under her feet and made hardships easy for me with her prayers.

To the great woman who always wished to see her eyes comforted on a day like this.

My wonderful mother 'Bachra'

To whom I strengthened my arm and were a spring from which I quench my thirst, to the comfort of my eyes.

My sweet sisters 'Soumia & Basma'

To my support and my backer who gives me love and strength.

My dear brothers 'Samir & Saif'

To the reason for my happiness and the sugar of my life, my nephew and sister's children.

My world 'Houdaifa & Sanad & Sabae'

To everyone who has been a help and support on this path.

To my loyal friends, companions of the years for those who face hardship and crises.

'Wissal -Aya-Manar-Malak-Aziza-Isra-Amira-Khaoula-Sara-Aya-Hassna-Ouissame-Haoua'.

Here is my family, young and old.

I dedicate to you this achievement and the fruit of my success that I have always wished for.

Here I am today completed and perfected the first fruits of it, thanks to God Almighty.

This is not the end of my journey; this is just the beginning of my successes.

Praise be to God for what he has given me.Thank you, love and gratitude at the beginning and the end.

And our final supplication is that all praise is due to God; Lord of the worlds.

Eng. Siham

Abstract:

This thesis explores the use of machine learning algorithms to predict the compressive and flexural strength of aluminum waste-based fiber-reinforced concrete as a green alternative for the development of mechanical properties at reduced environmental expense. Smart prediction models are to be created using three main algorithms: Artificial Neural Networks (ANN), Random Forest, and XGBoost. The procedure entailed experimental data collection, normalization as preprocessing, correlation analysis, and partitioning data into training, validation, and testing sets. Results indicated that XGBoost performed best with R^2 equal to approximately 0.80 for compressive strength and 0.97 for flexural strength. The study concludes that the use of AI in concrete mix design reduces reliance on costly experimental testing and sustainable building practices. Future studies are encouraged to expand data sources and explore hybrid model approaches for greater accuracy.

Keywords: Machine Learning, Artificial intelligence, Compressive strength, Flexural strength, Waste Aluminium fibre concrete, XGBoost, Artificial Neural Networks (ANN), Random Forest.

Résumé:

Ce mémoire étudie l'application des techniques d'apprentissage automatique pour prédire la résistance à la compression et à la flexion du béton armé de fibres issues de déchets d'aluminium, en tant qu'alternative écologique visant à améliorer les performances mécaniques du béton tout en réduisant son impact environnemental. L'objectif est de concevoir des modèles prédictifs efficaces à partir des algorithmes suivants : réseaux de neurones artificiels (ANN), forêts aléatoires (Random Forest) et XGBoost. La méthodologie adoptée comprend la collecte et le prétraitement de données expérimentales (normalisation, analyse de corrélation, division des données). Les résultats ont montré que le modèle XGBoost offrait les meilleures performances, avec un R^2 d'environ 0,80 pour la résistance à la compression et 0,97 pour la flexion. L'étude conclut que l'intelligence artificielle peut réduire considérablement la dépendance aux essais expérimentaux coûteux et favoriser une construction plus durable. Des recherches futures sont suggérées avec des jeux de données plus vastes et des approches de modélisation plus avancées.

Mots clés : Apprentissage automatique, Intelligence artificielle, Résistance à la compression, Résistance à la flexion, béton de fibres d'aluminium recyclé, XGBoost, Réseaux de Neurones Artificiels(RNA), Forêts aléatoires.

ملخص

تعرض هذه الأطروحة دراسة استخدام تقنيات التعلم الآلي لتوقع مقاومة الضغط والانحناء في الخرسانة المسلحة بألياف نفايات الألمنيوم، كحل بيئي مستدام لتحسين الأداء الميكانيكي للخرسانة وتقليل الأثر البيئي الناتج عن صناعة الأسمنت. تهدف الدراسة إلى تطوير نماذج ذكية قادرة على توقع الخصائص الميكانيكية للخرسانة بناءً على مكوناتها، باستخدام ثلاث خوارزميات رئيسية: الشبكات العصبية الاصطناعية(ANN) ، الغابات العشوائية(Random Forest) ، و XGBoost.اعتمدت المنهجية على جمع بيانات تجريبية، مع معالجتها وتحليلها إحصائيًا من خلال تطبيع القيم، تحليل الترابط، وتقسيم البيانات إلى مجموعات تدريب واختبار وتحقيق. أظهرت النتائج أن نموذج XGBoost حقق أعلى دقة، حيث بلغ معامل التحديد (R^2) حوالي 0.80 لمقاومة الضغط و0.97 لمقاومة الانحناء. توصلت الدراسة إلى أن دمج الذكاء الاصطناعي في تصميم الخلطات الخرسانية يمثل وسيلة فعالة لتقليل الحاجة إلى التجارب المخبرية المكلفة، ويوجه نحو تقنيات بناء أكثر استدامة. وأوصت بإجراء أبحاث مستقبلية تعتمد على قواعد بيانات أوسع ونماذج هجينة أكثر تقدمًا.

الكلمات المفتاحية : التعلم الالي , الذكاء الاصطناعي, قوة الانضغاط,قوة الانحناء , الخرسانة المسلحة بالألياف , XGBoost, الشبكات العصبية الاصطناعية,الغابات العشوائية .

Table of contents

Abstract..... I

Résumé..... II

ملخص.....III

List of tables..... VII

List of figures.....IX

List of Abbreviations XI

CHAPTER01: General Introduction I

1.1 Introduction:.....2

1.2 Research objectives:.....3

1.3 Original Contributions: 3

1.4 Thesis Structure: 3

CHAPTER02: General Overview of Machine Learning……..……...5

2.1 Introduction: 6

2.2 The revival of machine learning:.....6

2.3 Types of Machine Learning:.....7

2.3.1 Supervised Learning: 7

2.3.2 Unsupervised learning:.....8

2.3.3 Clustering: 9

2.3.4 Dimensionality reduction:.....9

2.3.5 Anomaly detection:.....9

2.3.6 Association Rule learning:.....10

2.4 Semi-supervised learning: 10

2.5 Reinforcement learning: 11

2.6 Supervised learning algorithms: 11

2.6.1 Linear regression:.....11

2.6.1.1 Simple linear regression…… 11

2.6.2 Multiple linear regression:.....14

2.6.3 Support Vector Machines (SVM):.....16

2.6.4 Decision Trees: 17

2.6.5 Random forests: 17

2.6.6 Gradient Boosting:.....19

2.7 Testing and validating a Machine Learning Model:.....19

2.8 Conclusion:.....20

CHAPTER03: Experimental Overview of Aluminium Waste Utilisation in Concrete …… 22

3.1 Introduction:.....23

3.2 Experimental Overview of Aluminium Waste Utilisation in Concrete:.....23

3.2.1 The Use of Aluminum Waste for Concrete Production (2011)……23

3.2.2 Prediction of the Compressive Strength of Aluminum Waste–Cement Concrete Using Scheffe's Theory (2012)…… 24

3.2.3 Comparison of Data Mining Techniques for Predicting Compressive Strength of Environmentally Friendly Concrete (2016)…… .25

3.2.4 Mechanical behaviour of fibre reinforced concrete using soft – drink can (2017)…… 25

3.2.5 Investigating the Behavior of Concrete and Mortar Reinforced with Aluminum Waste Strips (2018)…….....28

3.2.6 Mechanical properties of pavement quality concrete with secondary aluminium dross as partial replacement for ordinary Portland cement (2018)……..29

3.2.7 Effect of waste aluminium shavings on the bond characteristics of laterized concrete (2019)……...30

3.2.8 Experimental Study on Strength Properties of Aluminium Fibre Reinforced Concrete (2019)……31

3.2.9 Influence of Aluminum Waste on the performance of fluid sand Concrete (2020)…….....31

3.2.10 Mechanical Behavior of Concrete Reinforced with Waste Aluminium Strips (2021)……34

3.2.11 Effect of industrial waste metal chips on flexural behavior of reinforced concrete beams
(2023)……35

3.2.12 Evaluation of sisal fiber and aluminum waste concrete blend for sustainable construction using adaptive Neuro Fuzzy Inference System (2023)……35

3.2.13 The effect of Aluminum Waste Ash on the concrete compressive strength (2023……36

3.2.14 Use of experimental design methodology for the formulation of waste-based concrete (Utilisation de la méthodologie des plans d'expérience pour la formulation de béton à base de déchets) (2023)……38

3.2.15 Study on Properties of Concrete Using E-Waste Copper, Aluminium and Steel Fibers (2024)…… 41

3.2.16 Compression-Flexure Relationship in Fibre Concrete Based on Wasted Aluminium (2024)…… 42

3.3 Conclusion……42

CHAPTER04: Application of Machine Learning for Predicting Compressive and Flexural strength of Waste Aluminum Fiber concrete…… 44

4.1 Introduction:.....45

4.2 Libraries Used…… 46

4.2.1 NumPy:.....46

4.2.2 Pandas:.....46

4.2.3 Matplotlib: 47

4.2.4 Seaborn: 48

4.2.5 Scikit-Learn:.....48

4.2.6 TensorFlow: 48

4.2.7 Keras: 49

4.3 Data collecting and pre-processing.....49

4.3.1 Data Distribution Analysis:.....51

4.3.2 Correlation Between Variables: 51

4.3.3 Variable Normalization: 52
4.3.4 Hyperparameter Tuning: 53
4.4 Data scaling and splitting 53
4.4.1 Scaling techniques: 53
4.4.2 Training-validation-testing division: 53
4.4.3 Cross-validation methodology: 53
4.5 Model Performance Metrics 54
4.5.1 MAE (Mean Absolute Error): 54
4.5.2 MSE (Mean Squared Error): 54
4.5.3 R^2 (R-squared): 54
4.6 Model Creation and Results 55
4.6.1 Compressive strength: 57
4.6.2 Flexural strength: 60
4.7 Model comparison 63
4.7.1 Compressive Strength Prediction 63
4.7.2 Flexural Strength Prediction 63
4.7.3 Summary of Model Selection 64
4.8 Conclusion 64
CHAPTER 05: General conclusion and recommendations for future study 66
5.1 Conclusion 67
5.2 Recommendations for future study 68

REFERENCES 69

List of tables

Table3. 1: Compressive strength of AW-Concrete……..39

Table4. 1: Compressive Strength Prediction Results.....63

Table4. 2: Flexural Strength Prediction Results. 64

Table4. 3: Summary of model selection. 64

List of figures

Figure2. 1: A training package labelled for supervised learning (e.g., classification (spam))...7

Figure2. 2: Regression……8

Figure2. 3: Anomaly Detection……10

Figure2. 4: Reinforcement Learning……11

Figure2. 5: Linear regression……13

Figure2. 6: Multiple linear regression……..15

Figure2. 7 : SVM……..16

Figure2.8: Bagging **……..**18

Figure2. 9: GBRT with Early Stopping……... 19

Figure3. 1: Compressive strength of AW-Concret……. 23

Figure3. 2: Flexural strength of AW-Concrete…… 24

Figure3. 3: Compressive strength of AW-Concrete…….24

Figure3. 4: Experimental result for compressive strength of PC and PLC concrete……25

Figure3. 5: Compressive strength of AW Fiber in 7 days……26

Figure3. 6: Compressive strength of AW Fiber in 28 days……..26

Figure3. 7: Flexural strength of AW Fiber in 7 days……... 27

Figure3. 8: Flexural strength of AW Fiber in 28 days……27

Figure3. 9: Effect of Aluminum strip % on the compressive strength 28

Figure3. 10: Compressive strength of AW Concrete in 28 days……..28

Figure3. 11: Effect of Aluminum strip % on the flexural strength……... 29

Figure3. 12: Flexural strength of AW Concrete in 28 days……29

Figure3. 13: Variation of compressive strength for different replacements of aluminum dross…… .30

Figure3. 14: Bond Resistance of Laterized Concrete with Varying Percentages of Waste Aluminum Shaving (16 mm and 20 mm Diameter bars)……30

Figure3. 15: Effect of fiber volume on compressive strength……..31

Figure3. 16: The effect of Aluminum Waste on the compressive strength of Concrete …….....32

Figure3. 17:The effect of Aluminum Waste on the flexural strength of Concrete …….....32

Figure3. 18: The effect of Aluminum Waste on the compressive strength of Mortar …….....33

Figure3. 19: The effect of Aluminum Waste on the tensile strength of concrete …….. 33

Figure3. 20: Comparison of percentage increase of flexural strength…… 34

Figure3. 21: Compressive strength of standard cubic concrete specimens……..35

Figure3. 22: Experimental and ANFIS results……36

Figure3. 23: Concrete compressive strength to age at w/c=0.474……...36

Figure3. 24: Compressive strength at 10 % AWA concrete (w/c=0.525)……37

Figure3. 25: Compressive strength at 20% AWA concrete (w/c=0.575)…… 37

Figure3. 26: Compressive strength at 30% AWA concrete (w/c=0.611)…… 37

Figure3. 27: Compressive strength and age of concrete with varied w/c…….38

Figure3. 28: Compressive strength and age of concrete specimen with constant w/c…….38

Figure3. 29: Compressive strength of E-Waste Fibers……41

Figure3. 30: Flexural strength of E-Waste Fibers……41

Figure3. 31: Comparison of average compressive strength of test specimen……..42

Figure3. 32: Comparison of average flexural strength of test specimen……..42

Figure4. 1: Box Plot of Normalized Features…….. 50

Figure4. 2: Pearson Correlation Matrix for Compressive Strength. 51

Figure4. 3: Pearson Correlation Matrix for Flexural Strength.....52

Figure4.4: Modeling Process and Performance Evaluation for Predicting Mechanical Strengths of Materials. 56

Figure4. 5: The performance of XGBoost in predicting compressive strength. 57

Figure 4. 6: The performance of Random Forest in predicting compressive strength. 58

Figure4. 7: The Neural Network's performance in predicting compressive strength. 59

Figure4. 8:XGBoost's performance in predicting flexural strength. 60

Figure4. 9: Random Forest's performance for flexural strength. 61

Figure4. 10: The Neural Network's performance for flexural strength. 62

List of Abbreviations

Al: Aluminum.

WAF: Waste Aluminium Fibres.

ML: Machine Learning.

ANN: Artificial Neural Networks.

RF: Random Forest.

SVM: Support Vectors Machines.

HCA: Hierarchical Cluster Analysis.

PCA: Principal Component Analysis.

LLE: Local Linear Embedding.

T-SNE: T- Stochastic Neighbor Embedding.

RSS: Residual Sum Squares.

GBRT: Gradient Boosted Regression Trees.

AW: Aluminum Waste.

PC: Portland Cement.

PLC: Portland Limestone Cement.

ANFIS: Adoptive Neuro Fuzzy Inference System.

WAFC: Waste Aluminum Fibres Concrete.

MAE: Mean Absolute Error.

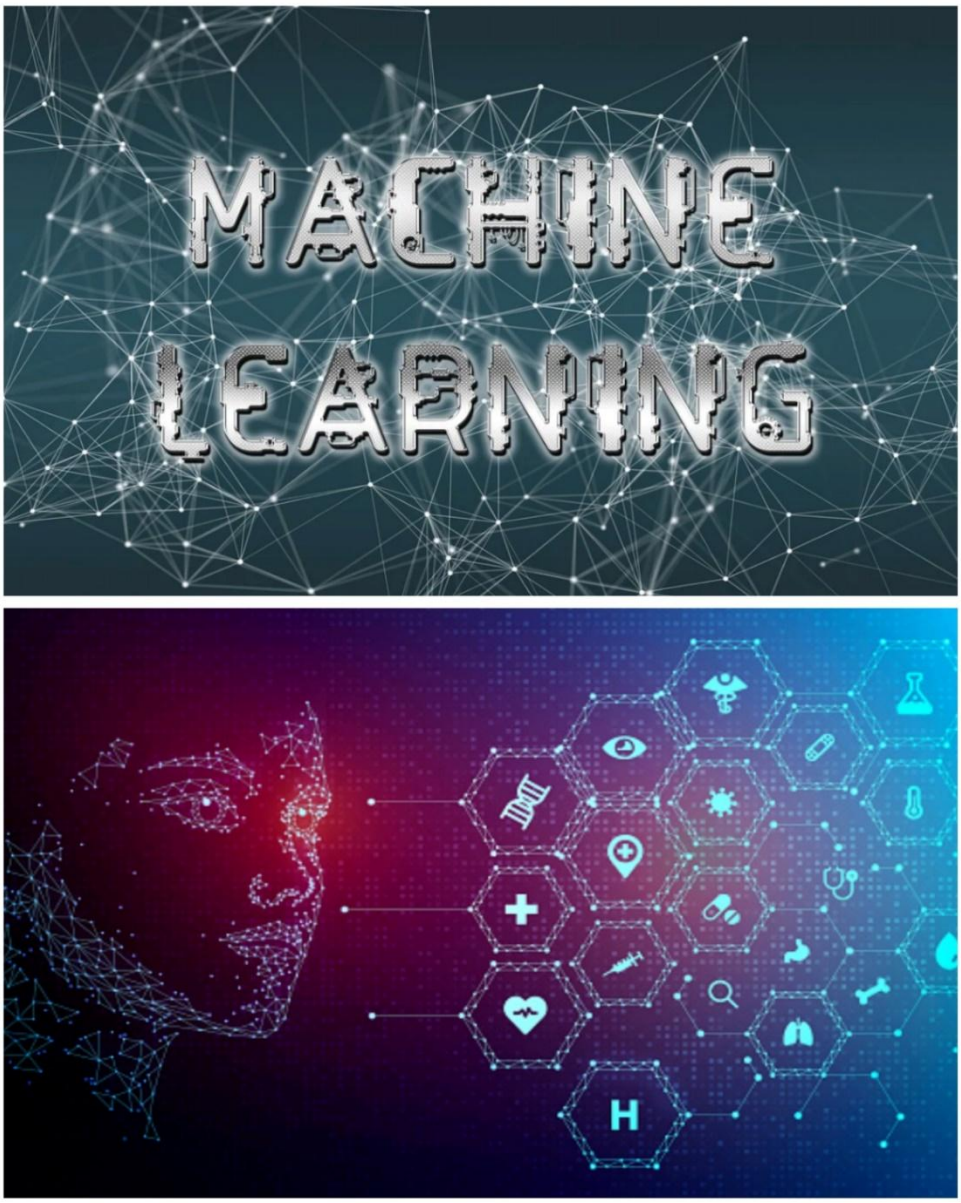
MSE: Mean Squared Error.

R^2: R-Squared.

NumPy: Numerical Python.

SciPY: Scikit learn Python.

XGB: XGBoost.


MACHINE
LEARNING

CHAPTER01: General Introduction

1.1Introduction:

Concrete remains the most widely used construction material worldwide, second only to water in total volume consumed. Its broad utility stems from its strength, durability, and adaptability to various structural demands. However, the rapid pace of urban development and industrialization has placed a significant strain on the availability of raw materials and has increased environmental concerns related to concrete production. The extraction of natural aggregates and the energy-intensive process of cement manufacturing contribute substantially to the depletion of natural resources and to greenhouse gas emissions.

In response to these challenges, the incorporation of industrial waste materials into concrete has gained increasing attention. Among such materials, waste aluminium fibres (WAF), a by-product of machining and recycling processes, offer potential for reinforcement while contributing to waste minimization. When incorporated into concrete, these fibres can improve mechanical properties, particularly in terms of crack resistance and energy absorption. The use of waste materials in concrete not only helps reduce environmental burdens but also enhances sustainability within the construction sector by promoting circular economy practices.

Despite the promising potential of WAF concrete, predicting its mechanical behaviour specifically its compressive and flexural strength—remains complex due to the highly non-linear interactions between various mix constituents. Traditional empirical or statistical methods often fall short in capturing the nuanced relationships inherent in such composite materials. Therefore, Machine Learning (ML) techniques, which have demonstrated significant success in data-driven modelling across many engineering disciplines, present a powerful alternative.

ML models can effectively learn from historical datasets and identify intricate patterns without requiring explicit physical models. These capabilities make ML highly suited for predicting the performance of new or modified concrete mixtures. In particular, ML methods can handle multivariate and nonlinear data, enabling accurate predictions of concrete properties based on constituent proportions, curing conditions, and fibre characteristics.

This thesis investigates the application of ML algorithms—including Artificial Neural Networks (ANN), Random Forest (RF), and XGBoost—to predict both the compressive and flexural strength of WAF concrete. Through the development of robust predictive models, this research aims to support the design and optimization of sustainable concrete mixtures utilizing industrial waste.

1.2 Research objectives:

The primary goal of this thesis is to explore the effectiveness of ML techniques in modelling the mechanical behaviour of waste aluminium fibre concrete. Specific objectives include:

- To develop a comprehensive experimental dataset for WAF concrete mixtures with varying proportions of aluminium fibres and cementitious components.
- To investigate the influence of input variables on the compressive and flexural strength using advanced data analysis and visualization techniques.
- To apply and compare the performance of multiple ML models in predicting compressive and flexural strength.
- To evaluate the reliability and generalization ability of each ML model through rigorous validation techniques.

1.3 Original Contributions:

This research makes several original contributions to the field of sustainable concrete technology and data-driven modelling:

- Creation of a novel and extensive experimental dataset on WAF concrete, incorporating compressive and flexural strength results from a wide range of mixture compositions.
- Comparative application of three different ML models for simultaneous prediction of two mechanical properties, enhancing understanding of multi-target learning in concrete science.
- Analysis of variable importance and model interpretability, providing insights into the influence of aluminium fibres on concrete performance.
- Contribution to the sustainable construction domain by demonstrating a scalable approach to reducing construction waste through WAF integration.

1.4 Thesis Structure:

The thesis is structured into five chapters as follows:

- **Chapter 1** presents the background, motivation, objectives, and contributions of the study.

- **Chapter 2** provides a comprehensive review of the existing literature on the application of ML in predicting concrete properties.
- **Chapter 3** details the experimental program, including material selection, mix design, and dataset compilation.
- **Chapter 4** outlines the data pre-processing steps, ML model development, training procedures, and performance evaluation, including comparative analysis between models.
- **Chapter 5** Conclusion and recommendations for future study.

In summary, this research paves the way for more sustainable material use in construction, leveraging modern technologies like machine learning to support green development goals.

CHAPTER 02: General Over view of Machine Learning

2.1 Introduction:

Machine learning (ML) is one of the subdivisions of computer science that deals with the creation of algorithms that, in order to be useful, relies on a set of some phenomenon's examples. Examples could be taken from nature, crafted by humans themselves, or made by some other algorithm.

Machine learning (ML) is a dynamic and rapidly expanding field of computer science whose purpose is to design algorithms capable of learning from data without programming. By analysing data patterns, ML systems are designed to make decisions, predictions, or classifications. The usefulness of machine learning lies in the ability to address difficult real-world problems—like spam filtering to predictive analytics—using data to build statistical models.

In this chapter, we present a state-of-the-art of machine learning. Firstly, we give a brief historical overview from the beginnings to the development of this field. Then we let's give its different types and some algorithms used. Before concluding this chapter, we will let’s describe how to test and validate a machine learning model.

2.2The revival of machine learning:

In 2006, Geoffrey Hinton et al. published a paper [1] showing how to train a deep neural network to recognize handwritten digits with industry-leading accuracy (>98%). They called this approach "Deep Learning". It was widely believed that it was impossible to train a deep neural network and had been given up on by most researchers by the 1990s. This research renewed interest in the scientific community, and soon enough many new publications demonstrated that deep learning was not just possible, but that it was capable of yielding stunning results no other method of machine learning could even hope to come close to matching (granted an enormous amount of computational power and enormous datasets). This enthusiasm quickly filtered into many other areas of machine learning.

In 10 years, machine learning has conquered the industry: it is now at the heart of the high-tech product magic, web search result ranking, powering voice recognition on smartphones and video recommendations, beating the world champion in the game of Go etc.

2.3Types of Machine Learning:

There are many types of machine learning systems. Here, we classify depending on whether or not they are human-directed (supervised, unsupervised, semi-supervised, and reinforcement learning).

2.3.1 Supervised Learning*:*

When training data presented to the algorithm is made up of solutions, which are known as labels.

Classification is a common supervised learning problem. The spam filter is a good example. For instance: It is trained on a lot of examples of emails with their class (ham or spam), and it must learn to classify new emails, as shown in Figure 2.1:



Figure2. 1: A training package labelled for supervised learning (e.g., classification (spam)). [2]

Another typical task is to predict a target numerical value, for example, the price of a car, based on a set of qualities (age, mileage, make, etc.) known as pre- dictators. Regression is the name of this kind of task. To train the system, we have to supply it numerous examples of automobiles, their predictors, and labels (i.e., their price). Figure 2.2 shows a regression spot .[2]

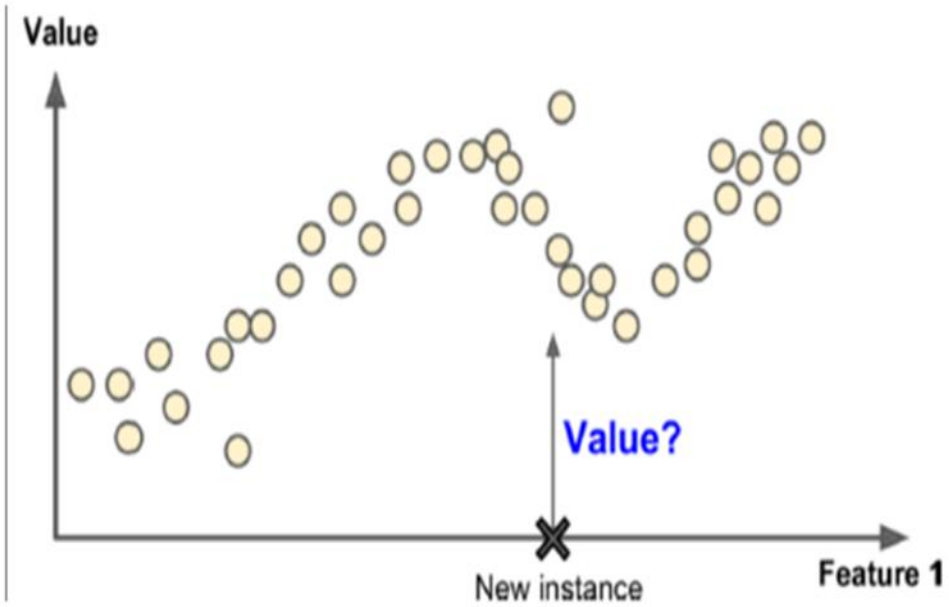


Figure2. 2: Regression. [2]

Also take into consideration that some of the regression algorithms are applicable to classification as well, and vice versa. For example, logistic regression can be widely applied for classification as well, because it can output a value that has something to do with the chance of belonging to a particular class (20% chance of spamming).
Some of the most important supervised learning algorithms are:

- K nearest neighbors.
- Linear regression.
- Logistic regression.
- Support Vector Machines (SVM).
- Decision trees and random forests.

2.3.2 Unsupervised learning:

In unsupervised learning, the training data is not tagged. The system tries to learn on its own. The model has no "answers" to learn lessons from; it must impose meaning on the data from the observations themselves.

Unsupervised learning allows us to address problems with minimal or no knowledge of what our results will be like. We can obtain a form from data that we do not necessarily know the impact of variables.
The following are some of the most important unsupervised learning algorithms.

2.3.3 Clustering:

K-Means; Hierarchical Cluster Analysis (HCA); Maximizing expectations.

- ✓ Visualization and dimensionality reduction:
- Principal Component Analysis (PCA).
- PCA kernel.
- Local linear embedding (LLE).
- T-distributed Stochastic Neighbor Embedding (t-SNE).
- ✓ Learning the association rules:
- A priori.
- Eclat.

2.3.4 Dimensionality reduction:

The aim is to reduce the data without losing too much information. One way to do it is to put a number of related characteristics into one. For example, the mileage of a car is highly correlated with age, thus the dimensionality reduction algorithm will unite them into one feature that defines the wear and tear of the vehicle. Feature extraction is known as this.[7]

2.3.5 Anomaly detection:

Anomaly detection is an interesting field, which aims to identify distant objects that are deviant in relation to the general data distribution. Outlier detection was found to be essential in many applications, including abnormal transaction detection by credit card companies to prevent fraud, defect detection in manufacturing, or automatic removal of outliers from a dataset before sending them to another algorithm learning process. The system learns from normal instances, and when a new instance is encountered, it can classify whether it seems to be a normal instance or whether it is likely an anomaly, as shown in the graph "Figure 2.3":



Figure2. 3: Anomaly Detection.[7]

2.3.6 Association Rule learning:

The goal is to plunge into enormous amounts of information and discover interesting correlations between the two between the attributes. For example, suppose that we have a grocery store.

Application of an association rule to sales data may reveal that people who purchase barbecue sauce and potatoes also tend to buy steak. By thus, possibly these objects may be positioned in very near distance from one another.

2.4Semi-supervised learning:

Semi-supervised learning approaches can accommodate partially labelled training data usually much more unlabelled data and limited number of labelled data and can provide predictions for any unseen point. Semi-supervised learning is observed in those circumstances where labels are expensive to get but the information is readily attainable in the form of being unlabelled. Different types of problems arising within applications such as classification, regression or ranking operations can be considered as instances semi-supervised learning. The hope is that the availability distribution of unlabelled data can allow it to perform better than in the supervised learning scenario. An investigation of the conditions under which this is so is the focus of much of modern machine learning theoretical and applied work. [3]

Most popular semi-supervised learning algorithms are hybrids of unsupervised and supervised algorithms. For example, Deep Belief Networks (DBNs) use unsupervised

building blocks named restricted Boltzmann machines (RBMs) stacked upon each other. RBMs are learned sequentially and in an unsupervised manner, and then the entire system is fine-tuned using supervised learning techniques.

2.5Reinforcement learning:

The entire reinforcement learning system is shown in Figure2.4. Unlike supervised learning setup, the learner in this case is not given a batch of data passively labelled. Instead, it learns through a series of steps by discovering the environment. After an action, the agent or student receives two types of information: their state in the environment at the moment and a real-value reward depending on the task and its respective intent. The goal of the agent is to maximize its reward and there by determine the best sequence of actions, or policy, to accomplish the goal. But the reward that he receives from the environment is just the near reward of the action just taken. Another important consideration of reinforcement learning is to account for deferred rewards or penalties. The agent is conflicted between exploring new states and actions to gain more information about the world and the rewards, and to act on the information already acquired in order to maximize its reward. It's the so-called trade-off between exploration and exploitation, that is, pertaining to learning by reinforcement. [3]

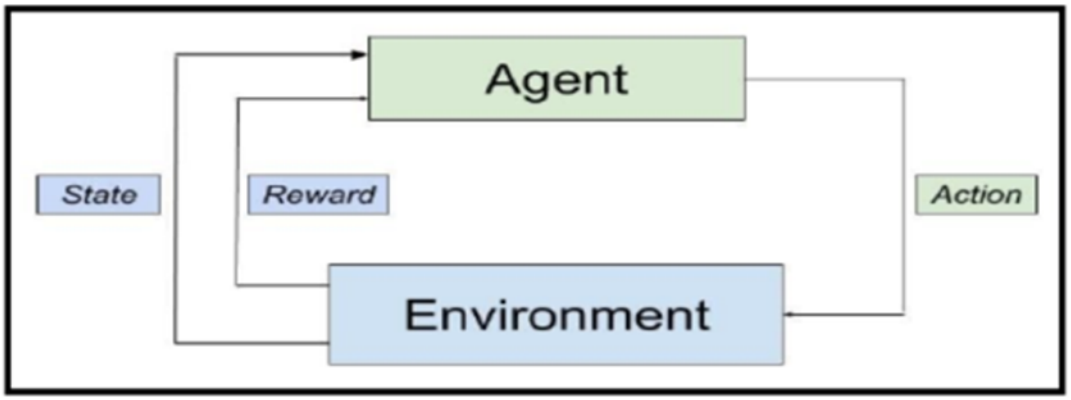


Figure2. 4: Reinforcement Learning. [3]

Deep Mind's Alpha Go program is an excellent illustration of reinforcement learning: it possesses made headlines in March 2016 by defeating world champion Lee Sedol at the game of Go. He learned the politics of victory by playing several parties against him after studying millions of games.

2.6 Supervised learning algorithms:

2.6.1 Linear regression:

Linear regression is a very simple supervised learning method. Linear regression is a good

quantitative response prediction tool, in any case. Although it might seem rather bland compared to some more statistical learning approaches linear regression remains a worthwhile and well-used statistical learning method. In addition, it is a good starting point for new techniques that are considered to be as generalizations or extensions of linear regression. [4]

2.6.1.1 Simple linear regression:

Simple linear regression is as easy as the name: a very basic linear approach to predict a quantitative response Y with a single predictor variable X. She assumes there is approximately a linear relationship between X and Y. We can write this linear relationship mathematically as:

$$\gamma \approx \beta_0 + \beta_1 \chi \quad (2.6.1.1)$$

« ≈ » can be read as "is roughly modeled as". We will sometimes explain in asserting that we are $\boldsymbol{\gamma}$ over $\boldsymbol{\chi}$ (or $\boldsymbol{\gamma}$ over$\boldsymbol{\chi}$).For example, $\boldsymbol{\chi}$can be employed to represent expenditure on TV ads and $\boldsymbol{\gamma}$ can be employed to represent sales. We can then reduce sales on television by adjusting the model.

$$Vents \approx \beta_0 + \beta_1 \times TV$$

In the final equation, β_0 and β_1 are two unknown constants that represent the terms and the slope in the linear model. Together as a pair, β_0 and β_1 are also known as the coefficients or parameters of the model.

Once we have used our training data to produce estimates β_0' and β_1' for the model coefficients, we can use them to predict sales for a given value of the amount spent on TV advertising by calculation:

$$\gamma' = \beta_0' + \beta'\chi$$

Where γ' is used for a forecast of γ for based $\chi = x$. Here we use a symbol, ', for indicate the estimated value of a coefficient of an unknown parameter, or to indicate the predicted value of the response. In practice, β_0 and β_1 are unknown. Thus, before we use (2.6.1.1) in making predictions, we need to use data to estimate the coefficients.

Either: (x1, y1), (x2, y2), . . ., (xn, yn)

Represent n pairs of observations, where each pair includes a measurement of χ and a measure of γ. In the case of advertising, this collection of data includes the expenditure on advertising and sales of the product. We wish to obtain estimates of coefficients β_0' and β_1' in

such a way that the linear model (2.6.1.1) will be closely fitted to the available data, i.e. that $\gamma' \approx \beta_0' + \beta_1'xi$ for i = 1,. n.That is, we want to find an intercept β_0' and a slope β_1' so that the resulting line is as close as possible to the data points. There are many ways to measure closeness. But the most common way is to minimize the least squares criterion. As in the example in Figure 2.5:

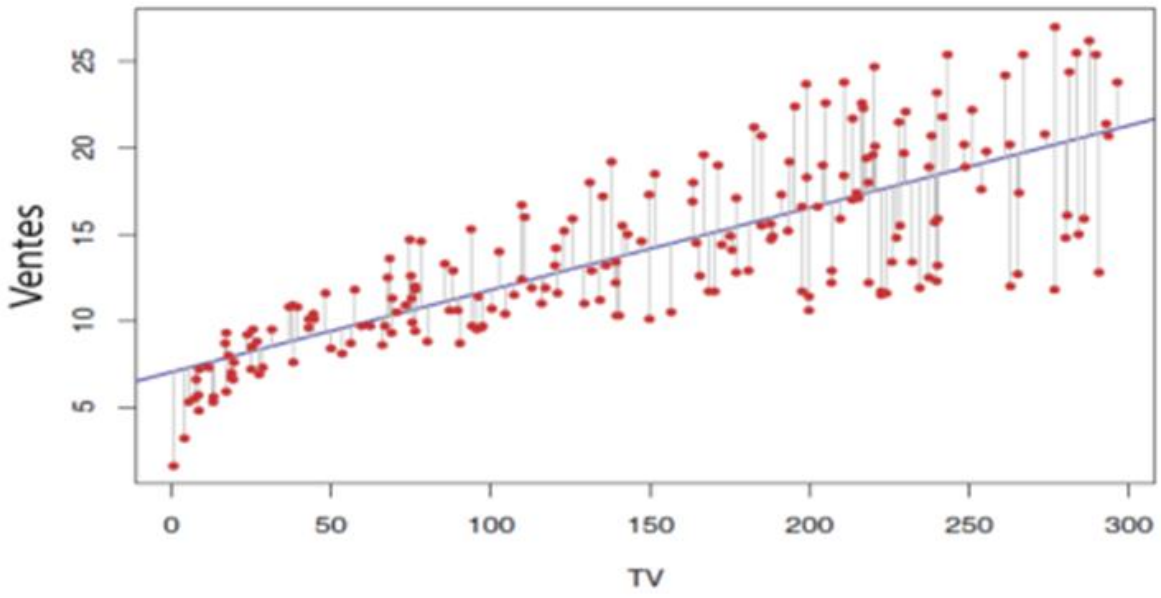


Figure2. 5: Linear regression. [5]

For advertising data, least squares corresponding to the regression of sales on the television advertising budget (TV). The adjustment is achieved by minimizing the sum squares. Each grey line segment is an error, and the adjustment compromises by averaging their squares. A linear fit captures the nature of the relationship here, although it is a bit deficient in the left part of the graph.

Let $\gamma_i' = \beta_0' + \beta_1'xi$ is the estimation of γ in terms of the ith value of χ. Then ei = γi-γi′ is the ith residual - that is, it is the deviation between the actual ith response and the ith fitted value of answer from our linear model. We define the residual sum of squares (RSS) like:

$RSS = e^2_1+e^2_2+\ldots+e^2_n$,

or equivalent to:

$RSS= (\gamma_{1-} \beta_0' -\beta'_1x_1)^2 +(\gamma_{2-}\beta_0'-\beta_1'x_2)^2 +\ldots+(\gamma_{n-}\beta_0'-\beta_1'x_n)^2$

The least-squares technique chooses β_0' and β_1' to minimize the RSS. It employs a precise computation, we can show that the minimizers are:

$$\beta_1' = \frac{\sum_{i=1}^{n}(x_i-\overline{x})(y-\overline{y})}{\sum_{i=1}^{n}(x_i-\overline{x})^2}$$

$$\beta_0' = \overline{y}-\beta_1'x$$

where $\overline{y} \equiv \frac{1}{n}\sum_{i=1}^{n} y_i$ and $\overline{x} \equiv \frac{1}{n}\sum_{i=1}^{n} x_i$ are the sample means. In other words, the above formulas are the least-squares estimates for simple linear regression. [4]

2.6.2 Multiple linear regression:

Simple linear regression is a convenient technique to predict a response based on a single predictive variable. In the real world, however, we typically have multiple predictors. For example, in advertising data, we considered the relationship between sales and television commercial. We also know the dollars paid for advertising in the and newspapers, and perhaps we would like to know whether one of those two media is associated with the to sales. How can we extend our analysis of advertising data to account of those two other predictors? One method is to estimate three regressions single systems of lines, each having a different advertising medium as a predictor.

However, the approach of estimating a separate simple linear regression model for not every indicator is not entirely satisfactory.

First, a forecast based on the levels of the three ad budgets is indeterminate, since the three budgets are specified by three different regression equations. Second, each of the three regression equations is derived solely from two of the three media in order to get the regression coefficient estimates. We will soon realize in an instant that if we have budgeted media that are correlated in our data set, this leads to some very misleading estimates.

Instead of using a simple, unique linear regression model for each predictor, a better strategy is to modify the simple linear regression model so that it will fit on multiple predictors. To achieve this, we can assign a unique slope coefficient to each predictor in a single model. In general, assume that we have unique predictors. Then, the multiple linear regression model is given by:

$\gamma = \beta_0 + \beta_1 X1 + \beta_2 X2 + \ldots + \beta p Xp + e;$

Just like in the case of simple linear regression, the regression coefficients β_0, β_1,. $\beta p'$

are unknown and need to be estimated. With estimates β_0', β_1',. .. $\beta p'$, we may do

Forecasts with the formula:

$\gamma' = \beta_0' + \beta'_1 X_1 + \beta_2' X_2 + \ldots + \beta' pXp,$

The parameters are being estimated using the identical least squares approach that

We have experienced in simple linear regression. We have selected β_0, β_1,. βp for

Minimize the sum of the squares of the residuals:

$RSS = \sum_{i=1}^{n} (\gamma_i - \gamma_i')^2$

$RSS = \sum_{i=1}^{n} (\gamma_i - \beta_0' - \beta_1' x_i 1 - \beta_2' x_i 2 - \cdots - \beta' p x_i p$

Figure 2.6 shows a three-dimensional structure, one response and two predictors, the least squares regression is a plane. The design is chosen to minimize the sum of squares of the vertical distances between each observation (indicate in red) and the plane**:**

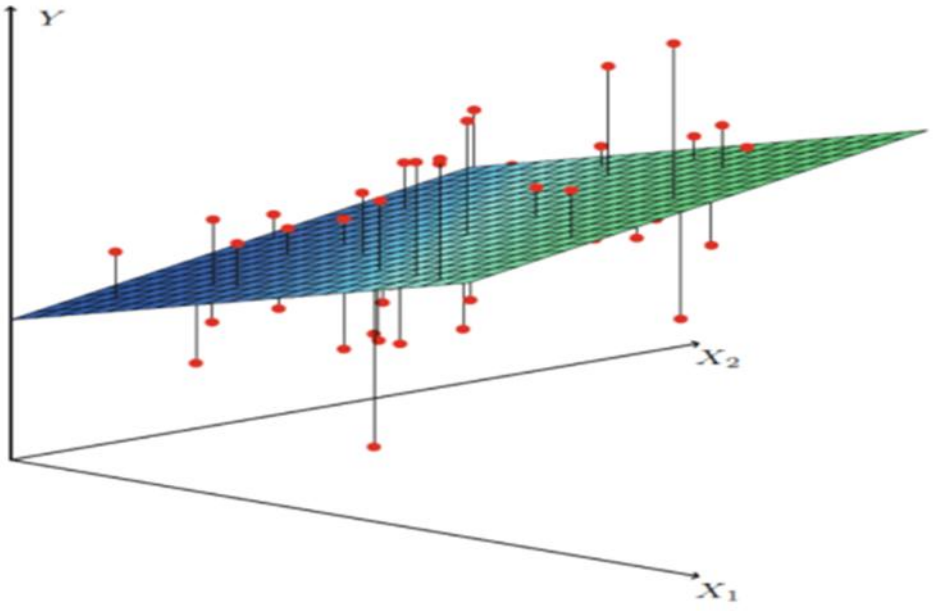


Figure2. 6: Multiple linear regression. [5]

In a three-dimensional framework, with two predictors and one response, the regression line of least squares becomes a plane. The design is chosen to minimize the sum of the squares of the distances between each observation (indicated in red) and the plan.

The values β_0'; β_1';…$\beta p'$ are the estimates of the multiple least squares regression coefficient. Unlike simple linear regression estimates, multiple regression coefficient estimates have somewhat complicated forms that are more easily represented using matrix algebra. [4]

2.6.3 Support Vector Machines (SVM):

A support vector machine (SVM) is a machine learning algorithm for supervised learning that is primarily used in classification. An SVM is trying to find a hyper plane, which divides the samples in the dataset. In Figure 2.7, we have two classes points (red and blue) which lie in a two-dimensional feature space (the x and y). If the x and y values of a point are both below five, the dot is blue. In all other cases, the dot is red. In this case, the classes are linearly separable, i.e., we can separate them by a hyper plane. [6]

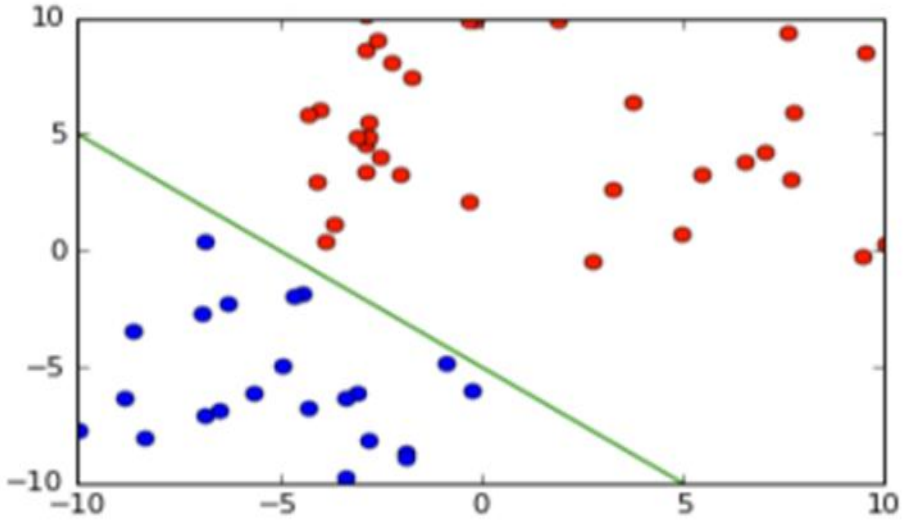


Figure2. 7 : SVM. [6]

An SVM tries to find a hyper plane with maximum distance between the hyperplane and the points. That is, out of all the hyper planes that can possibly distinguish samples, the SVM find the one that is farthest from all the points. In addition, SVMs also deal with data which are linearly inseparable. There are two ways this is achieved: addition of margins or use of the kernel trick. [7]

• Soft margins accomplish this by allowing a few instances to be misclassified while maintaining the maximum predictive power of the algorithm. In practice, it is preferable not to overtrain the machine learning model, and we can do this by relaxing some assumptions of the support vector machine.
• Kernel trick solves the same problem but differently. Let us assume that we map a two-dimensional space to let us but keep the classes linearly inseparable. Kernel trick uses a kernel function that works by mapping the data by having more size. In our scenario, the data, once mapped, is three-dimensional. The linearly inseparable classes in two dimensions will become linearly separable in three dimensions and our problem is solved.

2.6.4 Decision Trees:

Like SVMs, decision trees are universal machine learning algorithms that can be applied to classification and regression tasks, and even to tasks to multiple outputs. They are very powerful algorithms that can fit sets of complex data. Decision trees are also the building blocks of forests which are among the most powerful machine learning algorithms available today. [7]
Decision trees have two strengths over the majority of algorithms: the resultant model can readily be plotted and interpreted by non-technical people (at least for small trees), and the algorithms are completely scale-independent in handling data scaling. Since only one feature is handled at a time and each data partition does not rely on data scaling, pre-processing such as normalization or feature standardization is not required in decision tree algorithms. Specifically, decision trees are suitable when we have features at completely different scales, or both binary and continuous features. The worst disadvantage of decision trees is that they over fit and provide poor generalization performance. That is why, in most applications, ensemble techniques like random forests described below are usually used rather than a single decision tree.

2.6.5 Random forests:

As we have just seen, one of the biggest drawbacks of decision trees is that they over-fit to training data. Random forests are a way of correcting this problem. A random forest is simply

an ensemble of decision trees, each one a little different from the last. The idea is that each tree will be able to make a reasonably good prediction, but will over-learn some part of the data. If we build lots of trees, all of which do well and complement one another. In several ways, we can reduce the number of over additions by averaging out their predictions. This reduction in overfitting, without compromising the predictive power of the trees, can be demonstrated by extensive mathematical computation.

To implement this strategy, we need to build many decision trees, every tree must predict the target in an acceptable way and also must be different from the other trees. Random forests owe their name to the application of randomness in constructing trees so that every tree is different. Trees within a random forest are randomized in two aspects: by selecting the data points, with which a tree is constructed, and
by selecting the data points with which a tree is constructed. Selecting the features of every division test.
Random forests is the most popular algorithm of the set appentising technique "bagging" or "bootstrap aggregating" which includes building several instances of an same model (decision trees) and to train every one of these trees over a random subset of a data set, after training every tree we can aggregate the result of in to make the prediction, Figure 2.8 shows the "bagging" technique.

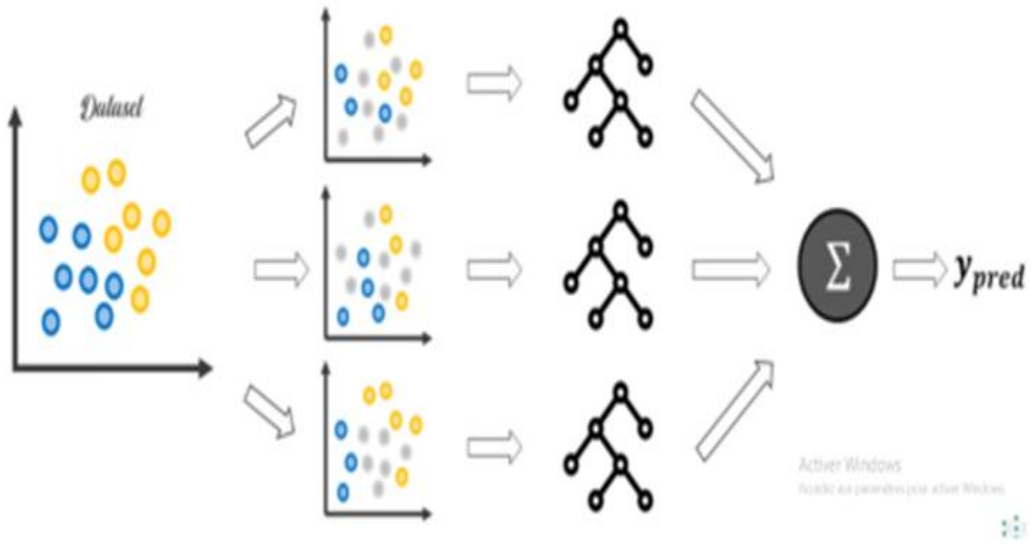


Figure2.8: **Bagging**[1].

[1] https ://machinelearnia.com/

2.6.6 Gradient Boosting:

Boosting is any ensemble method that can be used to combine multiple learners' weak learners. The overall idea behind most boosting methods is to train the models one at a time, with each model trying to correct the errors of its last iteration, so we end up with complementary models. There are various Boosting methods, one such being gradient Boosting which builds models one by one in an empty set, each correcting the last model. Rather than modifying instance weights with each iteration, this method attempts to modify the new model to the residual errors of the previous predictor.

Gradient Boosting works well with regression problems. It is called gradient Boosted Regression Trees (GBRT) when training a GBRT model. In order to search for the optimal number of trees, we can use an Early Stopping to avoid, among others, over fitting, the XGBoost library supports the capability to implement Gradient methods boosting, and Figure 2.9 depicts a GBRT model using an Early Stopping to search for the optimal number of trees.

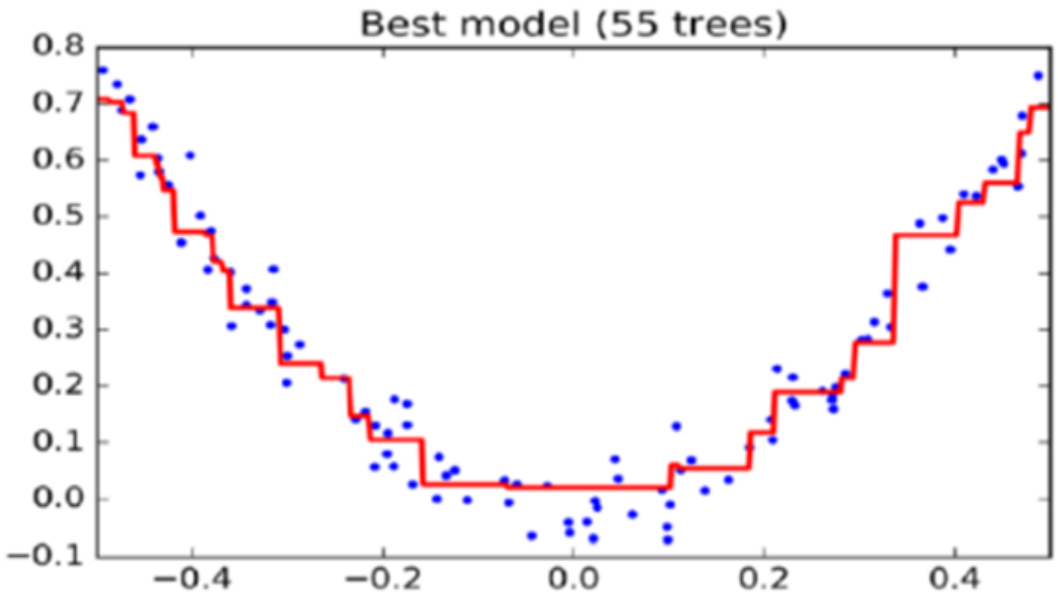


Figure2. 9: GBRT with Early Stopping. [2]

2.7Testing and validating a Machine Learning Model:

The most effective way to be able to see how generalizable a model is to new cases is to actually try it out on new cases. One method to do so is to put the model into production and monitor how it's doing. It is working okay, but if the model is abysmally bad, users will complain – not the best plan. A better way is to divide the data into two groups: the training set and the test set. As the names imply, the model is trained on the training set and tested on the test set. The proportion of error on new cases is called generalization error (or out-of-

sample error), and by running the model over the test set, we obtain an estimate of this error. If training error is zero (i.e., the model predicts very few errors on the training set) but the generalization error is enormous, it means that the is overfitting on training data. Training on 80% of data and keeping 20% for test is common.

To test a model is therefore very simple: you just need to pass it through a series of tests. Let us now consider the case that we are hesitating between two models (say a linear model and a polynomial model): how do we proceed? One approach is to train both models and to compare their generalization according to the series of tests. Suppose we find the best hyper parameter value that gives us a model with the lowest generalization error, e.g., an error of only 5%.

We deploy this model into production, but it sadly does not do as well as expected and produces 15% errors. What just happened? The problem is that the generalization error has been measured several times on the test set, and that the model and hyperparameters have been optimized to produce the best possible model. That means the model will not do as well on new data.

One standard approach to this problem is to have a second set of blocking referred to as validation set. We train multiple models with different hyperparameters on training set, we select the model and the hyperparameters that yield the best performance in the validation set, and once we are satisfied with the model, we perform a single final test against the test set to obtain an estimate of the generalization error.

To avoid "wasting" too much of the training data in validation sets, a common trick is to employ cross-validation: the training set is divided into complementary subsets, and each model is trained on a different combination of these sub-assemblies and tested against the rest. Once the model type and the hyperparameters are selected, the last model is trained with the hyperparameters on the whole training set, and the error that is generalized is tested on the test set.

2.8Conclusion:

Machine learning has emerged as a cornerstone of modern technology, revolutionizing industries with its data-driven approaches and adaptive algorithms. From its early conceptual roots to the resurgence of deep learning, the field has seen rapid advancements driven by increasing computational power and data availability.

The various learning paradigms—supervised, unsupervised, semi-supervised, and reinforcement learning-offer flexible solutions for a wide array of tasks, supported by powerful algorithms such as linear regression, decision trees, and support vector machines.

Proper validation and testing techniques, including cross-validation, are crucial to ensure model generalization and prevent overfitting. As machine learning continues to evolve, its integration into everyday applications will only grow, making a strong foundational understanding of its principles more important than ever.

CHAPTER03: Experimental Overview of Aluminium Waste Utilisation in Concrete

3.1Introduction:

Concrete is a fundamental material used in modern construction, but its manufacture comes at a very high environmental cost in terms of carbon emissions when cement is manufactured. To counter this, scientists have been trying to find alternatives by adding industrial waste materials like aluminum waste (AW) to concrete mixtures. Not only does this reduce environmental waste, but it also enhances certain mechanical properties of concrete. During the past ten years, AW has been investigated for its effects as a partial replacement for cement or as reinforcing fibers, with promising outcomes in strength, durability, and thermal insulation. In addition, advancements in machine learning and statistical modeling have enabled the formulation of better predictions of concrete behavior, enabling optimized mix designs for specific applications. This chapter provides an overview of pioneering experimental work on AW-concrete and the positive impact of pioneering methods towards encouraging sustainable construction.

3.2Experimental Overview of Aluminium Waste Utilisation in Concrete:

3.2.1 The Use of Aluminum Waste for Concrete Production (2011)

Elinwa and Mbadike investigated aluminum waste (AW) as a partial cement replacement (5–40%) in concrete and found that it functioned as a retarder, well suited for hot weather concreting. The best replacement was 10%, which gave maximum compressive and flexural strength. Statistical models validated significance. [8]

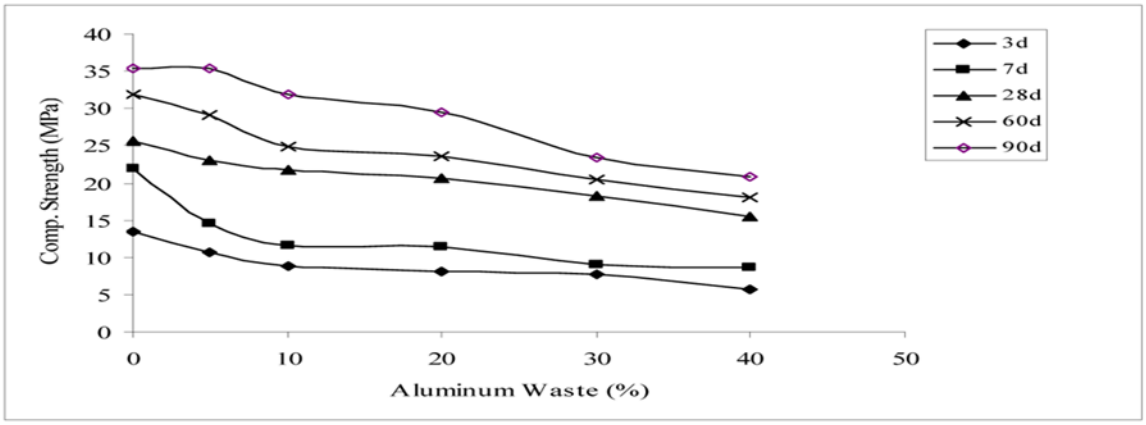


Figure3. 1: Compressive strength of AW-Concrete. [8]

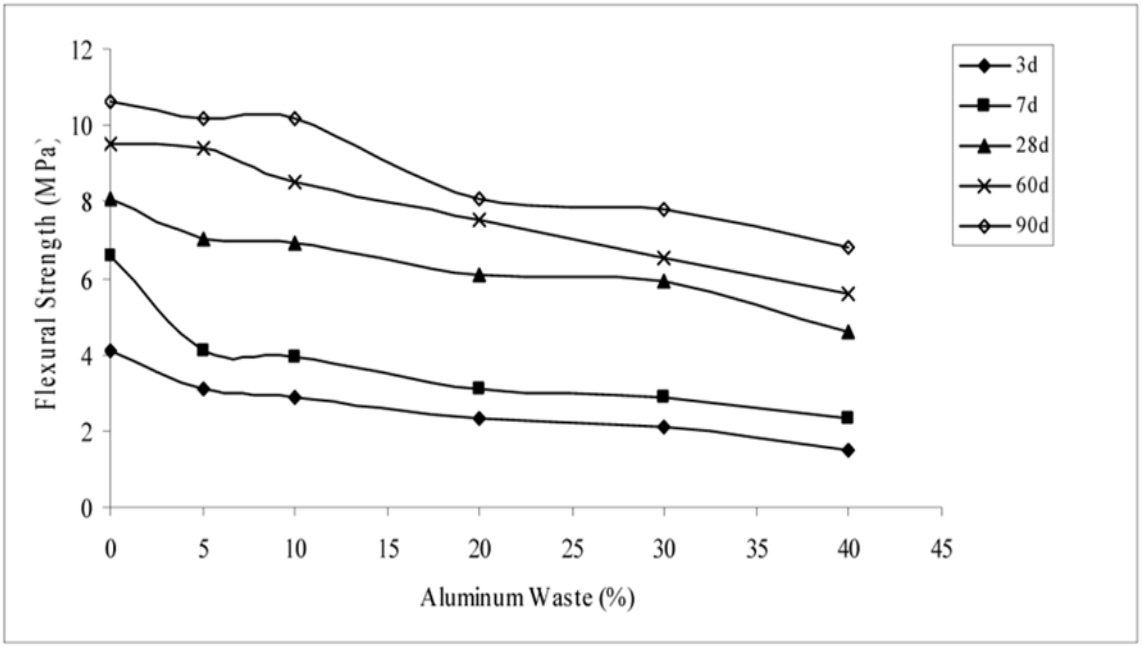


Figure3. 2: Flexural strength of AW-Concrete. [8]

3.2.2 Prediction of the Compressive Strength of Aluminum Waste–Cement Concrete Using Scheffe's Theory (2012)

Arimanwa et al. modeled AW-cement concrete strength, predicting 28.832 N/mm² at 10% AW. AW reduced setting times (initial: 36–58 min; final: 435–518 min) and workability but maintained constant density. The model passed in t- and f-statistic tests. [9]

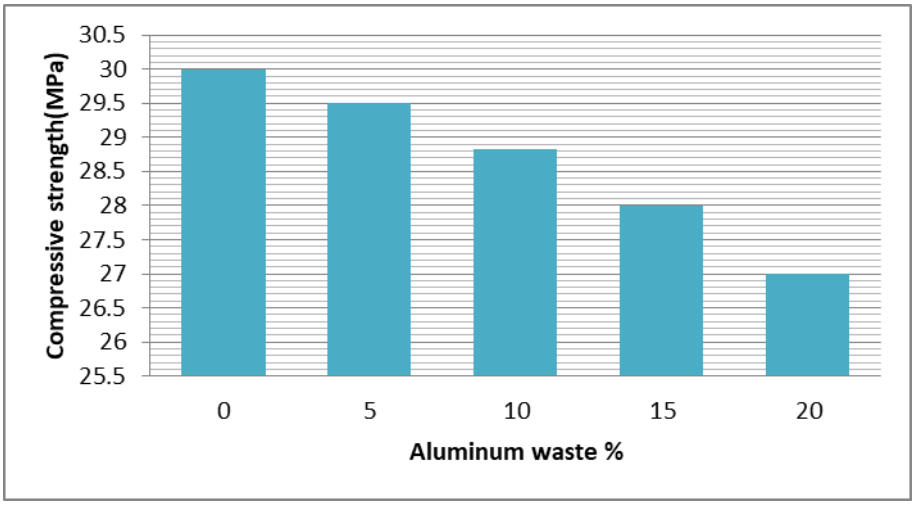


Figure3. 3: Compressive strength of AW-Concrete. [9]

3.2.3 Comparison of Data Mining Techniques for Predicting Compressive Strength of Environmentally Friendly Concrete (2016)

This study evaluated nine data-mining models, including advanced techniques like Gaussian processes and ensemble methods, to predict the strength of eco-friendly concrete. Gaussian processes regression achieved the highest accuracy, while ensemble methods improved the performance of regression trees but not advanced models like SVM. The findings provided insights into model selection for sustainable concrete applications. [10]

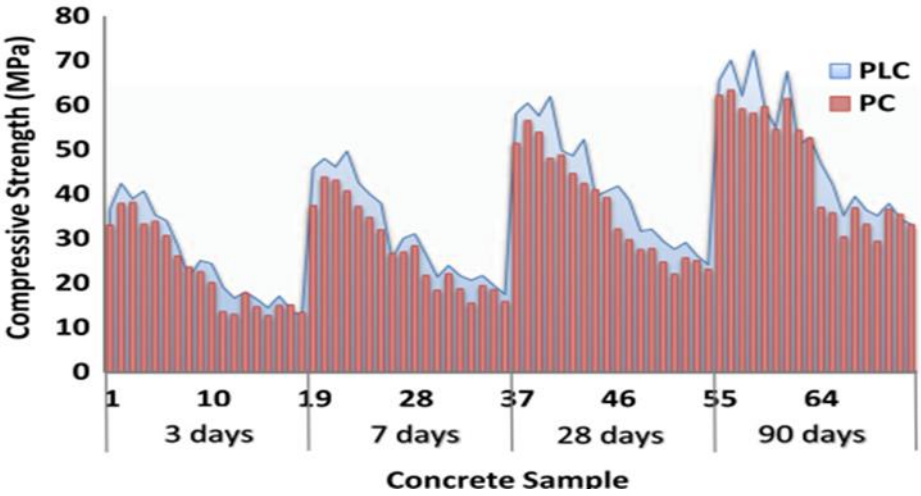


Figure3. 4: Experimental result for compressive strength of PC and PLC concrete. [10]

3.2.4 Mechanical behaviour of fibre reinforced concrete using soft – drink can (2017)

Ilya and Cheow Chea added recycled aluminum can fibers (0–2% of cement weight) to concrete. 1% fibers exhibited the highest compressive/flexural strength, which were effective as reinforcement. [11]

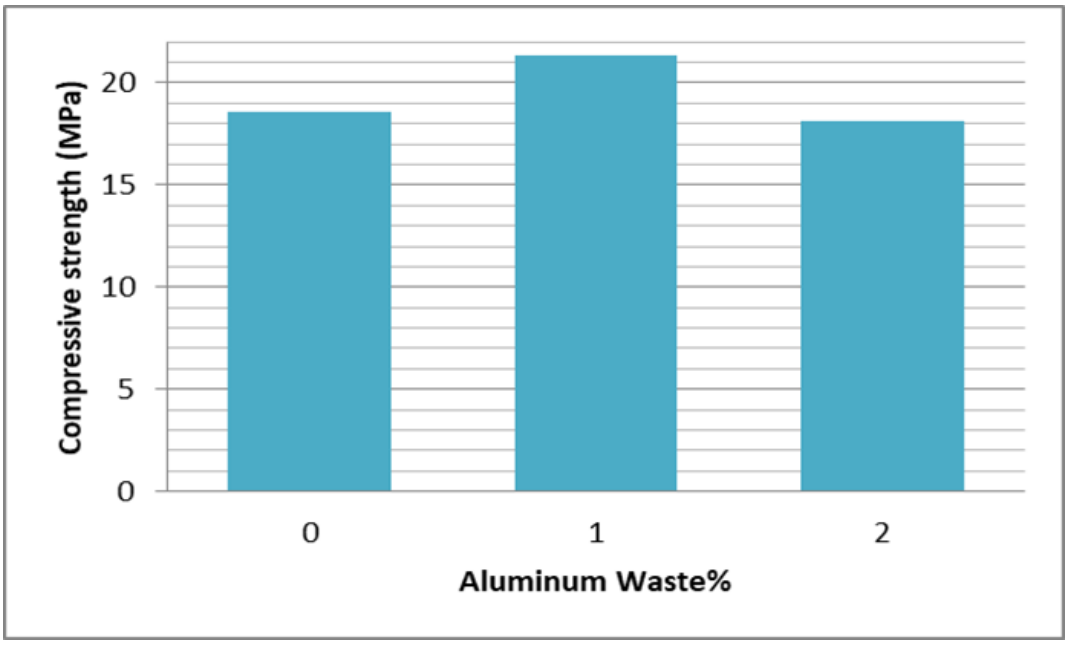


Figure3. 5: Compressive strength of AW Fiber in 7 days. [11]

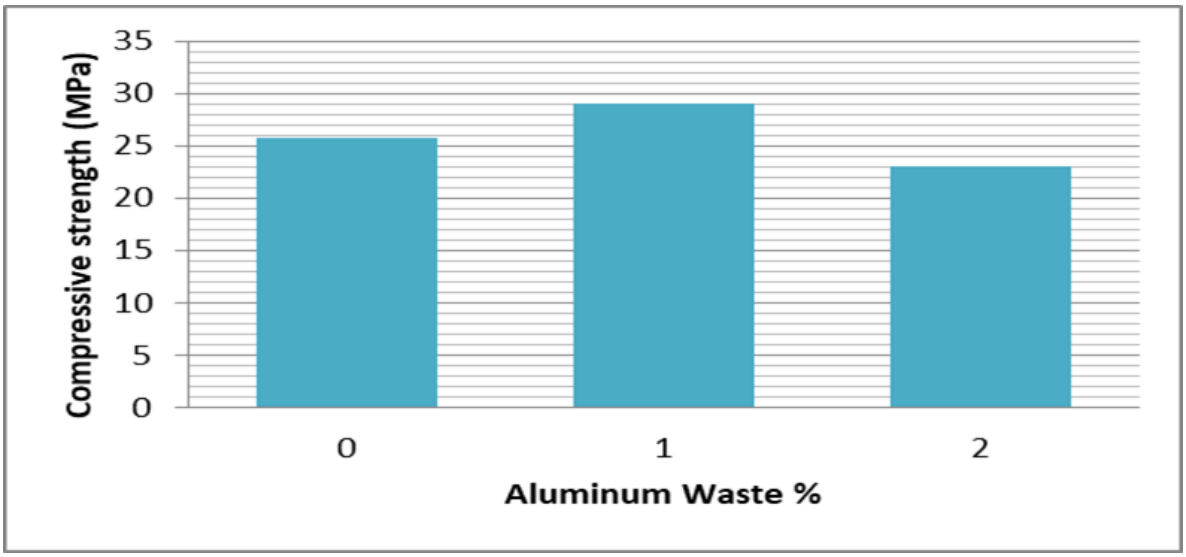


Figure3. 6: Compressive strength of AW Fiber in 28 days. [11]

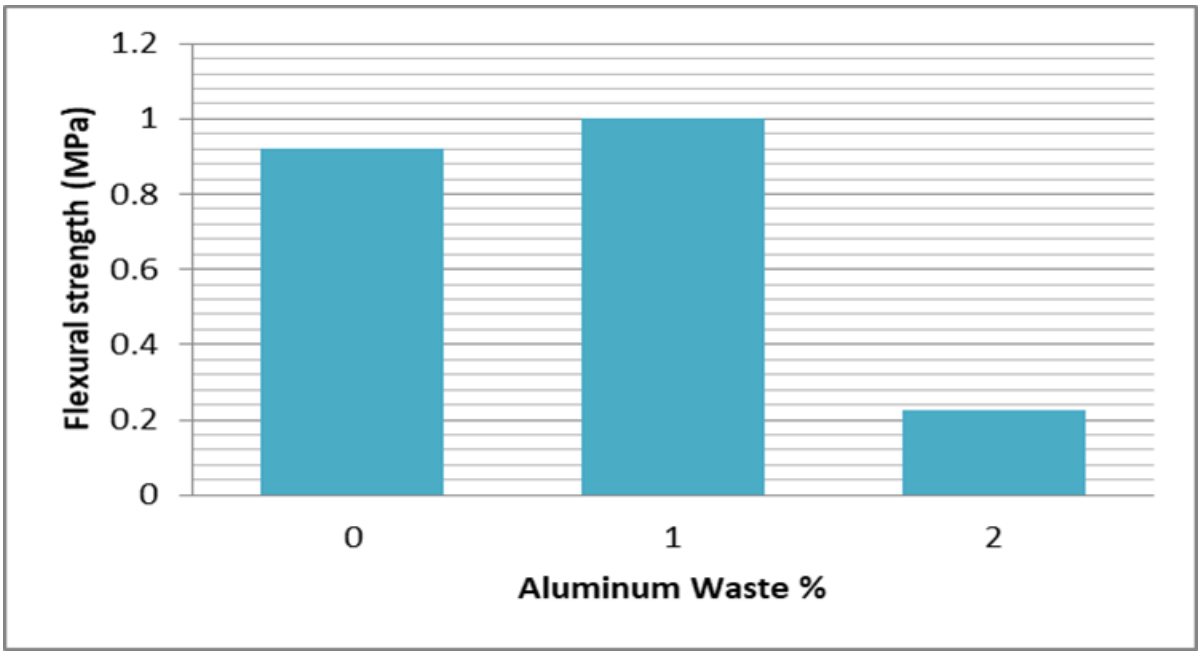


Figure3. 7: Flexural strength of AW Fiber in 7 days. [11]

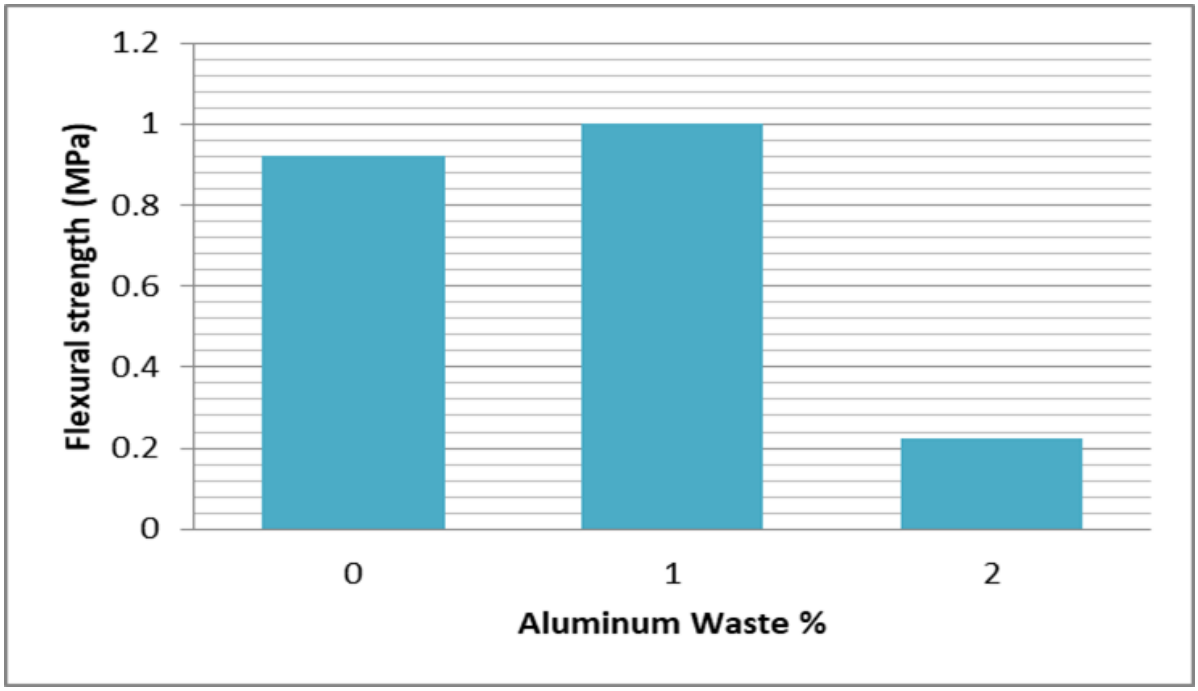


Figure3. 8: Flexural strength of AW Fiber in 28 days. [11]

3.2.5 Investigating the Behavior of Concrete and Mortar Reinforced with Aluminum Waste Strips (2018)

Muwashee et al. used reinforced concrete/mortar with 2.5% aluminum strips from tin cans, achieving 22% higher compressive strength and flexural strength of up to 11.20 MPa (from 3.31 MPa). Mortar showed >100% improvement in tensile/flexural strength. [12]

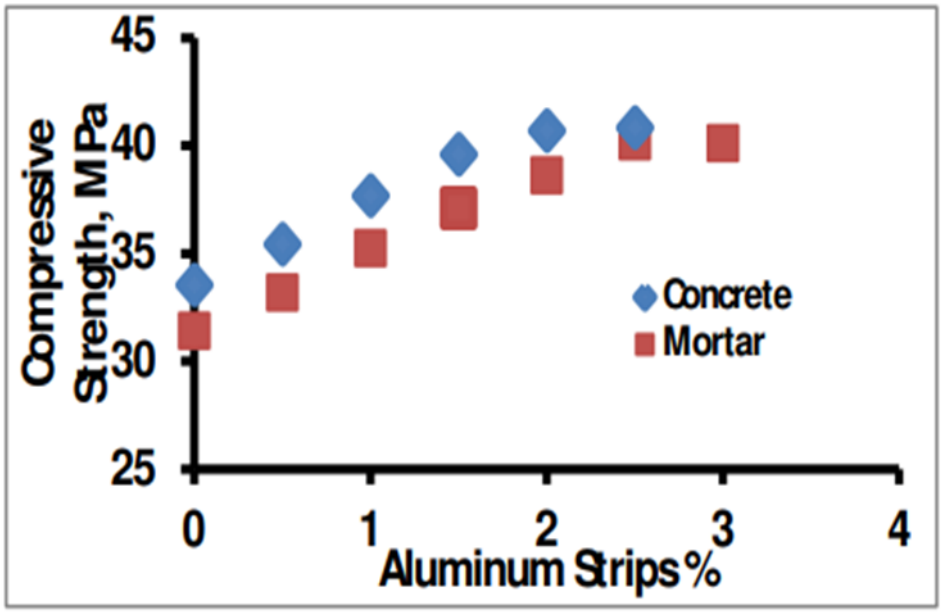


Figure3. 9: Effect of Aluminum strip % on the compressive strength for both concrete and mortar mixes. [12]

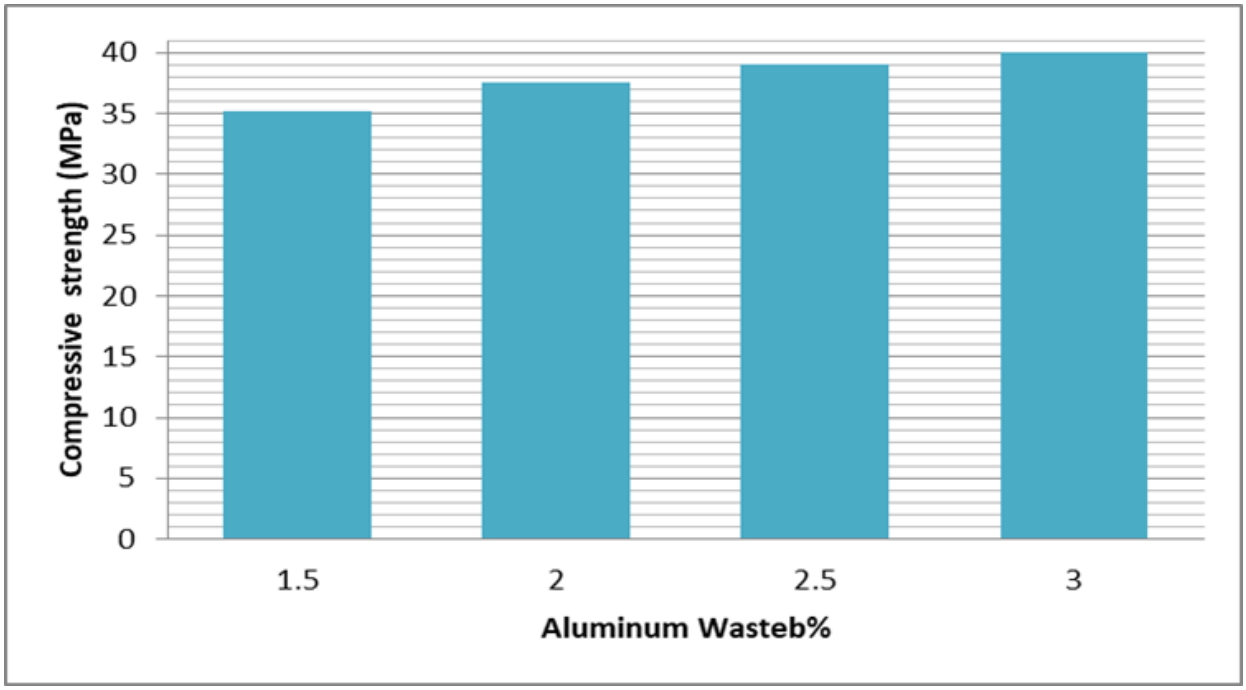


Figure3. 10: Compressive strength of AW Concrete in 28 days. [12]

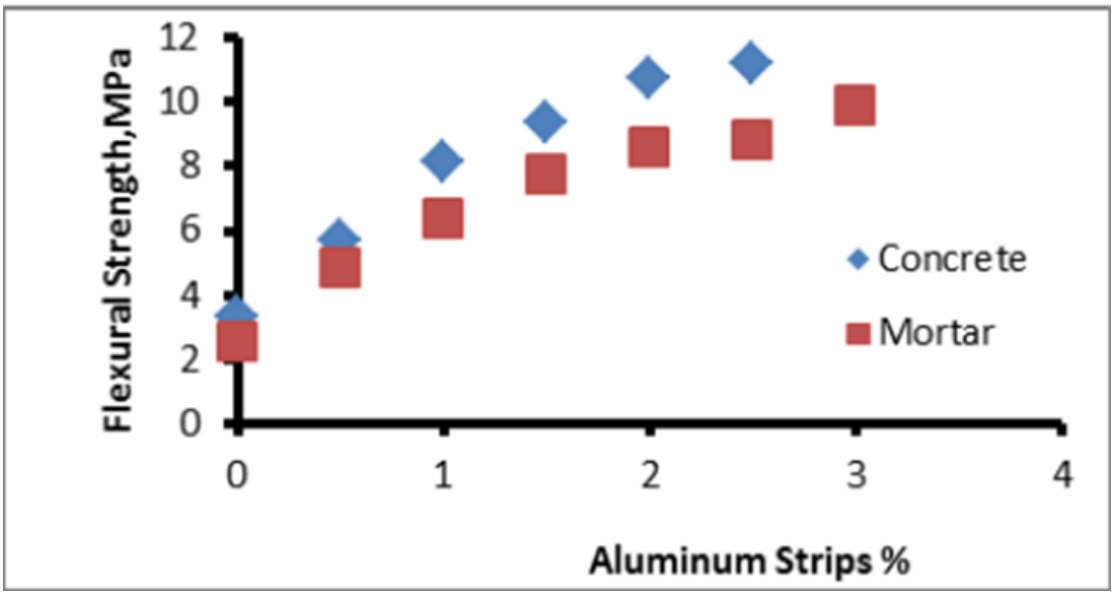


Figure3. 11: Effect of Aluminum strip % on the flexural strength for both concrete and mortar mixes. [12]

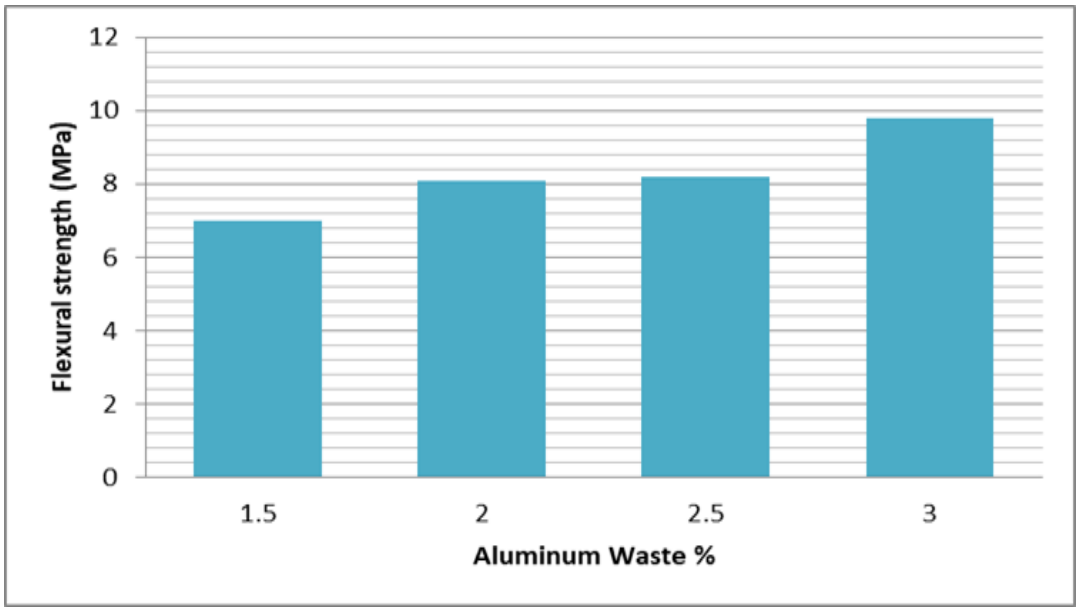


Figure3. 12: Flexural strength of AW Concrete in 28 days. [12]

3.2.6 Mechanical properties of pavement quality concrete with secondary aluminium dross as partial replacement for ordinary Portland cement (2018)

Panditharadhya et al. replaced cement with aluminum dross (5–20%). 15% dross exhibited better mechanical properties, enhanced initial setting time, and reduced final setting time, suitable for hot climates [13]

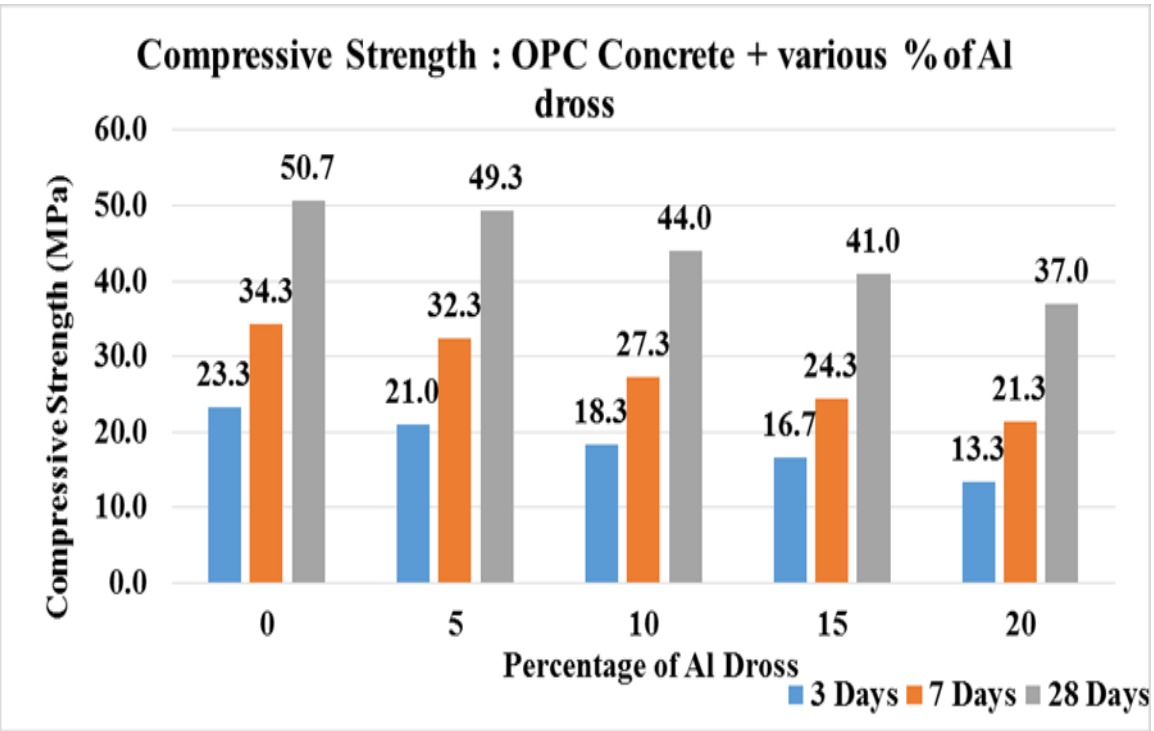


Figure3. 13: Variation of compressive strength for different replacements of aluminum dross. [13]

3.2.7 Effect of waste aluminium shavings on the bond characteristics of laterized concrete (2019)

Ofuyatan et al. determined that 1–2% shavings of aluminum increased bond strength in laterized concrete but reduced compressive strength. The smaller bars (16 mm) had higher bond resistance. [14]

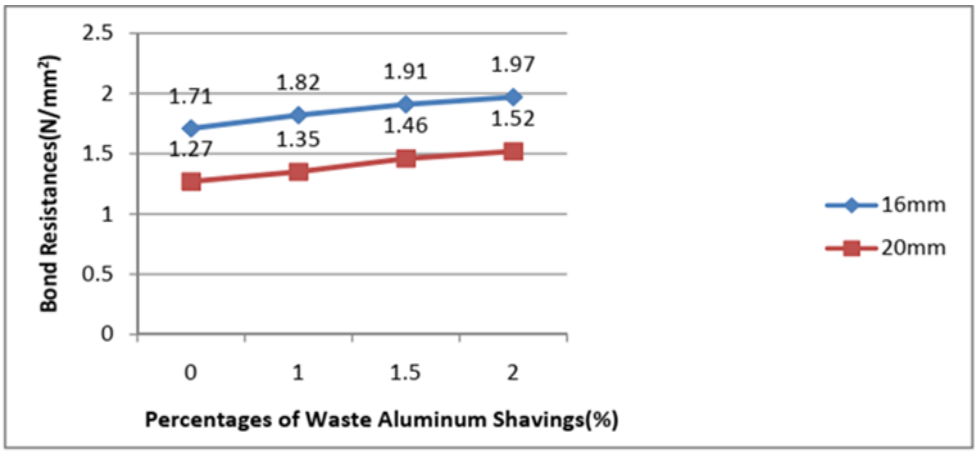


Figure3. 14: Bond Resistance of Laterized Concrete with Varying Percentages of Waste Aluminum Shaving (16 mm and 20 mm Diameter bars). [14]

3.2.8 Experimental Study on Strength Properties of Aluminium Fibre Reinforced Concrete (2019)

Sabapathy et al. tested looped-end aluminum fibers (0–2% by volume). 1% fibers enhanced compressive/split tensile strength, with regression models validating performance. [15]

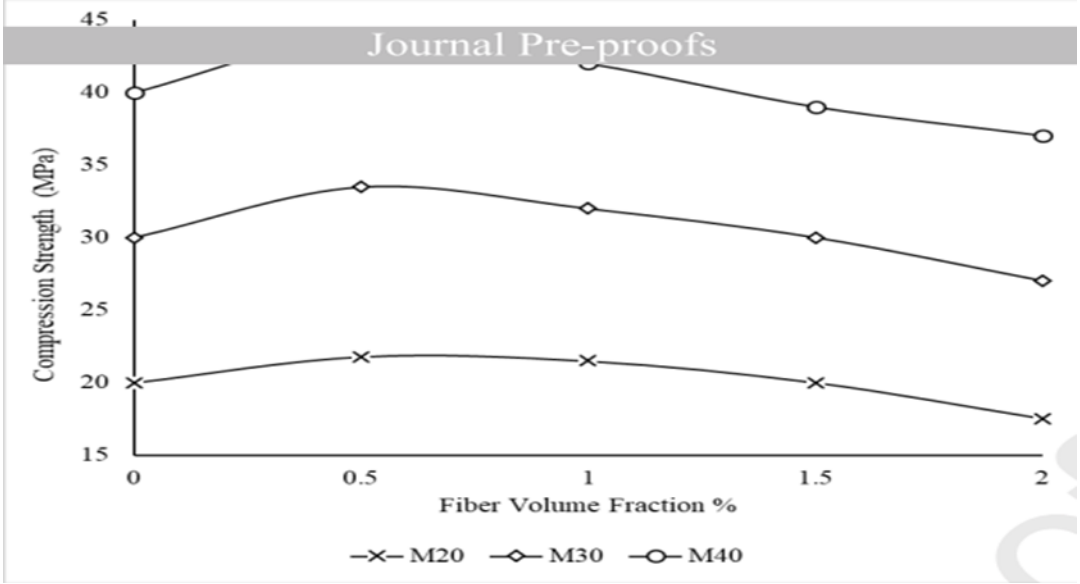


Figure3. 15: Effect of fiber volume on compressive strength.[15]

3.2.9 Influence of Aluminum Waste on the performance of fluid sand Concrete (Influence de déchets d'aluminium sur les performances de Béton de sable fluide) (2020)

Mohamed and Ahmed replaced sand with AW (0–30%), reporting decreased fluidity/density but increased thermal insulation, which indicates potential application in insulation materials. [16]

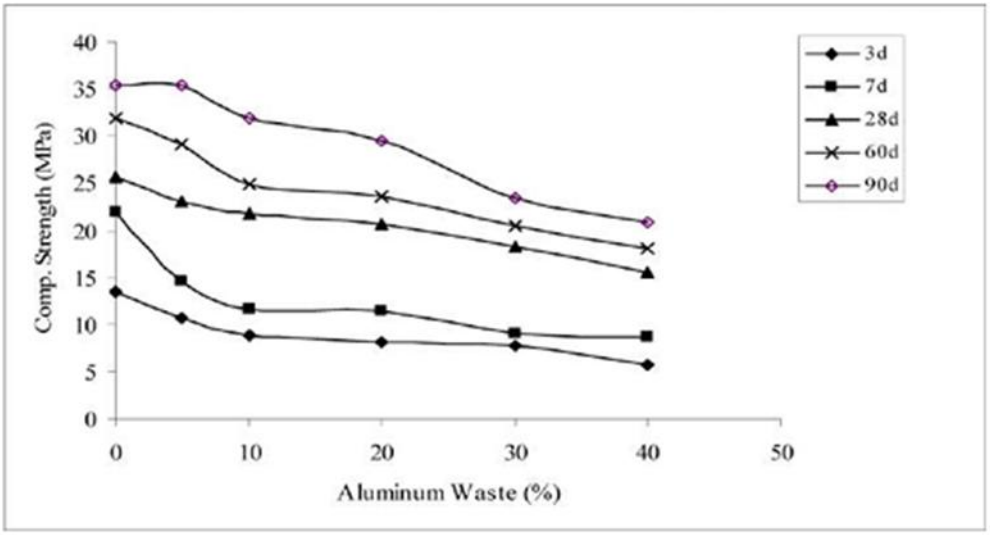


Figure3. 16: The effect of Aluminum Waste on the compressive strength of Concrete. [8]

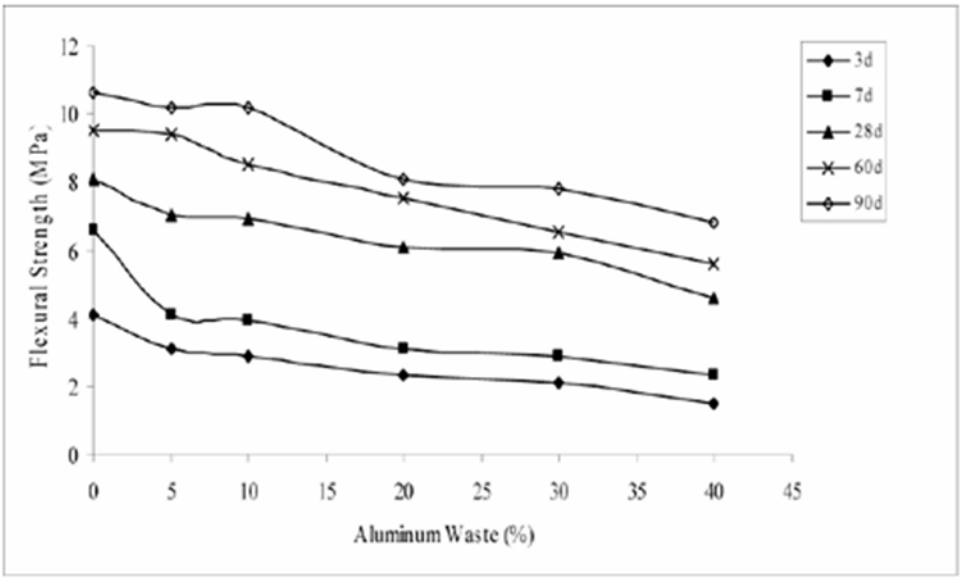


Figure3. 17: The effect of Aluminum Waste on the flexural strength of Concrete. [8]

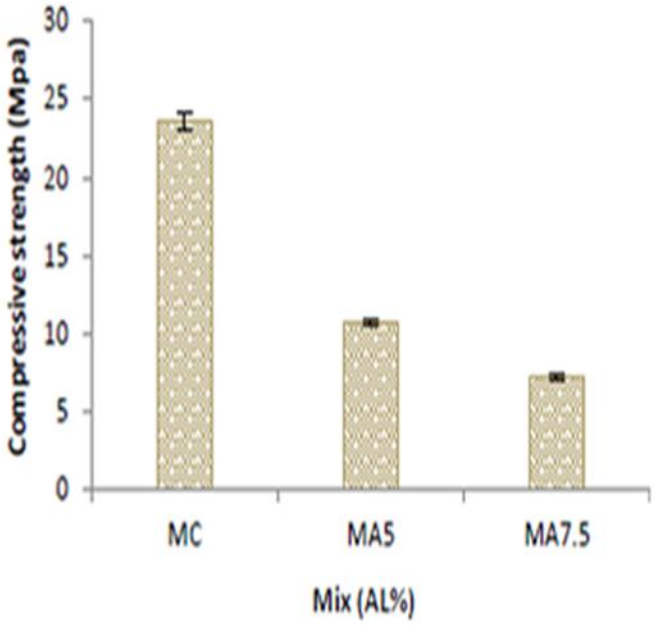


Figure3. 18: The effect of Aluminum Waste on the flexural strength of Mortar. [17]

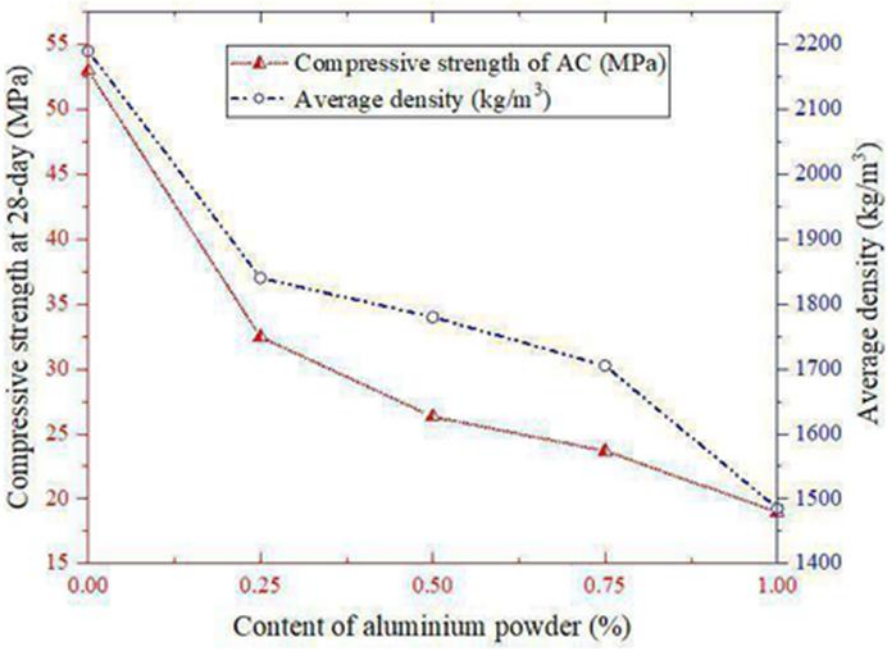


Figure3. 19: The effect of Aluminum Waste on the tensile strength of Concrete. [18]

3.2.10 Mechanical Behavior of Concrete Reinforced with Waste Aluminium Strips (2021)

Channa and Saand used soft drink tin fibers (1–5% by weight of cement). 4% fibers provided the highest split tensile/flexural strength, beyond which the performance declined. [19]

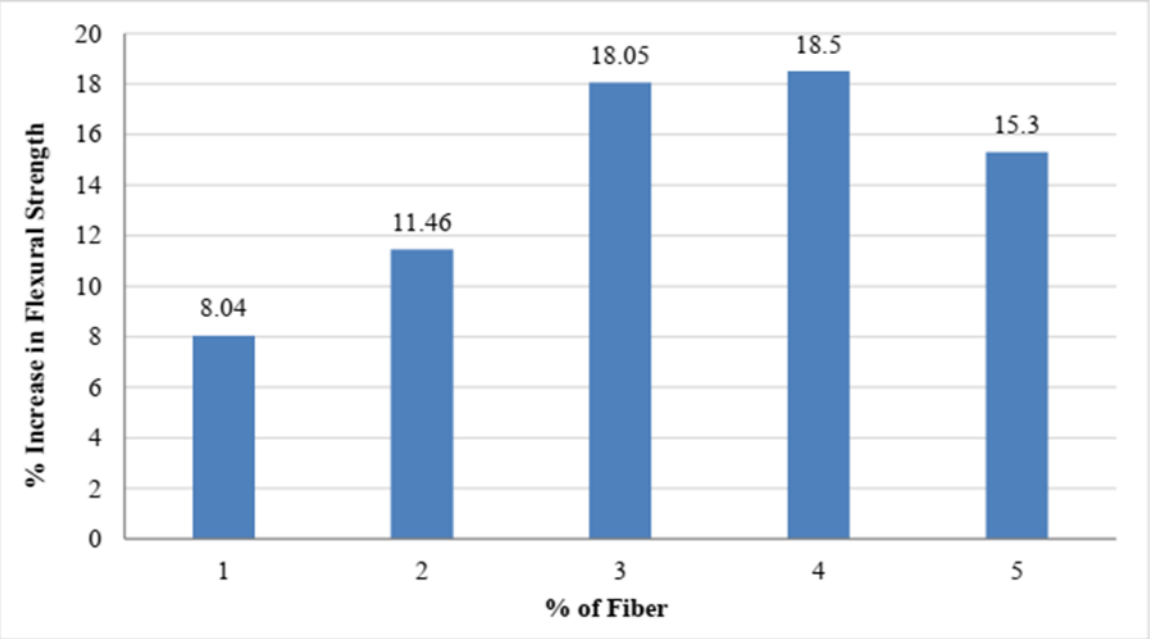


Figure3. 20: Comparison of percentage increase of flexural strength With percentage increase of fiber. [19]

3.2.11 Effect of industrial waste metal chips on flexural behavior of reinforced concrete beams (2023)

Aslan et al. utilized metal chips from industries in RC beams, achieving 1.75× load capacity (steel chips) and improved ductility over control. [20]

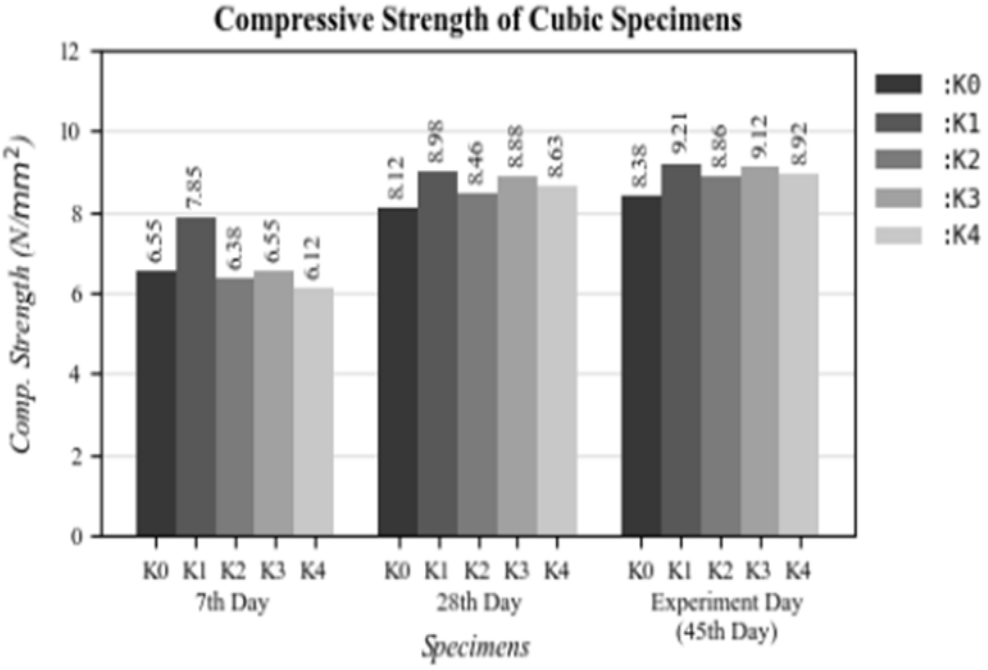


Figure3. 21: Compressive strength of standard cubic concrete specimens for 7th day, 28th day, and experiment day (45th). [20]

3.2.12 Evaluation of sisal fiber and aluminum waste concrete blend for sustainable construction using adaptive Neuro Fuzzy Inference System (2023)

ANFIS Model for AW-Sisal Fiber Concrete.

Agor et al. optimized AW-sisal fiber mixtures through AI. 10% sisal + 0.1% AW yielded 24.97 N/mm² compressive strength; ANFIS model yielded 99.57% accuracy. [21]

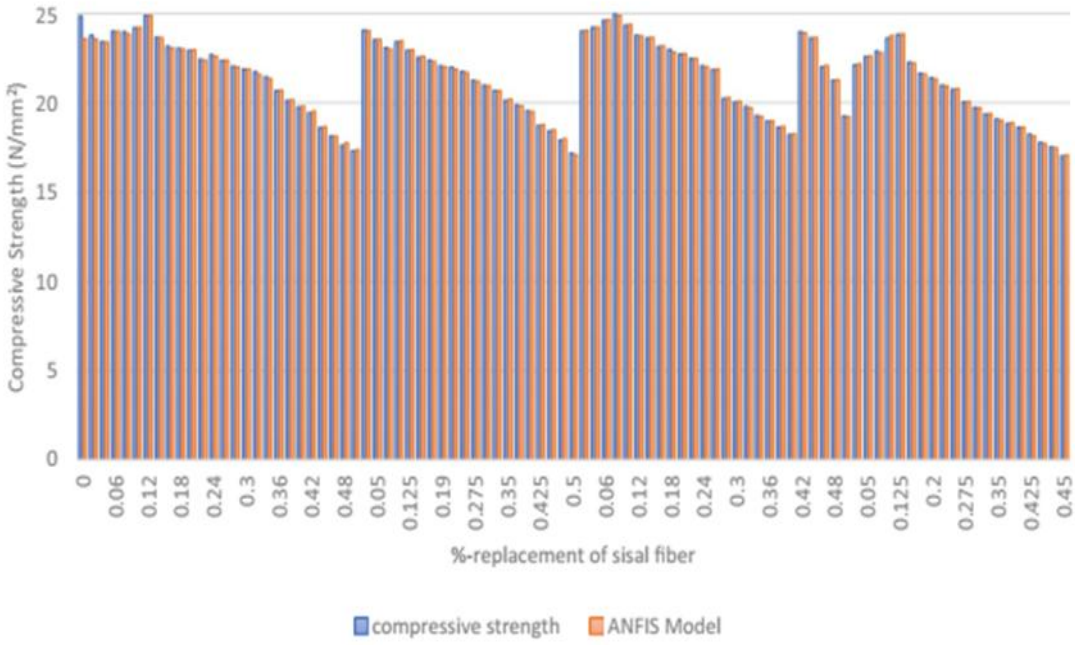


Figure3. 22: Experimental and ANFIS results. [21]

3.2.13 The effect of Aluminum Waste Ash on the concrete compressive strength (2023)

Aluminum Waste Ash in Concrete.

Suryadi and Rochman replaced sand with AW (10–30%), achieving 31.78–60.49% less compressive strength due to AW's smoothness, deeming it unsuitable. [22]

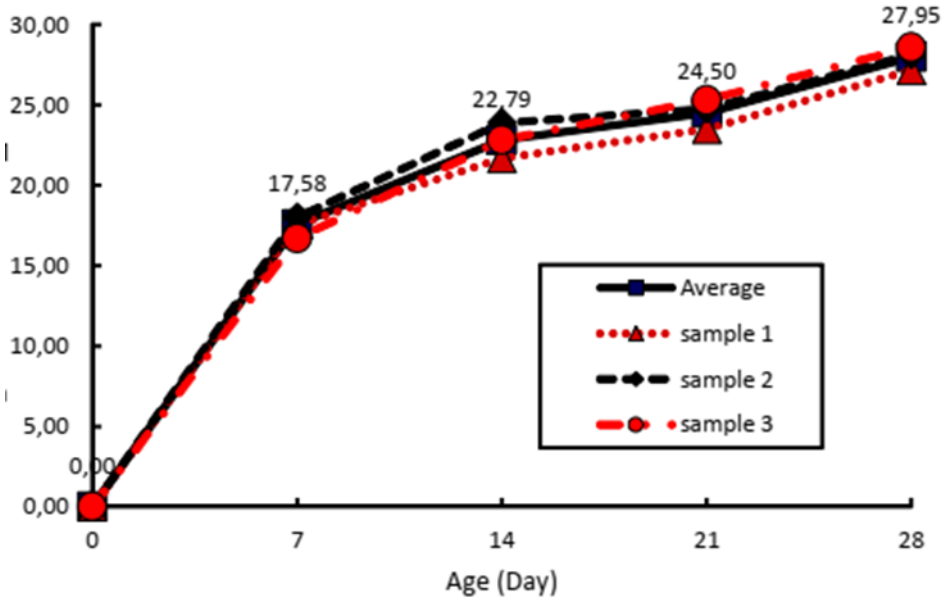


Figure3. 23: Concrete compressive strength to age at w/c=0.474. [22]

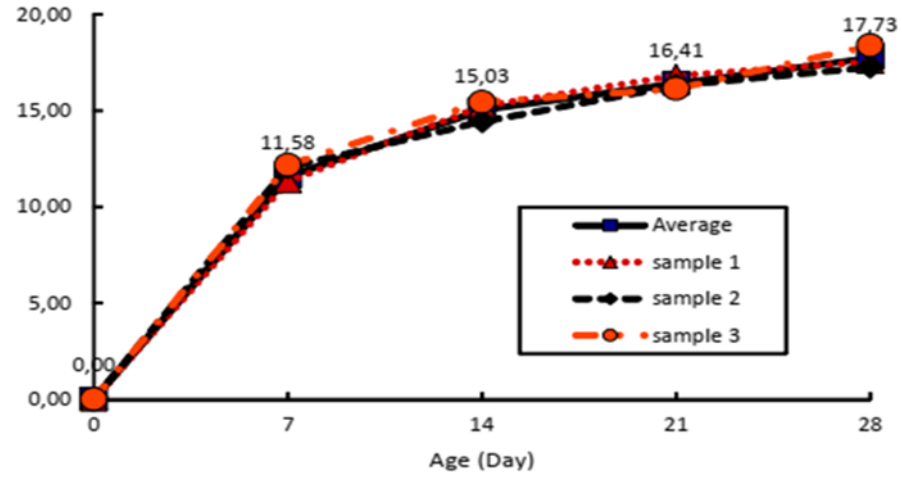


Figure3. 24: Compressive strength at 10 % AWA concrete (w/c=0.525). [22]

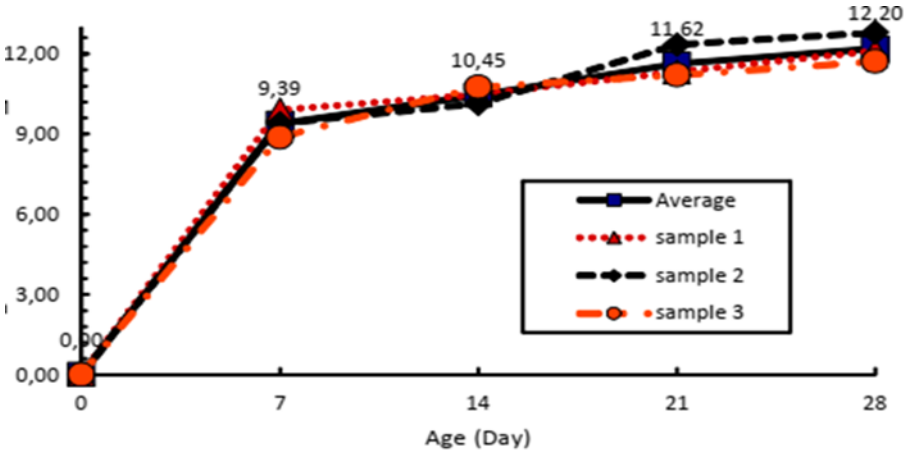


Figure3. 25: Compressive strength at 20% AWA concrete (w/c=0.575). [22]

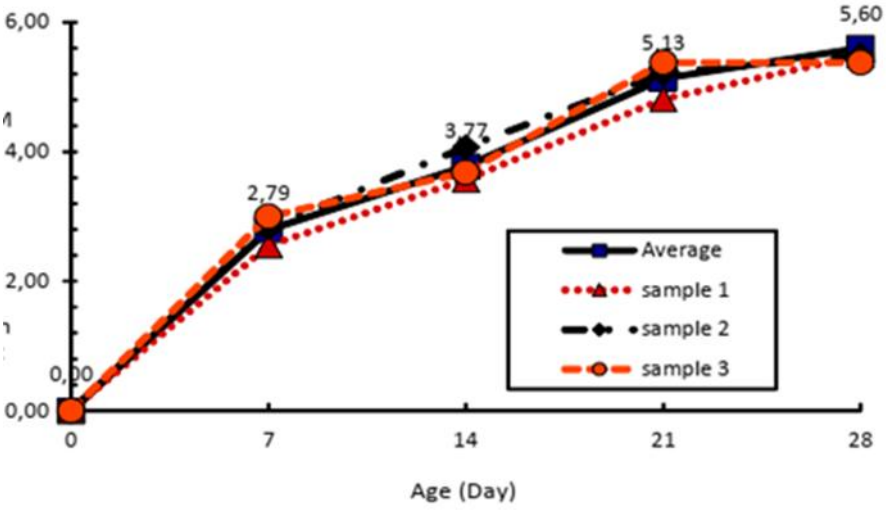


Figure3. 26: Compressive strength at 30% AWA concrete (w/c=0.611). [22]

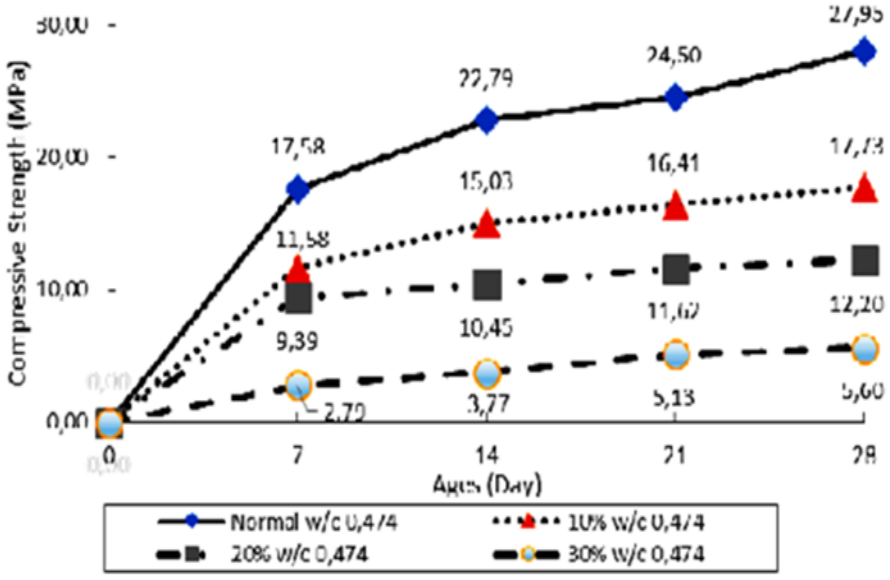


Figure3. 27: Compressive strength and age of concrete with varied w/c. [22]

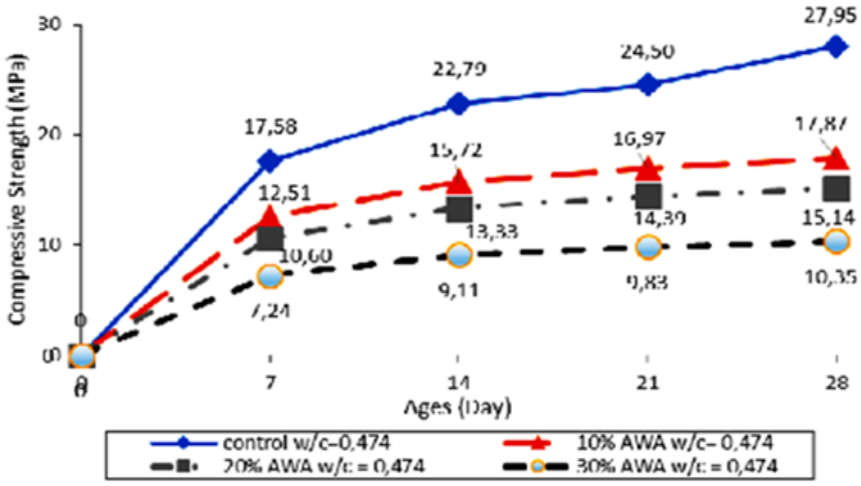


Figure3. 28: Compressive strength and age of concrete specimen with constant w/c. [22]

3.2.14 Use of experimental design methodology for the formulation of waste-based concrete (Utilisation de la méthodologie des plans d'expérience pour la formulation de béton à base de déchets) (2023)

Design-of-Experiments for AW Concrete Selmani Khalil utilized statistical methods (96 trials) to show 5% AW marginally superior strength, however, higher doses underperformed regular concrete. [23]

Table3. 1: Compressive strength of AW-Concrete. [23]

Al %	Age (Days)	Compressive strength (MPa)
0	28	20.27
5	28	19
5	28	7
0	28	17.57
0	28	24.73
5	28	10
0	28	27.81
5	28	11
0	28	14.35
5	28	19
5	28	8
0	28	20.3
0	28	20.34
5	28	11.3
5	28	11
0	28	19.03
5	28	17
0	28	32.99
5	28	6.2
0	28	32.27
0	28	31.94
5	28	20
0	28	16.7
5	28	10.1
5	28	11.2
5	28	12
0	28	38.44
0	28	33.02
5	28	6.3

5	28	9
0	28	34.74
0	28	18.19
0	28	28.54
0	28	33.53
5	28	11
0	28	17.34
5	28	5.5
5	28	8.5
0	28	22.57
5	28	6.8
0	28	23.38
5	28	10
0	28	46.5
0	28	38.34
5	28	7.5

0	28	24.07
5	28	4.5
5	28	11
5	28	7
5	28	22
0	28	33.7
0	28	34.75
5	28	14
5	28	13.1
0	28	31.48
0	28	32.09
5	28	6.5
0	28	14.19
0	28	30.84
0	28	28.9
0	28	31.4
5	28	11

5	28	12
0	7	11.3

3.2.15 Study on Properties of Concrete Using E-Waste Copper, Aluminium and Steel Fibers (2024)

E-Waste Fibers Reinforced M50 Concrete.

Gohil and Patel had experimented with e-waste fibers (steel, Cu, Al). 1% steel fibers achieved 66.89 N/mm² compressive strength; copper fibers were best in tensile/flexural strength with 1.5%. [24]

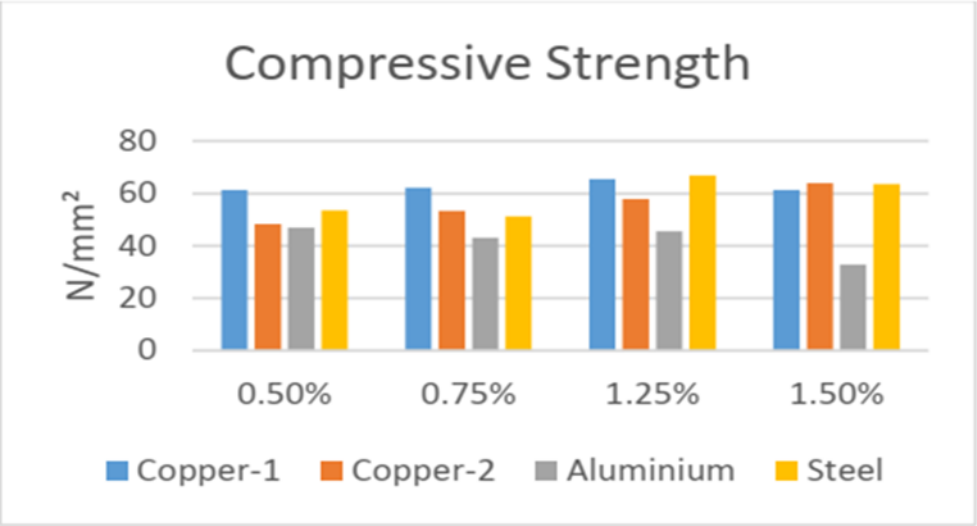


Figure3. 29: Compressive strength of E-Waste Fibers. [24]

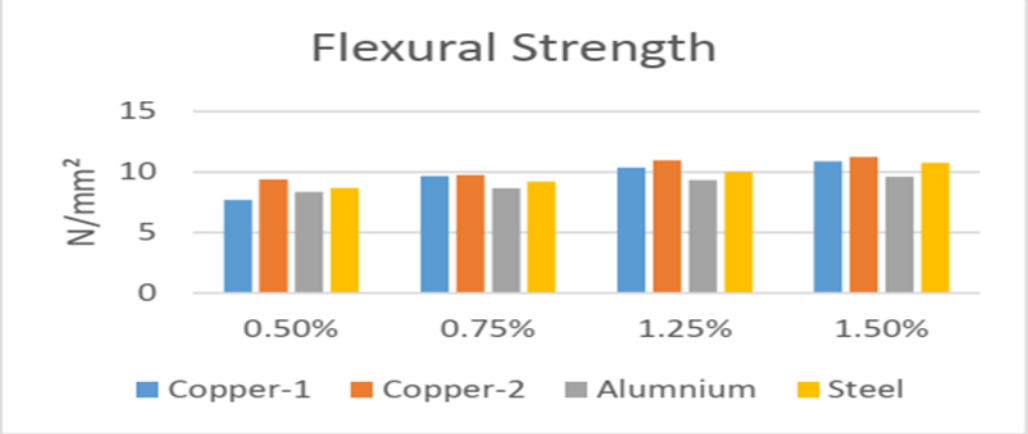


Figure3. 30: Flexural strength of E-Waste Fibers. [24]

3.2.16 Compression-Flexure Relationship in Fibre Concrete Based on Wasted Aluminium (2024)

Hafiane Ahmed Anisse experimented results demonstrated a clear relationship between the compressive and flexural strengths of self-compacting concrete reinforced with recycled aluminum fibers. The inclusion of aluminum fibers at volume fractions ranging from 1% to 5%.

The study confirms that recycled aluminum fibers can effectively enhance the mechanical properties of self-compacting concrete, but only within a limited range (≤3%).

Higher fiber contents disrupt the concrete matrix, leading to premature failure. These findings provide practical guidance for optimizing fiber-reinforced concrete mixtures. [25]

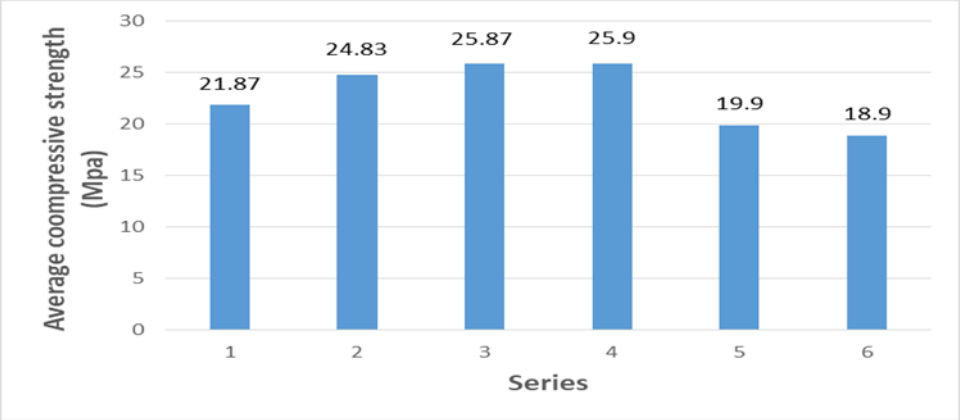


Figure3. 31: Comparison of average compressive strength of test specimen. [25]

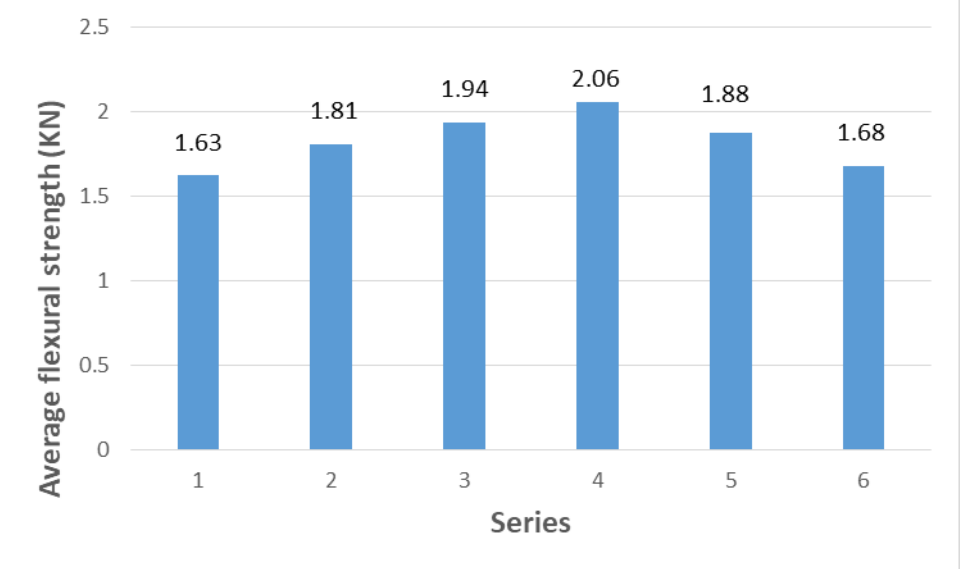


Figure3. 32: Comparison of average flexural strength of test specimen. [25]

3.3 Conclusion:

The use of waste aluminum in concrete has proven to have significant potential for improving mechanical characteristics while bringing about sustainability. Research has established that

optimal AW dosages, typically 5–15% for cement replacement and 1–4% for fiber reinforcement, can be utilized to enhance compressive, flexural, and tensile strength, particularly in special applications like hot-weather concreting or thermal insulation products. More advanced modeling approaches, such as machine learning, ANFIS, and statistical theory like Scheffe's, have also formulated the potential to forecast and maximize concrete performance. Despite the obstacles, such as reduced workability or strength with higher AW percentages, waste elimination and enhanced property are an ample reward to present AW-concrete as a very attractive option to environmentally conscious building. Future studies should aim at long-term performance, scalability, and integration of AW with other sustainable materials to expand its use in the construction sector.

CHAPTER04: Application of Machine Learning for Predicting Compressive and Flexural Strength of Waste Aluminium Fibre Concrete

4.1Introduction:

Machine learning application in material science has yielded new avenues for high-accuracy prediction of mechanical properties of material construction materials. In this study, an attempt is made to apply machine learning techniques for prediction of compressive and flexural strength of Waste Aluminium Fibre Concrete (WAFC). Through the utilization of deep learning techniques such as XGBoost, Random Forest, and Artificial Neural Networks (ANN), the research aims to establish robust correlations between input parameters (e.g., cement content, curing age, and fiber composition) and target mechanical properties.

The research begins with a proper data preprocessing step, including normalization and outlier identification, to furnish good-quality inputs to the models. Critical operations such as feature correlation analysis, hyperparameter tuning and cross-validation are employed to optimize model performance. Performance is evaluated based on typical measures of regression—Mean Absolute Error (MAE), Mean Squared Error (MSE), and R-squared (R^2)—to determine the optimum-performing predictive model. Results provide actionable insights to inform engineers in WAFC mixture design and optimization, reducing the application of costly and time-consuming experimental methods.

4.2 Libraries Used

4.2.1 NumPy:

NumPy stands for “Numerical Python” and is a fundamental package for scientific computing in Python. It provides a comprehensive mathematical library capable of performing numerical computations efficiently, enabling work with multidimensional arrays and matrix data structures, which are very common in data science and machine learning.

Why use NumPy?

- NumPy arrays use less memory than regular Python lists.
- A regular Python list is essentially a collection of pointers to separate Python objects (e.g., the numbers inside the list).
- A NumPy array is designed to be an array of uniform values, avoiding memory used by type pointers.
- NumPy can also read data faster and supports many useful broadcasting operations across array dimensions.

4.2.2 Pandas:

Pandas is a package for data manipulation and analysis in Python. The name Pandas is derived from the term “Panel Data.” Pandas introduce two additional data structures into Python: the Pandas Series and Pandas Data Frame. These structures allow us to work with labeled and relational data in a simple and intuitive way.

The recent success of machine learning algorithms is partly due to the large volumes of data available for training. However, when it comes to data, quantity is not the only factor—data quality is just as important. Large datasets are often not ready for immediate use in algorithms. Frequently, these datasets contain missing values, outliers, or incorrect data, all of which can impair the performance of machine learning models. Therefore, a crucial step in machine learning is to first examine the data and ensure it is suitable for training by performing basic data analysis. This is where Pandas becomes essential.

Pandas Series and **DataFrames** are designed for fast, flexible, and user-friendly data analysis and manipulation. Key features that make Pandas a great data analysis tool include:

- Support for row and column labels

- Ability to compute continuous statistics on time-series data
- Easy handling of NaN values
- Capability to load data from various formats into DataFrames
- Ability to join and merge different datasets
- Integration with NumPy and Matplotlib

For these and other reasons, Pandas DataFrames have become one of the most commonly used objects for data analysis in Python.

4.2.3 Matplotlib:

This is currently the most popular Python library for producing plots and other 2D data visualizations. Since data analysis often requires visualization tools, Matplotlib is a highly suitable library for this purpose.

With just a few lines of code in Python, you can easily generate line plots, histograms, bar charts, scatter plots, and more. Instead of spending time searching for solutions, you can focus on generating visualizations for faster data analysis and exploration.

Some of the key features that have made it the most widely used tool for graphical data representation include:

- Progressive development and interactive data visualization
- Fine control over graphical elements
- Export options to many formats such as PNG, PDF, SVG, and EPS

Built-in plotting features in Seaborn and Pandas are both based on Matplotlib.

4.2.4 Seaborn:

Seaborn is a visualization library built on top of Matplotlib, focused on creating standard statistical plots with a single-line function call. This simplicity comes at the cost of fewer customization options compared to raw Matplotlib.

However, it is still possible to modify Seaborn plots using Matplotlib's pyplot functions.

4.2.5 Scikit-Learn:

Scikit-learn is a machine learning library built on top of NumPy, SciPy, and Matplotlib. It provides simple and efficient tools for common data analysis tasks such as classification, regression, clustering, dimensionality reduction, and pre-processing.

4.2.6 TensorFlow:

TensorFlow is a scalable, cross-platform programming interface for implementing and running machine learning algorithms.

It was developed by researchers and engineers on the Google Brain team. Initially created for internal use at Google, TensorFlow was released as open-source in November 2015.

To enhance machine learning model performance, TensorFlow enables execution on both CPUs and GPUs. However, its full performance potential is best realized when using GPUs.TensorFlow currently supports frontend APIs in several programming languages, with the Python API being the most complete and widely used by machine learning and deep learning practitioners. Additionally, TensorFlow has an official C++ API.

TensorFlow has a vast ecosystem of related components, including libraries such as TensorBoard and deployment/production APIs.

TensorFlow 1.x had a complex Python class system for model construction. Due to Keras's immense popularity, TensorFlow 2.0 adopted Keras as its official API.

4.2.7 Keras:

Keras is a high-level deep learning API that makes it very easy to train and deploy neural networks. It can run on top of TensorFlow, Theano, or Microsoft Cognitive Toolkit (formerly CNTK). TensorFlow comes with its own implementation of the API, called tf.keras, which supports some advanced TensorFlow features (e.g., efficient data loading).

Although Keras remains a standalone library, it can now be officially imported via TensorFlow, so separate installation is not necessary.

The Keras API is user-friendly and allows models to be built simply by stacking layers one after another.

4.3 Data collecting and pre-processing

To understand the behaviour and spread of each feature, descriptive statistical analysis and visualization were performed. Box plots were generated for all normalized input variables to identify the presence of outliers and skewness. The majority of variables exhibited distributions within a normal range, although some (e.g., water content or silica fume) showed slight skewness or variance, indicating the need for normalization to ensure balanced model training.

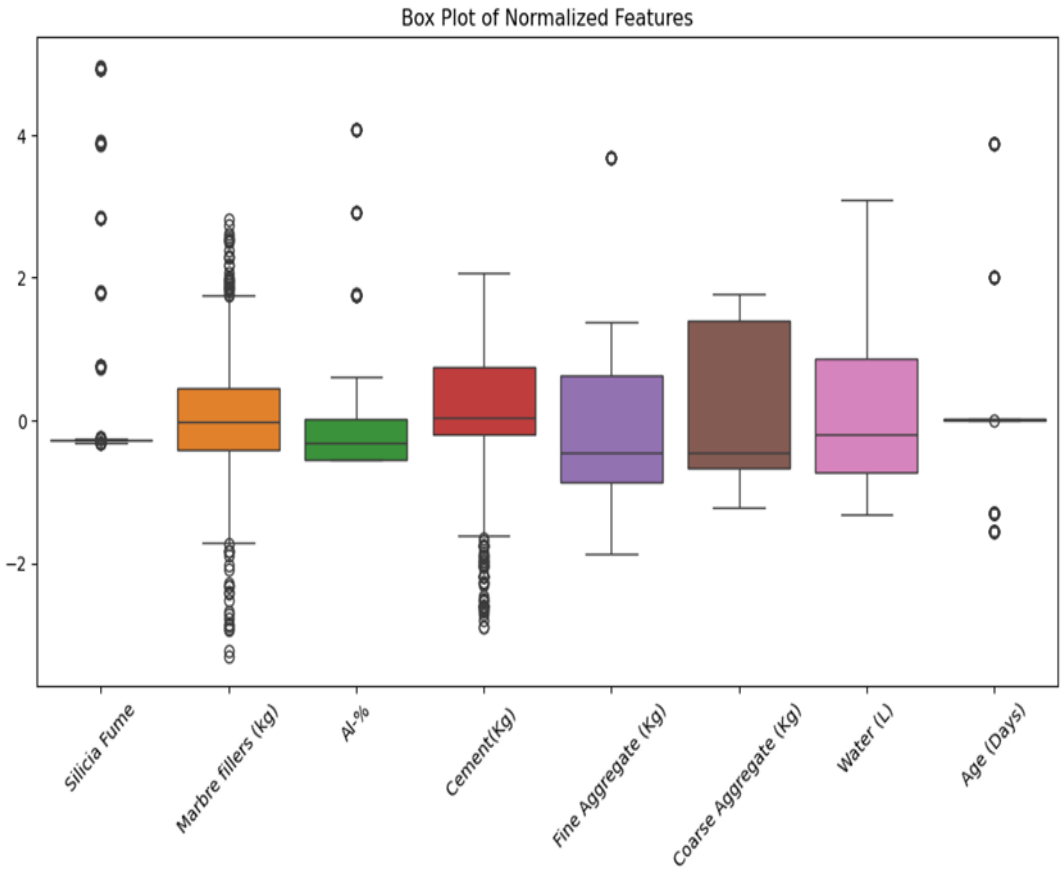


Figure4. 1: Box Plot of Normalized Features.

The plot reveals some outliers in variables like water content or silica fume, indicating the need for further data preprocessing to improve model performance.

The overall distribution of most variables appears balanced, suggesting successful normalization and suitability for modeling.

4.3.1 Data Distribution Analysis:

There are several ways to describe or classify correlation, but three of the most significant include:

4.3.2 Correlation Between Variables:

A Pearson correlation matrix was calculated to assess linear relationships between features and the target variable. This analysis helps identify potential multicollinearity between inputs and determines which features most significantly influence the target output (strength). Strong correlations were observed between cement content and compressive strength, while curing age also showed moderate positive correlation with strength values.

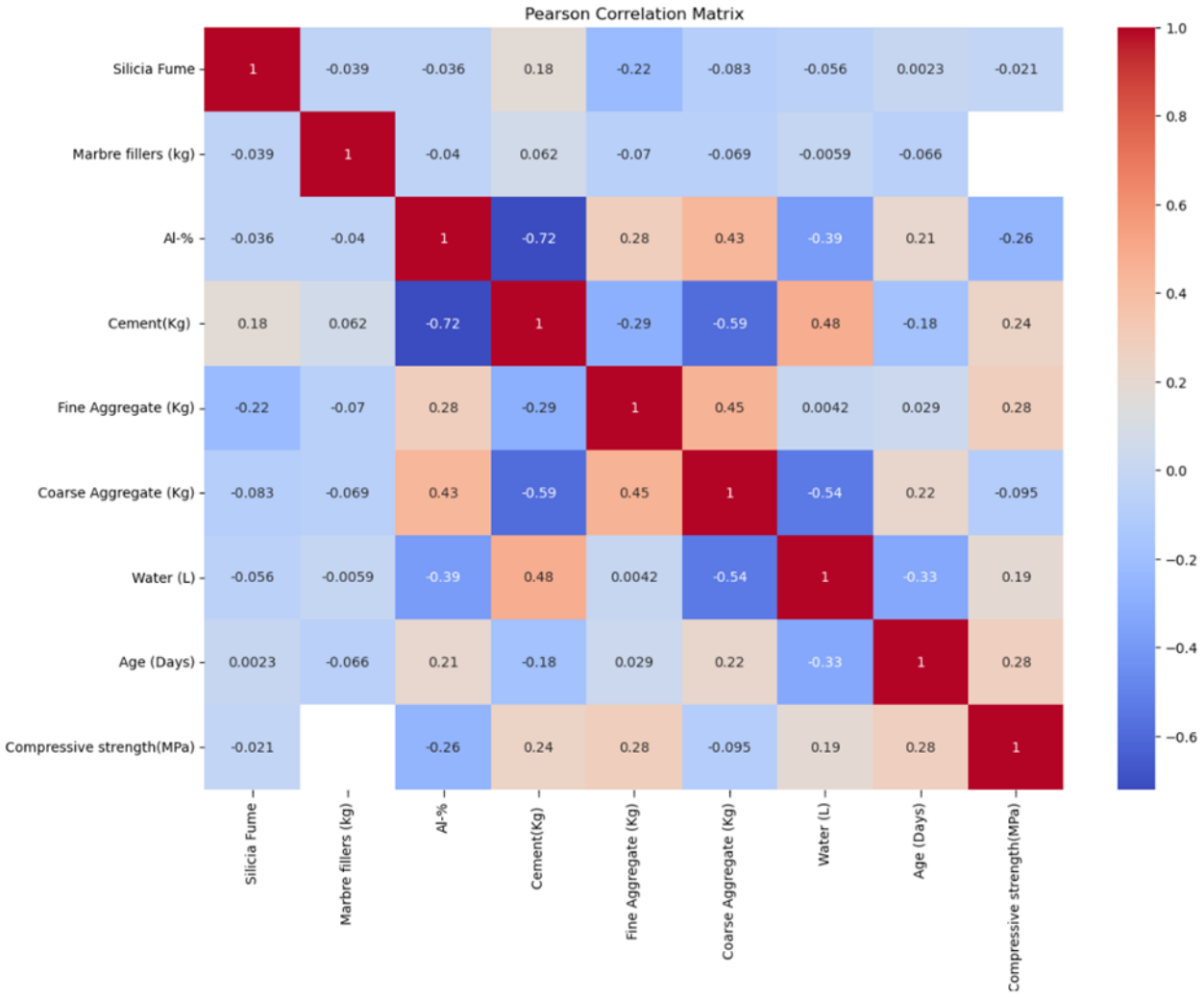


Figure4. 2: Pearson Correlation Matrix for Compressive Strength.

A strong positive correlation exists between cement content and compressive strength, which is expected given cement's primary role in concrete.

Curing age also shows a moderate positive correlation, reflecting strength gain over time.

Some variables exhibit weak correlations, suggesting minimal impact on compressive strength.

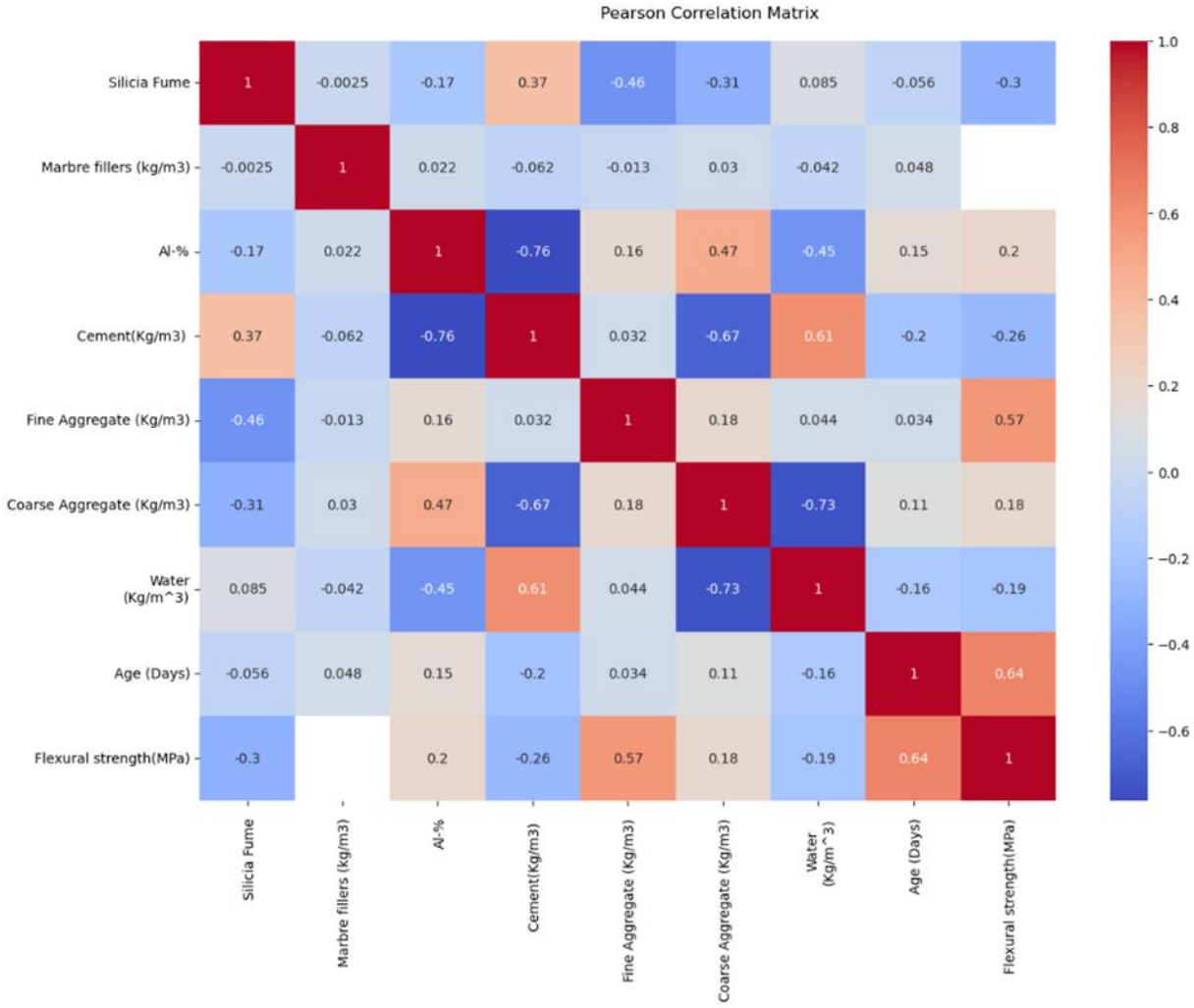


Figure4. 3: Pearson Correlation Matrix for Flexural Strength.

Correlations are less pronounced compared to compressive strength, possibly due to the more complex nature of flexural behavior.

Fiber content shows a moderate correlation, highlighting its importance in enhancing flexural strength.

4.3.3 Variable Normalization:

Since the dataset contains features with different units and scales (e.g., cement in kg/m³, water in liters, age in days), Min-Max normalization was applied to scale all input variables between 0 and 1. This normalization ensures that all features contribute equally to the model's training and prevents bias toward higher-magnitude variables.

4.3.4 Hyperparameter Tuning:

To optimize model performance, hyperparameter tuning was conducted using grid search and cross-validation. Key parameters such as learning rate, number of estimators (in ensemble models), maximum depth (for trees), and batch size (for neural networks) were varied systematically. The best combination was selected based on validation performance, primarily using metrics such as Mean Squared Error (MSE) and R^2 score.

4.4 Data scaling and splitting

Before feeding the dataset into machine learning models, appropriate scaling and splitting procedures were applied to ensure fair and unbiased training. These steps are critical to enhance model generalizability and avoid overfitting

4.4.1 Scaling techniques:

The dataset includes numerical features with varying magnitudes and units (e.g., kilograms, litters, days). To bring all input variables to a comparable scale, the Min-Max scaling method was employed. This technique transforms each feature to a [0, 1] range based on its minimum and maximum values. Such normalization is especially beneficial for algorithms sensitive to the scale of input data, including neural networks and distance-based models like K-Nearest Neighbours.

4.4.2 Training-validation-testing division:

To evaluate the model's performance objectively, the dataset was divided into three distinct subsets:

•Training set (70%): Used to train the machine learning model.

•Validation set (15%): Used to fine-tune model parameters and prevent overfitting.

•Testing set (15%): Reserved for final performance evaluation after the model is fully trained.

This division ensures that the model is evaluated on unseen data, giving a more accurate estimate of real-world performance.

4.4.3 Cross-validation methodology:

In addition to the standard split, k-fold cross-validation was employed to enhance robustness during training and hyperparameter tuning. In this approach, the training set is further divided into k subsets (folds). The model is trained k times, each time using a different fold as the validation set and the remaining k-1 folds as the training set. The final performance metric is obtained by averaging the results across all folds.

In this study, 5-fold cross-validation was chosen to balance computational efficiency and evaluation stability. This method helps reduce variance in performance estimates and makes the model selection process more reliable.

4.5 Model Performance Metrics

To evaluate and compare the effectiveness of the predictive models, several standard regression performance metrics were used. These metrics quantify the accuracy and reliability of the model's predictions compared to actual target values. The following metrics were adopted in this study:

4.5.1 MAE (Mean Absolute Error):

Mean Absolute Error measures the average magnitude of the errors in a set of predictions, without considering their direction. It is calculated as the average of the absolute differences between the predicted and actual values. MAE is intuitive and easy to interpret, offering a straightforward indication of how close the predictions are to the real outputs. Lower MAE values signify better model performance.

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

4.5.2 MSE (Mean Squared Error):

Mean Squared Error calculates the average of the squared differences between predicted and actual values. Unlike MAE, MSE penalizes larger errors more severely due to the squaring process, making it sensitive to outliers. It is particularly useful in optimization because it is mathematically differentiable, which facilitates gradient-based learning methods.

$$MSE = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

4.5.3 R^2 (R-squared):

R-squared, or the coefficient of determination, indicates how well the independent variables explain the variability of the dependent variable. It ranges from 0 to 1, where a value closer to 1 implies that the model explains a large proportion of the variance in the outcome variable. An R^2 of 0.80, for example, suggests that 80% of the variation in the target variable can be explained by the input features used in the model.

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}$$

4.6 Model Creation and Results

The analysis of the predictive accuracy across different series reveals notable insights into the model's performance. In Series 0, the predicted flexure strength exhibits an accuracy of:

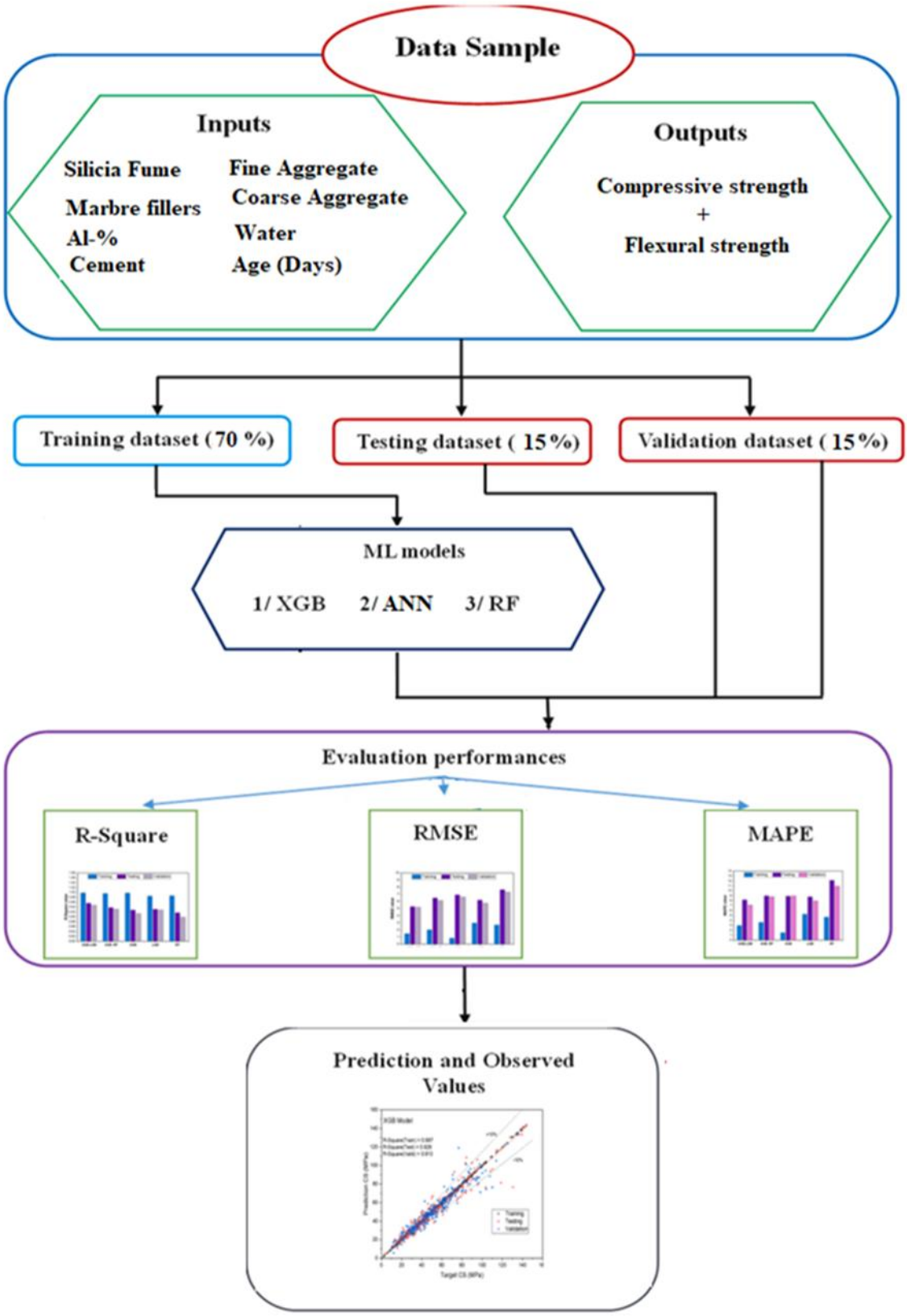


Figure4.4: Modeling Process and Performance Evaluation for Predicting Mechanical Strengths of Materials.

4.6.1 Compressive strength:

The intercept is the point where the regression line intersects the y-axis. It provides a measure of the mean of the dependent variable when the slopes are zero. If the slopes are not zero, the intercept represents the mean of the dependent variable minus the product of the slope and the mean of the independent variable.

- ***Model 01 Xgboost***

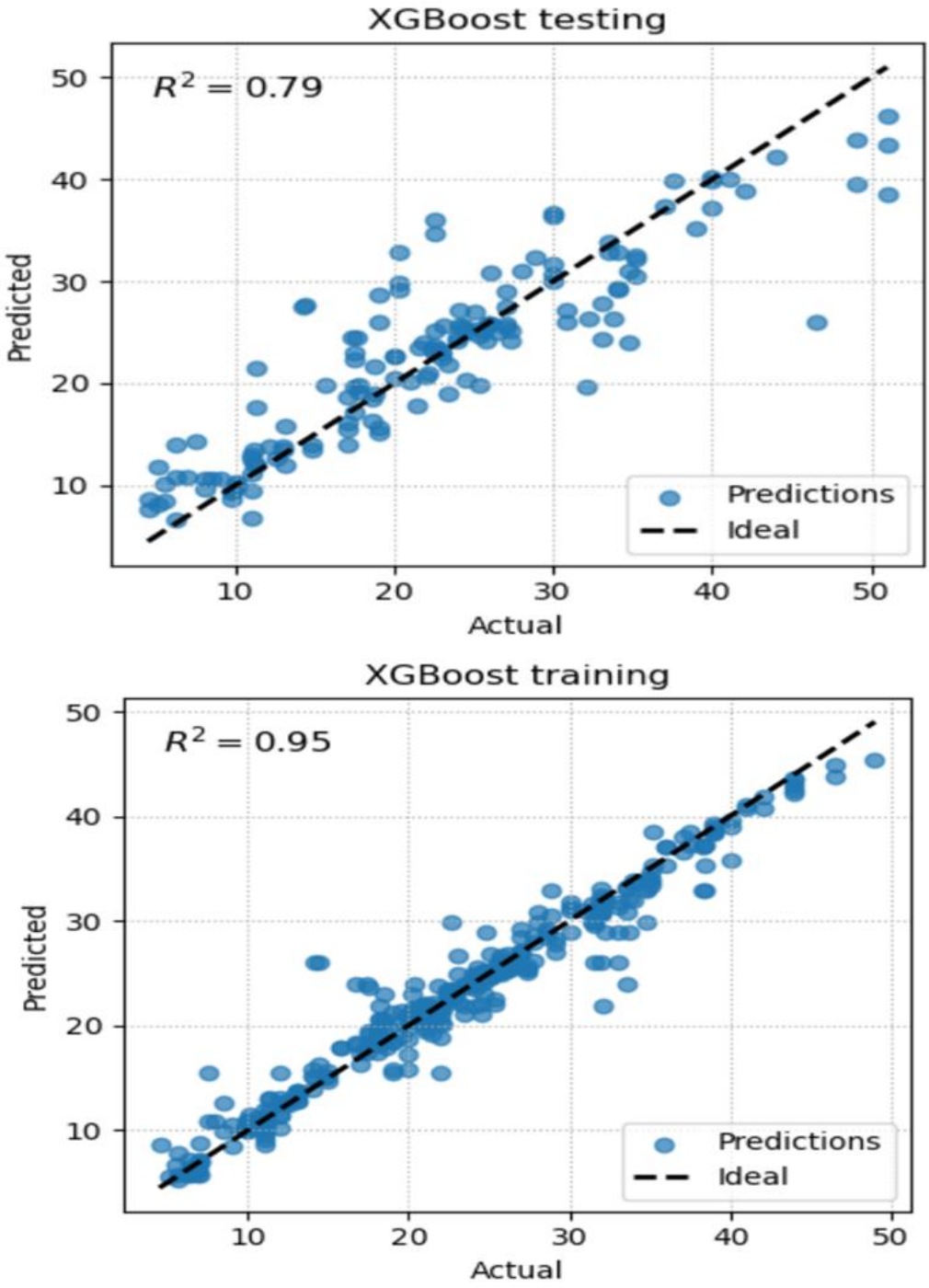


Figure4. 5: The performance of XGBoost in predicting compressive strength.

The XGBoost model shows a strong correlation between actual and predicted compressive strength values, with a close alignment to the ideal prediction line. Minor discrepancies and a few outliers suggest some potential for improvement, likely through enhanced feature engineering or tuning.

- ***Model 02 Random forest***

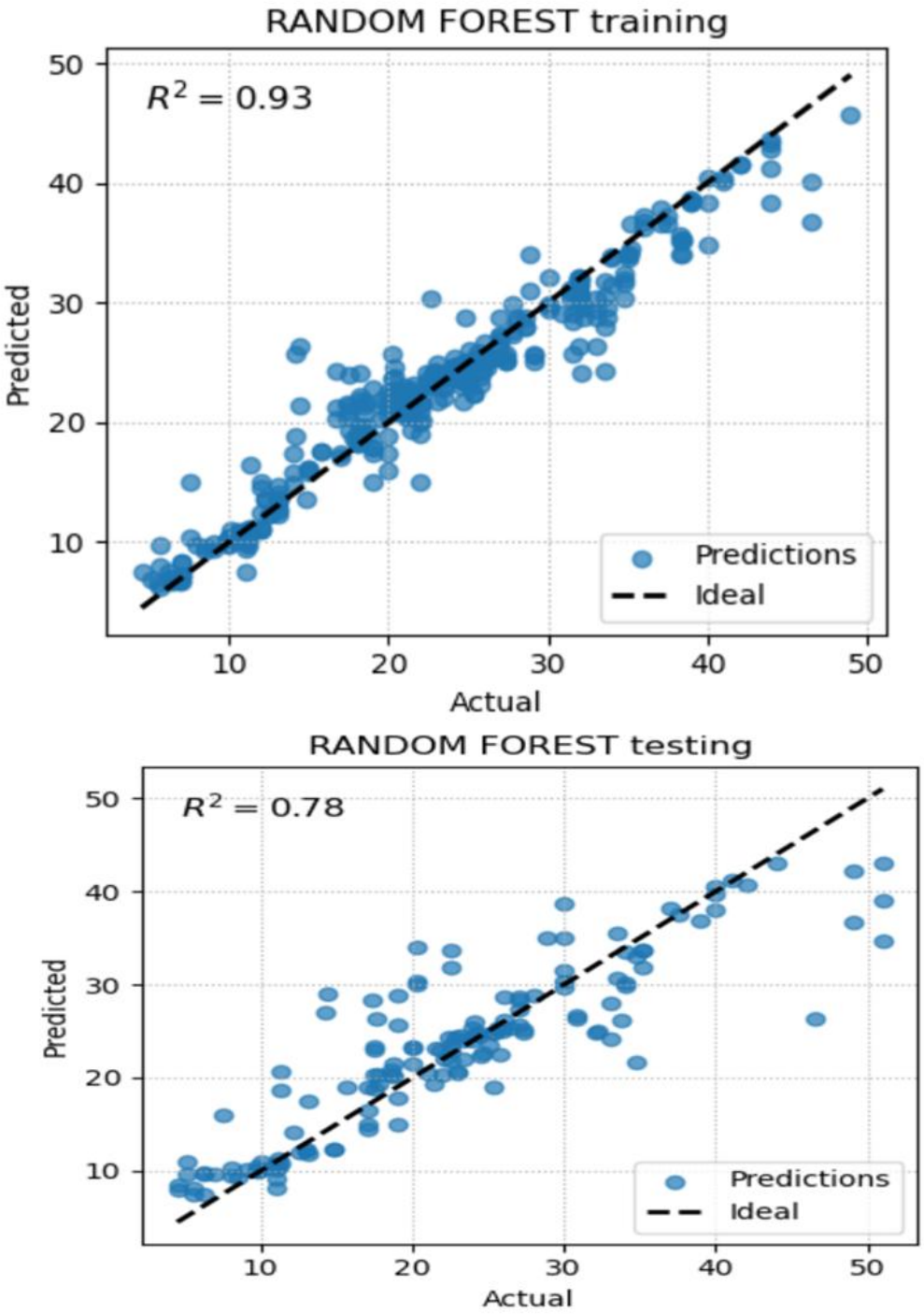


Figure 4. 6: The performance of Random Forest in predicting compressive strength.

Random Forest performs comparably to XGBoost but shows slightly greater prediction error, especially for higher compressive strength values. This indicates a possible limitation in extrapolating extreme values.

- ***Model 03 Neural network***

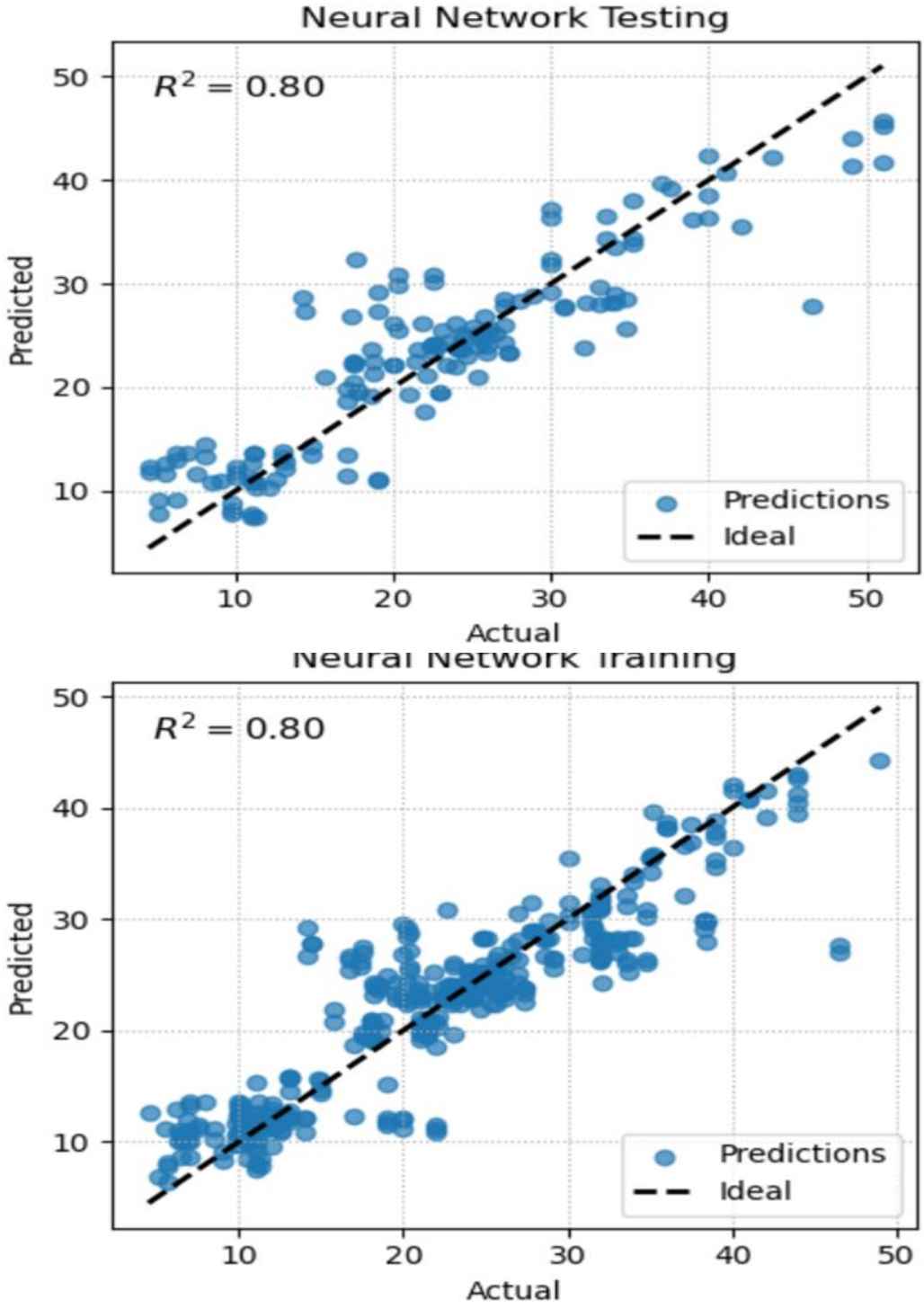


Figure4. 7: The Neural Network's performance in predicting compressive strength.

The neural network demonstrates solid prediction accuracy, particularly in the mid-range compressive strengths. Some deviations suggest possible overfitting or insufficient generalization at the extremes.

4.6.2 Flexural strength:

The slope represents the change in the dependent variable as the independent variable changes. A zero slope indicates that the independent variable has no influence on the dependent variable. In a linear model, the slope is not the same as elasticity.

- ***Model 01 Xgboost***

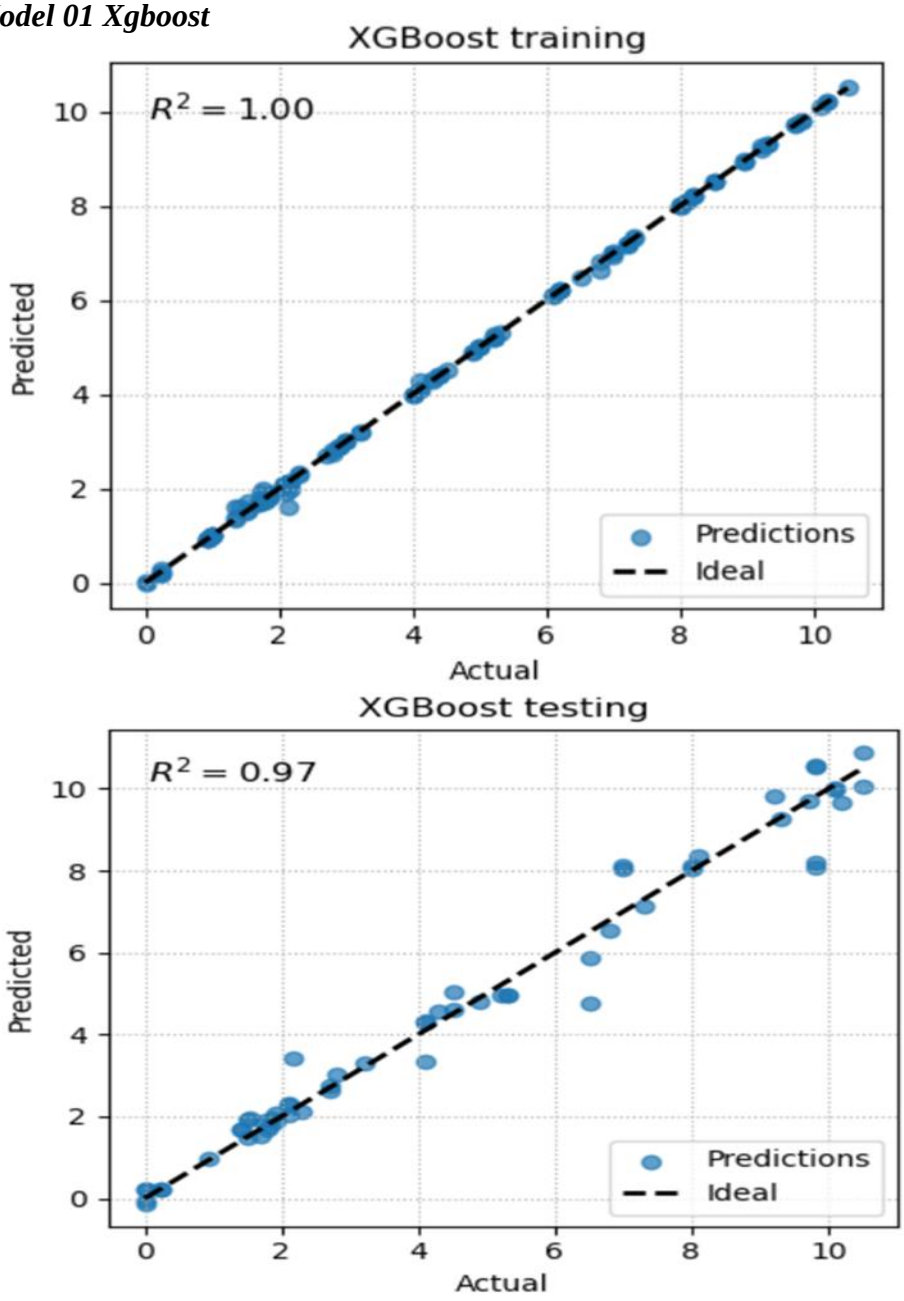


Figure4. 8:XGBoost's performance in predicting flexural strength.

The model achieves near-perfect accuracy, with actual and predicted values closely aligned. Performance is superior to compressive strength predictions, possibly due to more consistent data patterns.

- ***Model 02 Random forest***

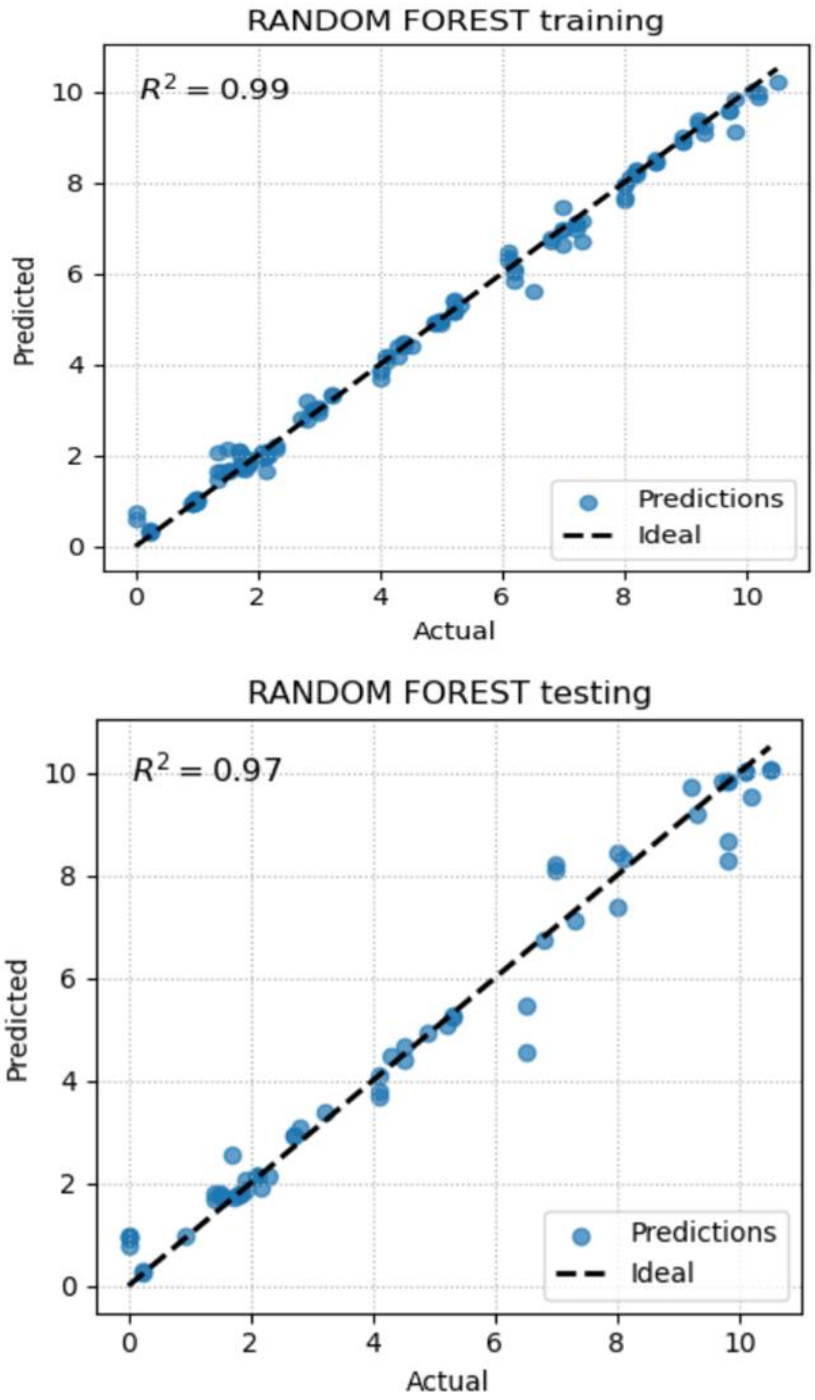


Figure4. 9: Random Forest's performance for flexural strength.

Performance is excellent but marginally inferior to XGBoost, particularly for extreme values.

The model maintains consistency across predictions.

- ***Model 03 Neural network***

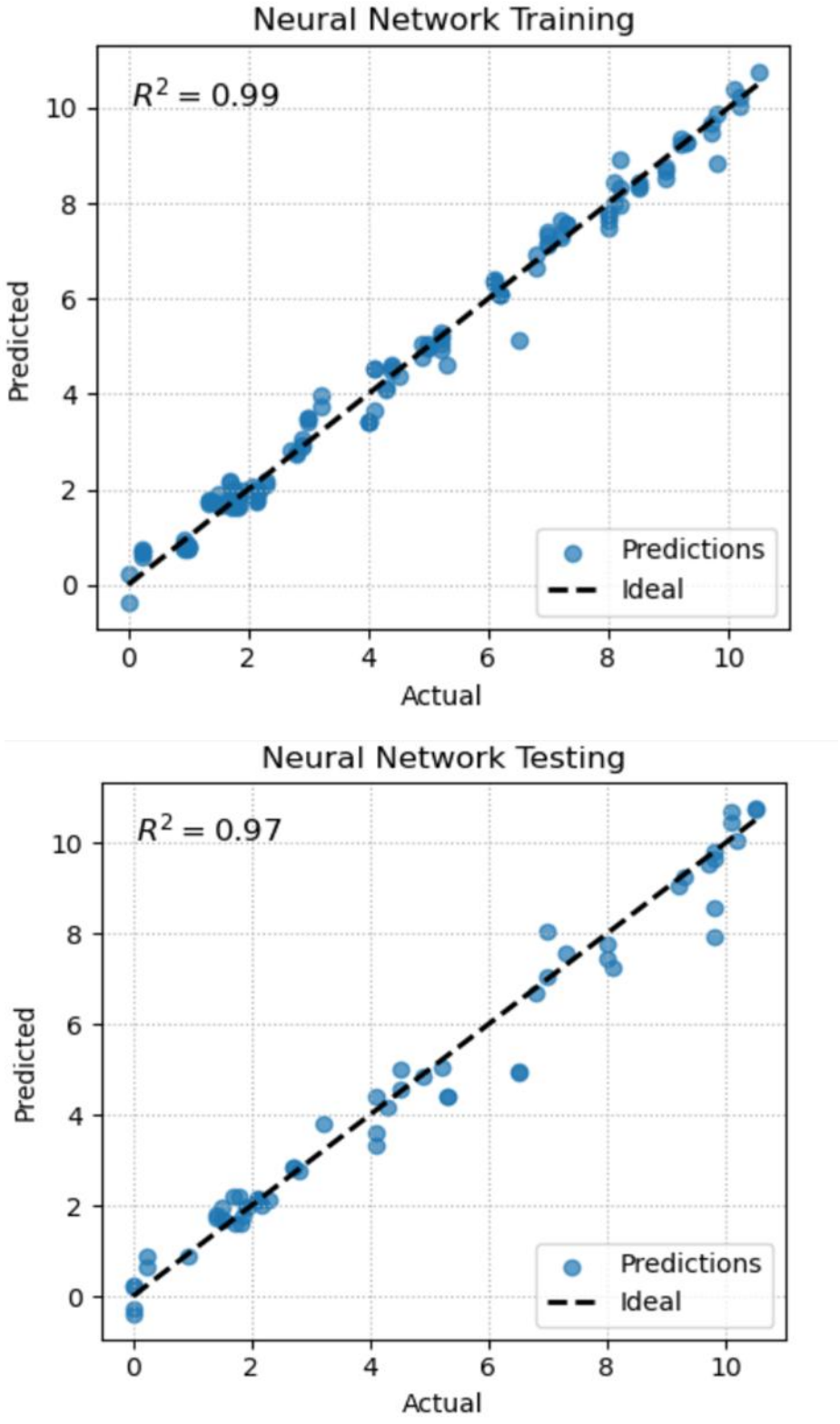


Figure4. 10: The Neural Network's performance for flexural strength.

The model performs well but lags slightly behind XGBoost. Minor deviations suggest opportunities for further optimization.

4.7 Model comparison

To evaluate the predictive performance of different machine learning models, three algorithms were tested: XGBoost**,** Random Forest**,** and Artificial Neural Network (ANN)**.** Their effectiveness was assessed on two distinct prediction tasks compressive strength and flexural strength using three key performance metrics: Mean Absolute Error (MAE)**,** Mean Squared Error (MSE)**,** and the coefficient of determination (R^2)**.**

4.7.1 Compressive Strength Prediction

The comparison results for compressive strength prediction are presented in the table below:

Table4. 1: Compressive Strength Prediction Results.

Model	**MAE**	**MSE**	**R^2**
XGBoost	3.3635	23.8020	0.7948
Random Forest	3.3400	25.2319	0.7825
ANN	3.5845	23.2365	0.7997

XGBoost achieved the best trade-off between low error and high R^2, although the ANN slightly outperformed it in MSE. Random Forest had slightly lower MAE but lagged in R^2 and MSE. Thus, **XGBoost** was identified as the most balanced model for compressive strength prediction.

4.7.2 Flexural Strength Prediction

The table below presents the model performance for flexural strength prediction:

Table4. 2: Flexural Strength Prediction Results.

Model	MAE	MSE	R^2
XGBoost	0.3514	0.2944	0.9748
Random Forest	0.3617	0.3074	0.9737
ANN	0.3803	0.3090	0.9736

Here, all models performed exceptionally well, with R^2 values exceeding 0.973. However, **XGBoost** again delivered the most accurate predictions across all metrics.

4.7.3 Summary of Model Selection

Table4. 3: Summary of model selection.

Prediction Type	Best Model	Reason
Compressive Strength	XGBoost	Best balance of MAE, MSE, and R^2
Flexural Strength	XGBoost	Highest accuracy and consistency

Based on this comparison, **XGBoost** stands out as the most robust and consistent model, making it the recommended choice for both prediction tasks in this study.

4.8 Conclusion

In this chapter, and based on the comparative analysis of XGBoost, Random Forest, and Artificial Neural Network (ANN) for predicting the compressive and flexural strength of concrete with waste aluminium fibres, XGBoost emerged as the most robust and accurate model across both tasks. For compressive strength, XGBoost provided the best balance of low error and high correlation (MAE = 3.36, MSE = 23.80, R^2 = 0.7948), outperforming Random Forest in consistency and ANN in average error. Although ANN showed slightly higher R^2 and lower MSE, its higher MAE suggests less reliability in individual predictions, possibly due to overfitting or insufficient regularization. Random Forest performed

moderately but struggled with extreme values due to its ensemble averaging nature. In flexural strength prediction, all models performed exceptionally well ($R^2 > 0.973$), yet XGBoost again demonstrated the highest accuracy (MAE = 0.3514, MSE = 0.2944, R^2 = 0.9748), showing near-perfect alignment between predicted and actual values. These results highlight XGBoost's capacity to model complex nonlinear relationships effectively while maintaining generalization and robustness, making it the recommended model for strength prediction of concrete materials in engineering applications.

CHAPTER 05: Conclusion and recommendations for future study

5.1Conclusion

Machine learning has become a fundamental component of modern technological progress, offering adaptive, data-driven solutions that span diverse fields. Its core learning paradigms supervised, unsupervised, semi-supervised, and reinforcement learning enable flexible modeling approaches, with algorithms like linear regression, decision trees, and support vector machines forming the backbone of intelligent systems. The importance of robust validation methods, such as cross-validation, cannot be overstated, as they ensure generalization and guard against overfitting. With the rapid expansion of data and computing capabilities, a strong grasp of machine learning fundamentals is crucial for future innovation and application.

The incorporation of waste aluminum in concrete demonstrates a promising pathway toward sustainable construction:

- Experimental results confirmed that substituting cement with 5–15% aluminum waste and reinforcing with 1–4% aluminum fibers improves compressive, flexural, and tensile strength under standard and harsh conditions.
- The most notable strength gains were observed in mid-range AW dosages (around 10% cement replacement and 2% fiber), where both mechanical performance and workability were reasonably balanced.
- The use of AW also offers thermal insulation benefits and suitability for hot-weather concreting, due to its reflective properties and microstructure.
- Despite a reduction in workability and early strength at higher dosages (above 15% AW), these drawbacks were partially mitigated by optimizing water-to-binder ratios and using superplasticizers.
- The research supports integrating aluminum waste with other supplementary cementitious materials (SCMs) like fly ash or blast furnace slag, creating hybrid sustainable concrete blends with enhanced durability and performance.

In terms of predictive modeling, several findings validate the strength of machine learning tools in this context:

- Among the tested models—XGBoost, Random Forest, and Artificial Neural Network (ANN) XGBoost achieved the highest accuracy and robustness for both compressive and flexural strength predictions, making it the most reliable model in this study.

- For compressive strength, XGBoost offered the best overall trade-off (MAE = 3.36, MSE = 23.80, R^2 = 0.7948), followed closely by ANN in terms of R^2 but with a higher prediction error, suggesting more variance in its individual predictions.
- For flexural strength, XGBoost performed exceptionally (MAE = 0.3514, MSE = 0.2944, R^2 = 0.9748), showing almost perfect alignment with actual test values and outperforming the other models in all metrics.
- Random Forest showed good consistency across both prediction tasks, but its tendency to average predictions across trees led to underestimation or overestimation of extreme values, reducing its effectiveness in datasets with wider strength variability.
- The ANN model showed promise, particularly in mid-range predictions, but was prone to slight overfitting, indicating a need for more sophisticated architecture tuning or data augmentation.
- The study highlights the value of data-driven approaches in concrete technology, where empirical testing can be augmented by machine learning to reduce cost, time, and experimental errors.
- Feature importance analysis within XGBoost further identified key predictors such as cement content, AW dosage, curing time, and water-cement ratio, offering insights into material design priorities for high-performance AW-concrete.

These findings reinforce that machine learning is not only capable of modeling concrete behavior but can also guide mix optimization, material selection, and performance forecasting in modern civil engineering.

5.2 Recommendations for future study

To further enhance the understanding and application of Machine Learning for predicting compressive and flexural strength of waste aluminium fibre-reinforced concrete, the following detailed recommendations for future research are proposed:

- ✓ **Increase Datasets:** Add additional variables (e.g., fibre geometry, curing conditions) and larger datasets to enhance model generalizability.
- ✓ **Construct Hybrid Models:** Integrate ML with physics-based models (e.g., finite element analysis) to increase interpretability and accuracy.

- ✓ **Examine Long-Term Durability:** Test WAFC's corrosion, freeze-thaw cycle, and other environmental resistance.
- ✓ **Explore Advanced ML Methods:** Utilize deep learning (e.g., CNNs for crack analysis) or transfer learning to leverage pre-trained models.
- ✓ **Perform Economic Analysis:** Evaluate cost-benefit trade-off and scalability for industrial application.
- ✓ **Optimize Multi-Objective Designs**: Use ML to optimize strength, sustainability, and cost simultaneously in concrete mixes.
- ✓ **Triangulate in Real-World Settings:** Test ML predictions in real construction contexts to assess practical usefulness.
- ✓ **Close Data Gaps:** Include underrepresented mix designs (e.g., high-volume fibre replacements) to increase predictability.
- ✓ **Interact with Industry:** Collaborate with construction companies to match research with actual material and regulatory limitations.

REFERENCES

[1] HINTON, Geoffrey E.; OSINDERO, Simon; TEH, Y-W Teh. “A fast learning algorithm for deep belief nets”. Neural computation, 2006, 18.7: 1527-1554.

[2] GÉRON, Aurélien. Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems. O’Reilly Media, Inc.", 2022.

[3] CHEN, Li-Pang. Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar: Foundations of machine learning: The MIT Press, Cambridge, MA, 2018, 504 pp., CDN $96.53 (hardback), ISBN 9780262039406. 2019.

[4] HASTIE, Trevor. Tibshirani R. Friedman J.: The elements of statistical learning. Friedman J.: The elements of statistical learning, 2001. [5] G. James, D. Witten, T. Hastie, and R. Tibshirani, An introduction to statistical learning. Springer, 2013, vol. 112.

[5]JAMES, Gareth, et al. An introduction to statistical learning. New York: springer, 2013.

[6]BELLAHMER, Hacene. Implémentation et évaluation d’un modèle d’apprentissage automatique pour l’estimation de la valeur marchande de propriétés immobilières. 2020. PhD Thesis. Université Mouloud Mammeri.

[7] VASILEV, Ivan, et al. Python Deep Learning: Exploring deep learning techniques and neural network architectures with Pytorch, Keras, and TensorFlow. Packt Publishing Ltd, 2019.

[8] ELINWA, Augustine U.; MBADIKE, Elvis. The use of aluminum waste for concrete production. Journal of Asian Architecture and Building Engineering, 2011, 10.1: 217-220..

[9] ARIMANWA, J. I., et al. Prediction of the compressive strength of aluminum waste–cement concrete using Scheffe’s theory. Journal of materials in civil engineering, 2012, 24.2: 177-183.

[10] B. A. Omran, Q. Chen, and R. Jin, "Comparison of Data Mining Techniques for Predicting Compressive Strength of Environmentally Friendly Concrete," Journal of Computing in Civil Engineering, vol. 30, no. 6, 2016, Art, no. 04016029.

[11]J. Ilya and C. Cheow Chea, "Mechanical behaviour of fibre reinforced concrete using soft-drink can," IOP Conf. Ser.: Mater. Sci. Eng., vol. 271, no. 1, p. 012079, 2017.

[12] MUWASHEE, Rawa Shakir; AL-JAMEEL, Hamid Athab; JABAL, Qusay Abdulhameed. Investigating the behavior of concrete and mortar reinforced with Aluminum waste strips. Int. J. Eng. Technol, 2018, 7.4.37: 211-213.

[13] PANDITHARADHYA, B. J., et al. Mechanical properties of pavement quality concrete with secondary aluminium dross as partial replacement for ordinary Portland cement. In: IOP conference series: materials science and engineering. IOP Publishing, 2018. p. 032011.

[14]OFUYATAN, Olatokunbo M., et al. Effect of waste aluminium shavings on the bond characteristics of laterized concrete. Advances in materials Research, 2019, 8.1: 25-36..

[15] SABAPATHY, Y. K., et al. Experimental study on strength properties of aluminium fibre reinforced concrete. Journal of King Saud University-Engineering Sciences, 2021, 33.1: 23-29.

[16] BOUDJEMA, Mohamed; KENNOUDI, Ahmed; GUENDOUZ, Mohamed. Influence de déchets d'aluminium sur les performances de Béton de sable fluide. 2020.

[17] Nadia Tebbal, Zine El Abidine Rahmouni, Valorization of aluminum waste on the Mechanical Performance of mortar subjected to cycles of freeze-thaw, Procedia Computer Science 158 (2019) 1114–1121.

[18] VAN, Lam Tang, et al. Effect of aluminium powder on light-weight aerated concrete properties. In: E3S Web of Conferences. EDP Sciences, 2019. p. 02005.

[19] CHANNA, Imran Ali; SAAND, Abdullah. Mechanical behavior of concrete reinforced with waste aluminium strips. Civil Engineering Journal, 2021, 7.7: 1169-1182.

[20] ASLAN, Çağatay, et al. Effect of industrial waste metal chips on flexural behavior of reinforced concrete beams. Revista de la construcción, 2023, 22.2: 368-381.

[21] AGOR, Chima Dike; MBADIKE, Elvis Michael; ALANEME, George Uwadiegwu. Evaluation of sisal fiber and aluminum waste concrete blend for sustainable construction using adaptive neuro-fuzzy inference system. *Scientific Reports*, 2023, 13.1: 2814..

[22] A. Suryadi and T. Rochman, "The effect of aluminum waste ash on the concrete compressive strength," Urbanism. Arhitectură. Construcţii, vol. 14, no. 1, pp. 83–94, 2023.

[23] SELMANI, Khalil; ENCADRE PAR BOULAARES, Said. Utilisation de la méthodologie des plans d'expérience pour la formulation de béton à base des déchets. 2023. PhD Thesis.

[24]Gohil, K. and Patel, N. Study on Properties of Concrete Using E-Waste Copper, Aluminium and Steel Fibers. *EasyChair Preprint*, no. 11996, Feb. 2024.

[25] HAFIANE, Ahmed Anisse; ENCADRÉ PAR BOURSAS, Farid. Compression-Flexure Relationship in Fibre Concrete Based on Wasted Aluminium. 2024. PhD Thesis. UNIVERSITE DE ECHAHID CHEIKH LARBI TEBESSI.