ISBN: 978-1-7636843-3-1

Machine Learning with WEKA for Predicting Concrete Compressive Strength: An Educational Approach for Civil Engineering Students

M Arifuzzaman[1], M Muhitur Rahman[1], M K Islam[1], M Assali[2], U Gazder[3], Wadea Sindi[4] and Manal Aburizaiza[5]

[1]Faculty, King Faisal University, Al Ahsa, Kingdom of Saudi Arabia
[2]Engineer, Spacemaker KSA LLC, Al Khobar, Kingdom of Saudi Arabia
[3]Faculty, University of Bahrain, Sakhir, Bahrain
[4]Director, Jeddah Development Authority, Jedda, Kingdom of Saudi Arabia
[5]Faculty, Department of Computer Science and Engineering, Yanbu Industrial College, Yanbu Industrial City, Kingdom of Saudi Arabia

Corresponding author's E-mail: marifuzzaman@kfu.edu.sa

Abstract

The research paper explores the application of machine learning algorithms using the WEKA platform to predict the compressive strength of concrete, a critical aspect of civil engineering. This study focuses on educating undergraduate students with no prior programming experience, offering a practical introduction to AI tools. Utilizing established datasets and advanced machine learning techniques, the research evaluates the performance of various ensemble-based models, including Random Committee, Bagging, and Rotation Forest, alongside instance-based models like KStar and IBk. The findings reveal that ensemble methods outperform instance-based techniques, providing higher prediction accuracy, lower error rates, and improved reliability. By integrating machine learning into civil engineering education, the study aims to enhance students' understanding of key concepts like compressive strength while preparing them for real-world engineering challenges. This interdisciplinary approach also underscores the potential of AI in optimizing construction materials and advancing sustainable infrastructure.

Keywords: Compressive Strength, Machine Learning, WEKA.

1. INTRODUCTION

Concrete compressive strength is a critical aspect of civil engineering education at universities. It refers to the ability of concrete to withstand axial loads without failure, measured in megapascals (MPa). Understanding this property is essential for students, as it influences design decisions in construction projects. Engineering curricula often include practical labs where students test concrete samples, reinforcing theoretical knowledge with hands-on experience. This integration of concrete strength principles prepares future engineers to ensure safety and durability in their designs, ultimately contributing to the advancement of infrastructure and construction practices.

Concrete compressive strength is a fundamental property that determines the load-bearing capacity of concrete structures. It is essential for ensuring the safety, durability, and longevity of buildings, bridges, and other infrastructure. Understanding this property is crucial for civil engineers, as it directly influences design decisions and material selection as suggested by Osman (2021).

In civil engineering education, concrete compressive strength serves as a cornerstone for various subjects, including structural analysis, materials science, and construction management. By integrating this concept into the curriculum, universities prepare students to tackle real-world engineering challenges effectively. Concrete compressive strength is defined as the maximum axial load that a concrete specimen can withstand before failure, typically measured in megapascals (MPa) or pounds per square inch (psi). Standard tests, such as the ASTM C39, are employed to evaluate this property using cylindrical specimens. Several factors influence concrete compressive strength, including the water-cement ratio, aggregate type, curing conditions, and the presence of admixtures. Understanding these variables is essential for students to optimize concrete mixtures for specific applications. Fig 1 demonstrates laboratory testing for concrete compressive strength.

Fig 1: Concrete Cylinder Compressive Strength Testing Demonstration

To effectively teach concrete compressive strength, universities should integrate theoretical knowledge with practical applications. Courses should cover the principles of concrete technology, structural design, and the implications of compressive strength on safety and performance. Hands-on laboratory experiences are vital for reinforcing theoretical concepts. Students should engage in experiments that involve mixing, casting, and testing concrete specimens to observe the effects of different variables on compressive strength firsthand.

Incorporating case studies of real-world projects allows students to analyze how compressive strength impacts design decisions. Evaluating successful and failed structures provides valuable lessons on the importance of accurate strength assessments. Emerging technologies, such as virtual simulations and interactive software, can enhance the learning experience. These tools allow students to visualize the behavior of concrete under various loads and conditions, fostering a deeper understanding of compressive strength.

Concrete compressive strength is a critical aspect of civil engineering education, influencing both theoretical knowledge and practical skills. By understanding its fundamentals, factors, and applications, students are better equipped to address engineering challenges.

High-performance concrete (HPC) represents a contemporary advancement in construction, necessitating the use of supplementary materials such as fly ash and blast furnace slag, alongside superplasticizers. The seminal water-to-cement (w/c) ratio theory proposed by Abrams in 1918 remains foundational, positing an inverse correlation between the w/c ratio and concrete strength. Abrams' principle asserts that concretes with identical w/c ratios exhibit equivalent strength, irrespective of their specific compositions as suggested by Rao (2001). However, recent experimental analyses challenge this notion, revealing that the volume of the paste also influences the strength of cement, even when other variables are controlled. This revelation adds a layer of complexity to our comprehension of HPC and the factors governing concrete strength. Figure 2 illustrates the practical application of high-strength concrete in modern construction, reflecting its role in enhancing structural integrity, efficiency, and sustainability in the engineering and construction industries.

Fig 2: High strength Concrete Application Example

Today, predicting concrete mixture properties such as compressive strength increasingly relies on advanced machine learning algorithms, including Artificial Neural Networks (ANN), Adaptive Neuro-Fuzzy Inference Systems (ANFIS), and Random Forest (RF) as per Armaghani (2020). These AI-driven approaches can derive concrete mixture strength and other related properties based on input data, offering more precise results compared to traditional regression methods. Research has extensively explored the application of machine learning models for predicting concrete strength as suggested by Asteris (2021), Hossain (2023) and Nguyen-Sy (2020). In one study, Panagiotis et al. employed a hybrid machine learning approach called Hybrid Ensemble Neural Softmax (HENSM), incorporating four conventional methods like ANN. Their findings indicated that the HENSM model achieved a high level of accuracy and showed potential as a robust alternative to address overfitting issues common in conventional models as per Bardhan (2024).

Khan (2012) investigated the use of ANN for predicting high-performance concrete strength and found that ANN outperformed traditional regression models in accuracy. They also noted the ease of using ANN to analyze how various concrete mix proportions affect the final properties. Similarly, DeRousseau et al. (2019) explored ANN methods for predicting concrete mixture strength and reported highly accurate compressive strength predictions.

Khademi et al. (2017) compared multiple methods, including multiple linear regression, ANN, and ANFIS, for predicting 28-day concrete strength. Their research highlighted that ANN and ANFIS provided the most accurate and reliable results for this purpose.

The primary objective of this research is to harness machine learning (ML) techniques to predict concrete compressive strength, a paramount parameter in the construction industry. This study employs both instance-based and ensemble-based ML methods to conduct a thorough analysis, thereby exceeding the scope of studies confined to a singular methodological framework.

2. DATA COLLECTION, STATISTICAL ANALYSIS AND ML APPLICATION

This work is based on the Yeh (1998) dataset which is famous for testing results and interpretation in Concrete Engineering Community.

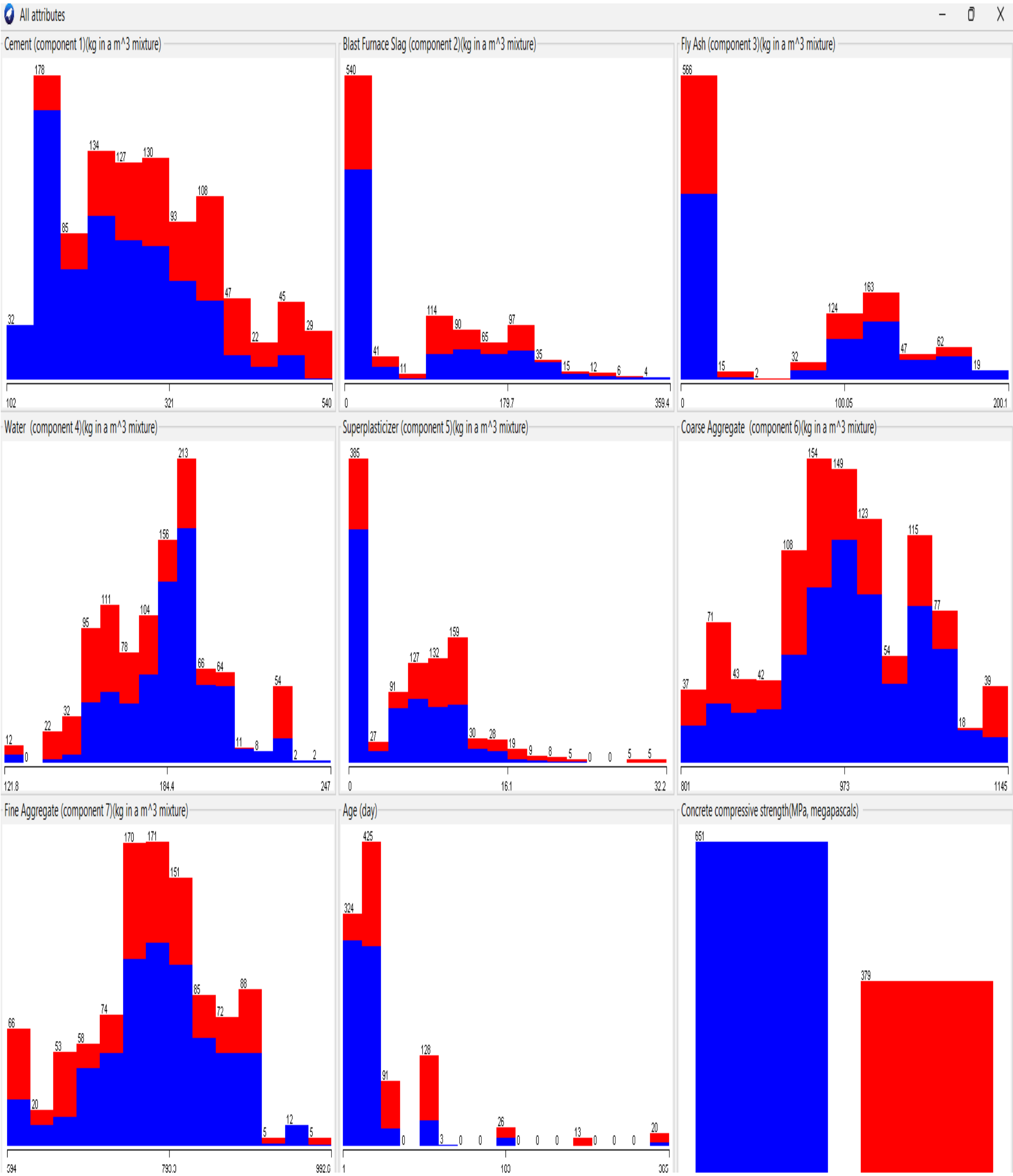

Fig 3: Statistical Diagram for the distribution of parameters in dataset

The dataset used in this study was meticulously assembled to cover a broad spectrum of concrete mixes, incorporating various combinations of cement, aggregates, water, and supplementary materials like fly ash and blast furnace slag. These mixes were crafted to mirror common formulations employed in construction, ensuring that the results are both relevant and applicable. Environmental factors, such as curing temperature and humidity, were carefully controlled and recorded during sample preparation to ensure consistency and replicability of the test conditions. This thorough approach to data collection was designed to capture the detailed behavior of concrete under various conditions, providing a solid basis for applying machine learning algorithms.

3. RESULTS AND DISCUSSIONS

After applying the six machine learning models, the results are presented in Table 2. Based on the information in Table 2, we can compare three ensemble-based machine learning models (Random Committee, Bagging, Rotation Forest) with three instance-based models (KStar, LWL, IBk). The performance of these models is evaluated across several metrics.

Table 2: Summary Results from the ML's

	Ensemble ML			**Instance Based ML**		
	Random Committee	Bagging	Rotation Forest	KStar	LWL	IBk
Correctly Classified Instances	91.77%	90.27%	91.17%	87.46%	79.83%	89.36%
Incorrectly Classified Instances	8.2 %	9.7 %	8.8%	12.5%	20.1%	10.6%
Kappa statistic	0.8172	0.7836	0.8045	0.7145	0.5463	0.7658
Mean absolute error	0.1046	0.17	0.1467	0.1492	0.3497	0.107
Root mean squared error	0.2476	0.2731	0.2558	0.3166	0.4076	0.324
Relative absolute error	22.84%	37.10%	32.03%	32.57%	76.33%	23.35%
Root relative squared error	51.73%	57.06%	53.44%	66.16%	85.16%	67.69%
F-Measure	0.917	0.902	0.911	0.871	0.795	0.893
MCC	0.819	0.785	0.806	0.722	0.549	0.766
ROC Area	0.965	0.961	0.970	0.939	0.825	0.879
Time (s)	0.05	0.08	0.41	0	0	0

The performance of various classification models was meticulously evaluated using a comprehensive suite of metrics. The ensemble methods, including Random Committee, consistently outperformed the instance-based methods, such as LWL, in terms of accuracy and overall predictive power. Random Committee achieved the highest accuracy, correctly classifying 91.77% of instances, while LWL exhibited the highest error rate, misclassifying 20.1% of instances.

Beyond accuracy, the Kappa statistic, mean absolute error, root mean squared error, relative absolute error, root relative squared error, F-Measure, Matthews Correlation Coefficient, and ROC area all demonstrated the superiority of ensemble methods. These metrics assess the agreement between predicted and actual values, the magnitude of errors, the balance between precision and recall, the ability to distinguish between classes, and the computational efficiency. Ensemble methods consistently achieved higher scores on these metrics, indicating their superior performance in classification tasks. For instance, the Kappa statistic, which measures the agreement between predicted and actual values while accounting for chance, was significantly higher for ensemble methods. This suggests that their predictions were not simply due to random guessing but were based on meaningful patterns in the data. Similarly, the mean absolute error, which measures the average magnitude of errors, was lower for ensemble methods, indicating that their predictions were more precise.

Furthermore, the F-Measure, which combines precision and recall, was also higher for ensemble methods. This suggests that they were able to achieve a better balance between correctly identifying positive instances (precision) and correctly identifying negative instances (recall). The Matthews Correlation Coefficient, which is a balanced measure that can be used even for imbalanced classes, also favored ensemble methods, indicating their reliability in classification tasks as per Sevim (2021), Yaseen (2018) and Yuan (2014).

Finally, the ROC area, which measures the model's ability to distinguish between classes, was larger for ensemble methods. This suggests that they were more effective at separating positive and negative instances in the data. However, it's important to note that instance-based methods, while less accurate, may be more computationally efficient, as evidenced by the zero computational time reported for KStar and LWL. This could make them more suitable for real-time applications where computational speed is a critical factor. From these numerical findings, it is evident that ensemble-based machine learning models outperform instance-based models across most metrics for predicting the compressive strength of concrete. This suggests that for this particular application, ensemble methods are more suitable for achieving higher prediction accuracy and reliability.

4. CONCLUSIONS

Based on the provided document excerpts, several conclusions can be drawn regarding the use of machine learning (ML) techniques for predicting the compressive strength of concrete:

1. Ensemble-based ML models generally outperform instance-based models in predicting the compressive strength of concrete. This is evidenced by the higher accuracy, lower error rates, and better performance across various metrics for ensemble methods as shown in Table 2.
2. The study highlights the importance of data preprocessing, particularly the removal of outliers and extreme values, which is crucial for improving the reliability of ML models.
3. The research contributes to the field by comparing different ML approaches, such as ANN, ANFIS, and ensemble methods, and finding that ensemble methods are more suitable for achieving higher prediction accuracy and reliability in this specific application.
4. The research is part of a broader educational context, as indicated by the references to civil engineering education and the importance of integrating theoretical knowledge with practical applications for students.
6. The document also touches on the advancements in concrete technology, such as the use of high-performance concrete (HPC) and the influence of the water-to-cement ratio and paste volume on concrete strength, reflecting a move towards more sustainable and efficient construction materials.

ACKNOWLEDGMENTS

We would like to extend our heartfelt thanks to the Weka Team at the University of Waikato, New Zealand, for their crucial support in analyzing our data. Their expertise with the Weka software suite was instrumental to the success of our research project. The team's dedication significantly improved both the quality and efficiency of our analysis, underscoring how essential external collaboration can be in driving scientific progress.

REFERENCES

Osman, S., Alluqmani, A., Ariuzzaman, M., Aburizaiza, M., Sindi, W., & Alam, M. (2021). Application of Machine Learning for Prediction of Early Seven-Day Strength of Concrete in Factories. *Third International Sustainability and Resilience Conference: Climate Change - The Sustainability and Resilience 2021*.

Rao, G. A. (2001). Generalization of Abrams' law for cement mortars. *Cement and Concrete Research*, *31*(3), 495-502.

Armaghani, D. J., & Asteris, P. G. (2020). A comparative study of ANN and ANFIS models for the prediction of cement-based mortar materials compressive strength. *Springer London*. https://doi.org/10.1007/s11771-019-4114-5

Asteris, P. G., Skentou, A. D., Bardhan, A., Samui, P., & Pilakoutas, K. (2021). Predicting concrete compressive strength using hybrid ensembling of surrogate machine learning models. *Cement and Concrete Research*. https://doi.org/10.1016/j.cemconres.2021.106449

Hossain, M., Sarkar, D., Zaman, A., Arifuzzaman, M., Gazder, U., Tito, M., & Al-Duais, A. (2023). Revolutionizing solar energy: enhancing panel efficiency with advanced machine learning techniques and data filtering. *7th IET Smart Cities Symposium (SCS 2023)*.

Nguyen-Sy, T., Wakim, J., To, Q.-D., Vu, M.-N., Nguyen, T.-D., & Nguyen, T.-T. (2020). Predicting the compressive strength of concrete from its compositions and age using the extreme gradient boosting method. *Construction and Building Materials, 260*, 119757. https://doi.org/10.1016/j.conbuildmat.2020.119757

Bardhan, A., Ozcan, N. T., Asteris, P. G., & Gokceoglu, C. (2024). Hybrid ensemble paradigms for estimating tunnel boring machine penetration rate for the 10-km long Bahce-Nurdagi twin tunnels. *Engineering Applications of Artificial Intelligence*, *136*, 108997.

Khan, M. I. (2012). Predicting properties of high performance concrete containing composite cementitious materials using artificial neural networks. *Automation in Construction, 22*, 516–524.

DeRousseau, M. A., Laftchiev, E., Kasprzyk, J. R., Rajagopalan, B., & Srubar, W. V. (2019). A comparison of machine learning methods for predicting the compressive strength of field-placed concrete. *Construction and Building Materials, 228*, 116661. https://doi.org/10.1016/j.conbuildmat.2019.08.042

Khademi, F., Akbari, M., Jamal, S. M., et al. (2017). Multiple linear regression, artificial neural network, and fuzzy logic prediction of 28 days compressive strength of concrete. *Frontiers of Structural and Civil Engineering, 11*, 90–99. https://doi.org/10.1007/s11709-016-0363-9

Yeh, I. C. (1998). Modeling of strength of high-performance concrete using artificial neural networks. *Cement and Concrete Research, 28*(12), 1797–1808.

Sevim, U. K., Bilgic, H. H., Cansiz, O. F., Ozturk, M., & Atis, C. D. (2021). Compressive strength prediction models for cementitious composites with fly ash using machine learning techniques. *Construction and Building Materials, 271*, 121584. https://doi.org/10.1016/j.conbuildmat.2020.121584

Yaseen, Z. M., Deo, R. C., Hilal, A., Abd, A. M., Bueno, L. C., Salcedo-San, S., & Nehdi, M. L. (2018). Predicting compressive strength of lightweight foamed concrete using extreme learning machine model. *Advances in Engineering Software, 115*, 112–125. https://doi.org/10.1016/j.advengsoft.2017.09.004

Yuan, Z., Wang, L.-N., & Ji, X. (2014). Prediction of concrete compressive strength: Research on hybrid models genetic based algorithms and ANFIS. *Advances in Engineering Software, 67*, 156–163. https://doi.org/10.1016/j.advengsoft.2014.03.007