Evaluation and Prediction of W/C Ratio vs. Compressive Concrete Strength Using A.I and M.L Based on Random Forest Algorithm Approach

R. Jamalpour*

Department of Civil Engineering, Karaj Branch, Islamic Azad University, Karaj, Iran.
Received: 28 February 2024 - Accepted: 08 June 2024

Abstract

Concrete, an artificial stone composed of cement, aggregate, water, and additives, is extensively utilized in contemporary civil projects. A pivotal characteristic of concrete is its capacity to efficiently serve various purposes and structural requirements. Cement, water, aggregate, and additives are pivotal parameters wherein even minor alterations can significantly impact concrete strength. Among these parameters, the Water/Cement (W/C) ratio holds particular significance due to its inverse correlation with strength. Traditionally, predicting concrete strength solely based on the water-to-cement ratio has been challenging. However, with advancements in AI and machine learning techniques coupled with ample data availability, accurate strength prediction is achievable. This paper presents an analysis of a diverse dataset comprising various concrete tests utilizing machine learning methodologies, followed by a comparative examination of the outcomes. Furthermore, this study scrutinizes a renowned dataset encompassing 1030 experiments, featuring diverse combinations of cement, water, aggregate, etc., employing artificial intelligence and machine learning techniques. Model accuracy and result fidelity are evaluated through rigorous sampling methodologies. Initially, the dataset is subjected to analysis utilizing the linear regression algorithm, followed by validation employing the random forest algorithm. The random forest algorithm is employed to predict the water-to-cement ratio and corresponding compressive strength for concrete with a density of 300 kg/m3. Notably, the obtained results exhibit a high level of concordance with experimental and laboratory findings from prior studies. Hence, the efficacy of the random forest algorithm in concrete strength prediction is established, offering promising prospects for future applications in this domain.

Keywords: Concrete, AI, Machine Learning, W/C Ratio, Prediction, Strength.

1. Introduction

1.1. Concrete

Concrete is a versatile and widely-used construction material composed of a mixture of cement, water, aggregates (sand or gravel), and sometimes additives or admixtures. (Fig. 1.)It is known for its strength, durability, and versatility in various construction applications. The fabrication of concrete entails the amalgamation of cement, water, and aggregates to create a paste that progressively solidifies through a chemical reaction known as hydration. This reaction forms a robust material adept at bearing substantial loads and enduring severe environmental conditions. Concrete's composition can be tailored to fulfill specific project demands by modifying the ratios of its constituents and integrating additives or admixtures to enhance attributes such as workability, strength, or durability. It can be molded into diverse shapes and sizes, enabling the construction of varied structures including buildings, bridges, pavements, and dams.A significant benefit of concrete is its extended lifespan and minimal maintenance demands, rendering it a cost-effective option for numerous construction endeavors. Moreover, concrete is recognized for its fire resistance, energy efficiency, and environmental sustainability when produced through eco-friendly methods [1]. Fundamentally, concrete is integral to contemporary construction owing to its adaptability, longevity, and capacity to meet extensive structural and aesthetic needs. Its pervasive utilization within the industry underscores its essential role as a primary construction material that is continually enhanced by progressive research and innovation [2].

Fig. 1. Concrete as a natural and man-made stone with many uses.

1.2. Concrete Application

Concrete serves as a versatile and robust material for construction, extensively utilized across a broad spectrum of applications today. It is employed in various sectors, including building construction, infrastructure, landscaping, water management,

*Corresponding author
Email address: jamalpour-reza@kiau.ac.ir

architectural features, industrial applications, marine construction, the energy sector, art, sculpture, and transportation. Concrete's use in diverse projects worldwide has been prominently demonstrated in structures built from 2010 to the present. An illustrative example is the One World Trade Center in New York City, USA. Completed in 2014, this iconic skyscraper is a testament to resilience and strength. Its foundation and core structure, comprised of reinforced concrete, provide crucial stability and support across its 104 floors(Fig.2.). Burj Khalifa (Dubai, UAE), the tallest building in the world, completed in 2010, features a concrete core that extends up to the 160th floor. The use of high-strength concrete allowed for the construction of this record-breaking structure (Fig. 3.). The Shard (London, UK), Completed in 2012, this 95-story skyscraper is known for its distinctive glass façade and concrete core. The Shard's concrete core provides structural integrity and support for the building's height (Fig. 4.). Marina Bay Sands (Singapore), completed in 2010, this integrated resort complex features three towers connected by a unique sky park [3].

Fig. 2. One World Trade Center (New York City, USA), Completed in 2014.

Fig. 3. Burj Khalifa (Dubai, UAE) ,the tallest building in the world.

Fig. 4. Shard (London, UK),Completed in 2012.

The towers are constructed with reinforced concrete, showcasing the material's versatility in creating complex structures (Fig. 5.).

The Lotte World Tower (Seoul, South Korea), completed in 2016, this 123-story skyscraper boasts a reinforced concrete core that supports its impressive height. The tower's innovative design showcases the use of concrete in modern high-rise construction (Fig. 6.). The Louvre Abu Dhabi (Abu Dhabi, UAE), opened in 2017, this iconic museum features a stunning dome structure made of steel and concrete. The concrete elements provide a solid foundation for the building's unique design and architectural features (Fig. 7.). The National Museum of Qatar (Doha, Qatar), completed in 2019, this museum showcases a striking design inspired by desert rose crystals.

Fig. 5 Marina Bay Sands (Singapore), Completed in 2010.

Fig. 6. The Lotte World Tower (Seoul, South Korea), Completed in 2016.

Fig. 7. The Louvre Abu Dhabi (Abu Dhabi, UAE), Opened in 2017.

The building's concrete structure complements the intricate façade and interior spaces, highlighting the material's adaptability in modern architecture (Fig. 8.). The Leeza SOHO Tower (Beijing, China), completed in 2019, this impressive skyscraper features a distinctive twisting design with a central atrium.

The building's reinforced concrete core provides stability and support for its 46 floors, showcasing the use of concrete in innovative high-rise construction (Fig. 9.). The Vessel (New York City, USA), completed in 2019, this interactive art installation and observation deck is made up of a network of interconnected staircases.
The structure's concrete foundation supports the intricate geometry of the design, creating a unique urban landmark (Fig. 10.). The Museum of the Future (Dubai, UAE), Set to open in 2020, this futuristic museum is characterized by its dynamic torus shape.

Fig. 8. The National Museum of Qatar (Doha, Qatar), Completed in 2019.

Fig. 9. The Leeza SOHO Tower (Beijing, China), Completed in 2019.

Fig. 10. The Vessel (New York City, USA), Completed in 2019.

The building's concrete structure forms the foundation for the innovative design, demonstrating the material's role in shaping cutting-edge architectural projects (Fig. 11.).

Fig. 11. The Museum of the Future (Dubai, UAE), Set to open in 2020.

2. Concrete Component

Concrete is a composite material consisting of several essential components that combine to form a robust and durable construction material (Fig. 12.). At its core, concrete is made up of cement, water, and aggregates like sand and gravel. Cement serves as the binder, holding the mixture together, while water activates the cement, enabling it to harden. Aggregates contribute to the bulk and strength of the mix, with sand occupying the gaps between the larger gravel particles. Furthermore, additives such as admixtures may be incorporated to improve specific characteristics of the concrete, such as enhancing its workability or minimizing its water requirement. The precise combination of these ingredients yields a versatile construction material that can be shaped into a variety of forms and structures. Proper mixing, placement, and curing of concrete are critical to ensuring its structural integrity and longevity.

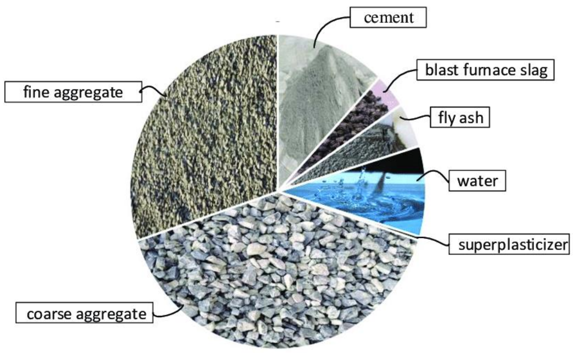

Fig. 12. Concrete Component.

2.1. Cement in Concrete

Cement is an integral component in concrete production, essential for its strength and durability. It is a fine powder derived from a combination of limestone, clay, and various minerals, which are subjected to high temperatures in a kiln through a process known as calcination. This procedure produces clinker, which is subsequently ground into a fine powder to form cement. Within the concrete mixture, cement serves as a binding agent, adhering the other components together. When it interacts with water, a chemical reaction known as hydration occurs, resulting in a paste that gradually hardens and binds the aggregates. This reaction is crucial for concrete's ability to bear substantial loads and maintain structural integrity. The specific type and quality of cement utilized can markedly influence the characteristics of the final concrete product. Various types of cement, such as Portland cement or blended cement, provide differing degrees of strength, durability, and workability to the concrete (Fig. 13.).

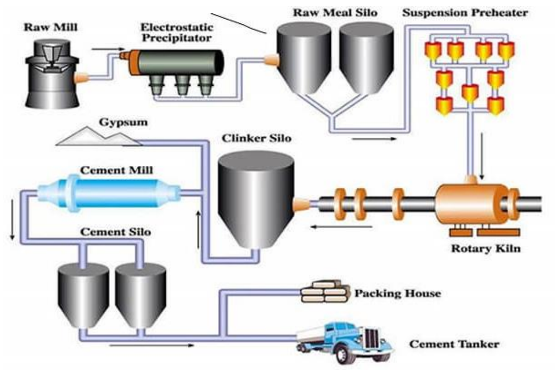

Fig. 13. Portland Cement Factory Process.

Furthermore, additives or admixtures may be integrated into the concrete mix to enhance specific properties of the concrete. Cement is fundamental in the construction sector, serving as the primary component of concrete, which is among the most extensively utilized construction materials globally. The robustness and durability imparted by cement are crucial for the longevity and structural integrity of concrete structures such as buildings, bridges, roads, and dams. The pivotal role of cement in concrete is undeniable, as it forms the cornerstone of modern construction projects. Its capacity to forge sturdy and enduring structures has rendered it indispensable in infrastructure development and building construction. With the increasing emphasis on sustainable and resilient construction methods, the contribution of cement to producing high-quality concrete remains critical in addressing the dynamic demands of the evolving construction industry.

2.2. Water in Concrete

Water is pivotal in the production and quality of concrete, as both the quantity and quality of water utilized in the mix substantially affect the workability, strength, durability, and overall performance of the final product. The purity of water in concrete production is critical. Contaminants such as high levels of chloride, sulfate, or organic matter can adversely impact the setting time, strength, and durability of the concrete. Employing clean, potable water devoid of impurities is essential for producing high-quality concrete. Adequate moisture is crucial for effective hydration, enabling the concrete to reach its intended strength. (Fig. 14.). Proper curing methods, including moist curing or the application of curing compounds, are necessary to preserve the required moisture levels within the concrete. Additionally, water assists in managing the temperature of the concrete mix during placement and curing. Extreme temperatures can disrupt the hydration process and result in defects like cracking in the hardened concrete. Appropriate temperature control, facilitated through the use of chilled water or ice in warm conditions or heated water in colder environments, is vital to ensuring the concrete's quality and performance. In summary, the quantity and quality of water in concrete production are essential determinants of the strength, durability, and overall efficacy of the final concrete product.

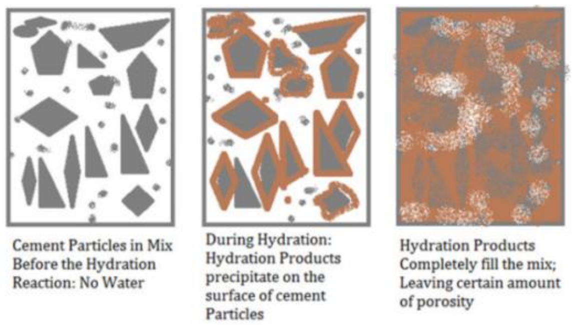

Fig. 14. Hydration Process of cement.

2.3. Water/Cement (W/C) Ratio

The proportion of water to cement in a concrete mixture, referred to as the water-cement ratio, is commonly abbreviated as W/C. This ratio, along with factors such as aggregate type, cement type, and the specific properties desired in the concrete, plays a critical role in the formulation of a concrete mix design. Understanding the significance of the water-cement ratio is essential, as it significantly influences overall characteristics of concrete. It is imperative to delve into this aspect to accurately determine and optimize the ratio to ensure the desired performance of the concrete.

2.4. Importance of W/C Ratio

The water-cement ratio is a crucial factor in the concrete mix that determines the amount of water needed to produce a workable and durable mixture.

A. Workability: The application and usage of concrete in construction are largely facilitated by achieving the right w/c ratio. A proper W/C ratio guarantees that the mixture is just moist enough to prevent leakage of cement or excessive settling, and it also eases the flow of the mixture while reducing the builders' workload. Only with perfect consistency can it have its best durability.

B. Strength: The W/C ratio significantly impacts the strength of concrete. If the water used is not in excess or less than required, the cement walls weaken or leave voids in the finished product.

C. Durability: The durability of concrete over the long term is influenced by ideal W/C ratios. A lower ratio produces concrete that is denser and less porous, which makes it resistant to chemicals, moisture intrusion, and other types of threats leading to degradation.

D. Shrinkage and cracking: Increased shrinkage can cause possible cracking, which can result from an unbalanced water-to-cement ratio. The cement's water content can affect how much it shrinks, whereas a material with a low water content may not be properly hydrated and be more prone to breaking over time.

E. Permeability: An excessive amount of water in the mix can make concrete more permeable, enabling dangerous elements to seep within, causing internal damage and reducing durability.

2.5. Water Amount (Less or Much) in Concrete

When too much water is added to the concrete, thus altering the W/C ratio, it can reduce the strength of the mixture and cause cracks or shrinkage in a short period of time. On the other hand, if too little water is added to the mixture, it increases the effort required to mix and use it during construction. The mixture becomes rigid and too hard to use. We need to strike the perfect balance in order to obtain the perfect bonding material.(Fig. 15. (a). Fig.(b).)

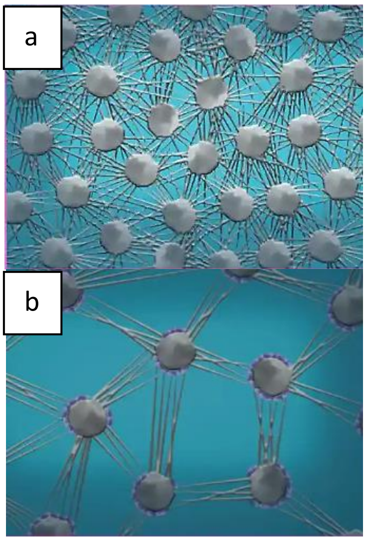

Fig. 15. (a) Good Hydration in Idea Water/Cement Ratio, (b) Bad Hydration in Badi Water/Cement Ratio Effect of water-cement ratio on cement hydration.

2.7. Balanced W/C Ratio achieving

The optimal water-cement ratio for a specific concrete mix must be carefully determined, taking into account factors such as the intended strength of the concrete, environmental conditions, aggregate characteristics, and the type of cement utilized. Concrete mix designers adhere to established standards, guidelines, and recommended best practices, such as those provided by the ACI code, to achieve the ideal ratio. The relationship between the concrete's compressive strength and the water-cement ratio is illustrated in Fig. 15. Notably, a lower water-cement ratio correlates with decreased compressive strength, and an increase beyond the appropriate level also leads to a reduction in strength. It is important to recognize that the ideal water-cement ratio is influenced not only by the concreting process—whether it involves vibration or not—but also by the type of aggregate (Coarse or Fine) and the nature and quantity of additives used. (Fig. 16 and Fig. 17)

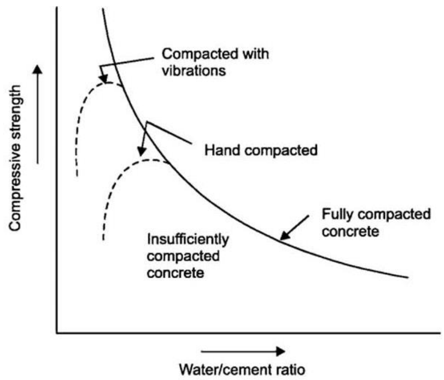

Fig. 16. Relationship between compressive strength and water/cement ratio.

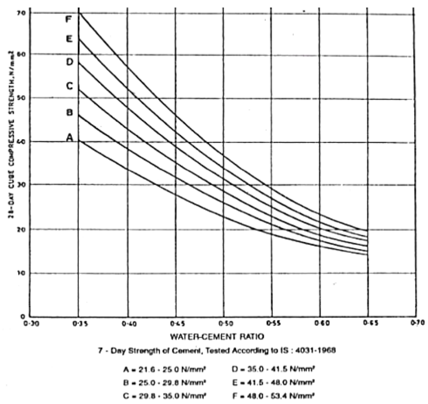

Fig. 17. Relationship between compressive strength and water/cement ratio for Different cement.

3. Artificial Intelligence and Machine Learning (ML)

3.1. Machine Learning and Methods

Machine learning is a subset of artificial intelligence that focuses on the development of algorithms and statistical models that enable computer systems to learn from and make predictions or decisions based on data, without being explicitly programmed. Machine learning algorithms can be classified into three main types: supervised learning, unsupervised learning, and reinforcement learning.

a) Supervised Learning: In supervised learning, the algorithm is trained on labeled data, where each data point is associated with a label or target variable. The algorithm learns to map input variables to the output variable by minimizing the error between predicted and actual values. Examples of supervised learning include classification and regression.

b) Unsupervised Learning: In unsupervised learning, the algorithm is trained on unlabeled data, where there are no target variables. The algorithm learns to identify patterns or structures in the data by clustering or dimensionality reduction techniques. Examples of unsupervised learning include clustering and anomaly detection.

c) Reinforcement Learning: In reinforcement learning, the algorithm learns by interacting with an environment and receiving feedback in the form of rewards or penalties. The goal is to learn a policy that maximizes the cumulative reward over time. Examples of reinforcement learning include game playing and robotics.

Machine learning methods also include techniques such as deep learning, which uses neural networks with multiple layers to extract features from data, and decision trees, which partition the data into smaller subsets based on decision rules. Other methods include support vector machines, k-nearest neighbors, and random forests. The choice of method depends on the nature of the problem and the characteristics of the data.

3.2. Types of Machine Learning Algorithms

There are several types of machine learning algorithms, including:

1-Regression algorithms: These algorithms are used to predict a continuous value, such as the price of a house or the temperature.

2-Classification algorithms: These algorithms are used to classify data into different categories, such as spam or not spam emails.

3-Clustering algorithms: These algorithms are used to group data points into clusters based on similarity.

4-Dimensionality reduction algorithms: These algorithms are used to reduce the number of features in a dataset while retaining as much information as possible.

5-Neural network algorithms: These algorithms use artificial neural networks to learn patterns in data.

6-Decision tree algorithms: These algorithms use decision trees to make predictions based on a series of if-then statements.

7-Support vector machine algorithms: These algorithms are used for classification and regression tasks and work by finding the best boundary between different classes.

8-Random forest algorithms: These algorithms use multiple decision trees to make predictions and reduce overfitting. The Random Forest algorithm is a popular and versatile machine learning method used for both classification and regression tasks. It belongs to the family of ensemble learning methods, which combine the predictions from multiple models to produce a more accurate or robust prediction than any single model.

9-Reinforcement learning algorithms: These algorithms learn through trial and error by receiving feedback in the form of rewards or penalties.

There is no silence between the machine learning algorithms. Each algorithm serves a specific purpose and can be used in combination with other algorithms to improve the accuracy and efficiency of a machine learning model. The choice of algorithm depends on the type of problem, the available data, and the desired outcome. The choice of algorithm depends on the nature of the problem and the characteristics of the data.

4. Literature Survey

In an investigation three machine learning (ML) techniques, namely K-nearest neighbor (KNN), linear regression (LR), and artificial neural network (ANN), were employed to forecast the compressive strength (fc') of nanocomposites containing Carbon Nanotubes (CNTs). A comprehensive database comprising 282 data entities for CNT-based concrete was systematically compiled, and the model's reliability was evaluated using the R2 test and statistical error analysis. The study reveals that the properties of CNT-based concrete composites are significantly influenced by the water-to-binder ratio, followed by the proportions of cement and coarse aggregates. Notably, the ML algorithms demonstrated superior generalization capabilities, indicating their potential for precise predictions of CNTs-based concrete properties [1].. In a separate study the electrochemical parameters of cement-based materials subjected to varying water-cement ratios during carbon curing and water curing were assessed using electrochemical impedance spectroscopy (EIS). Through this technique, an optimized circuit model and corresponding electrical parameters were derived to elucidate the alterations in the microstructure of cementitious materials following carbon capture. The findings suggest a notable increase in the semicircle diameter in the high-frequency region with the progression of both carbon curing and water curing processes. Particularly noteworthy is the observation that the electrochemical parameters ρct2 of cement mortar specimens subjected to carbon curing were approximately three times higher than those subjected to water curing, primarily due to the reduction of soluble materials and refinement of connecting pores within the microstructure of cementitious materials. Moreover, a quadratic function correlation between the rate of mass increase and ρct2 during the carbonation process of cement mortar was established, affirming the applicability of EIS analysis for monitoring carbon capture in cement-based materials, whether newly mixed concrete or recycled concrete aggregates. In another explored the application of electrochemical impedance spectroscopy (EIS) combined with artificial neural networks (ANNs) for estimating the water-to-cement (w/c) ratio in uncured cementitious materials. This methodology leverages the sensitivity of EIS to changes in the electrochemical properties of the material, such as capillary porosity in cement paste, facilitating the estimation of the w/c ratio. The researchers demonstrated that EIS, when interpreted through ANNs trained with principal components, could predict the w/c ratio with a mean absolute error of 0.014. This approach enhances the objectivity in interpreting EIS data and shows promise for in-situ applications in the construction industry [2]. In a related development, employed machine learning techniques with a multiclass classification strategy to predict the slump of industrially produced concrete. Their approach underscores the growing integration of advanced computational methods in concrete technology [3]. Furthermore, evaluated the efficacy of three optimization techniques—electromagnetic field optimization (EFO), teaching-learning-based optimization (TLBO), and water cycle algorithm (WCA), alongside multi-tracker optimization algorithm (MTOA)—in optimizing certain performance parameters [4]. In another innovative application, implemented deep learning algorithms to monitor the progression of alkali-silica reaction (ASR) in concrete. Their study used acoustic emission (AE) signals from concrete samples embedded with reactive coarse aggregates and steel rebars. The AE signals were processed and categorized into two distinct classes. The deployment of convolutional neural networks (CNN) and stacked autoencoders for signal classification revealed that the CNN-based model achieved higher accuracy compared to the autoencoder-based model, illustrating the potential of deep learning in structural health monitoring [5]. In their 2021 study, Song et al. [6] undertook an analysis involving the collection of data from experimental trials to apply machine learning (ML) techniques for predicting the compressive strength (CS) of concrete containing fly ash. The chemical and physical characteristics of all materials used were thoroughly assessed. The study gathered a total of 98 data points, with seven variables—cement, fly ash, superplasticizer, coarse aggregate, fine aggregate, water, and curing days—used as inputs to predict the CS. The reliability of the experimental data was enhanced through k-fold cross-validation, and the model's accuracy was assessed using the coefficient of determination (R^2), root mean error (RME), and root-mean-square error (RMSE). Additionally, statistical analyses were conducted to further validate the performance of the model. Notably, the bagging algorithm demonstrated superior predictive capabilities, as evidenced by its high R^2 value of 0.95, compared to the genetic expression programming (GEP), artificial neural network (ANN), and decision tree (DT) models, which recorded R^2 values of 0.86, 0.81, and 0.75, respectively. In a related research proposed a novel hybrid artificial intelligence (AI) model that integrates the least squares support vector regression (LSSVR) with grey wolf optimization (GWO). This model was designed to effectively consider various influencing factors and enhance the predictive accuracy concerning the compressive strength of foamed concrete. The evaluation of this GWO-LSSVR model indicated excellent congruence between actual and predicted values, with a correlation coefficient of 0.991 and a mean absolute percentage error (MAPE) of 3.54%. Consequently, this innovative AI model is recommended as an effective tool for the design and optimization of foamed concrete materials [7]. In another study, developed an AI-based framework that effectively models the complex behavior of fire-exposed reinforced concrete (RC) structural members. This framework accounts for the high-temperature material properties of concrete and steel reinforcement, including phenomena such as creep deformation and fire-induced spalling. Notably, the model operates without requiring temperature-dependent material properties or the use of specialized simulation/analysis software [8]. In another study proposed a hybrid model that combines the Random Forest (RF) technique with the Firefly Algorithm (FA) to predict damage index factors (DIF) for steel fiber reinforced concrete (SFRC). The model was trained and validated using 193 and 314 DIF data samples for the compressive and tensile strengths of SFRC, respectively, sourced from existing studies. Input variables included strain rate, matrix strength, fiber dosage, and fiber properties such as shape, aspect ratio, and tensile strength. The results indicated that the model efficiently and accurately predicts DIF values for SFRC, with matrix strength identified as the most influential factor on DIF values. Rachel Cook et al. (2019) introduced a novel hybrid machine learning (ML) model that integrates Random Forests (RF) with the Firefly Algorithm (FFA) for predicting the compressive strength of concrete. This RF-FFA model was compared against other ML models, including Support Vector Machine (SVM), Multilayer Perceptron Artificial Neural Network (MLP-ANN), M5Prime model tree (M5P), and RF.

The evaluation employed various statistical parameters and a Composite Performance Index (CPI), demonstrating that the RF-FFA model consistently outperformed the standalone models across diverse datasets in terms of prediction accuracy [9]. Effects of water-cement ratio and chloride ions on the meso-structure of concrete studied. This research involved immersing concrete cubes with varying water-cement ratios in both fresh and salt water, followed by an Electrochemical Impedance Spectroscopy (EIS) analysis using an electrochemical workstation. Findings indicated that a reduction in the water-cement ratio does not affect the electric double-layer capacitance of concrete; instead, it decreases porosity and densifies the internal structure of the concrete [10]. Study on optimizing Artificial Neural Networks (ANN) using the Ant Lion Optimization (ALO) for predicting concrete slump demonstrated promising results. With calculated Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) values of 3.0286 and 3.7788, respectively, the ALO proved to be an effective optimizer. This model outperformed other optimization techniques, including the Grasshopper Optimization Algorithm and Biogeography-Based Optimization, establishing ALO as a robust method for this application [11,12].

5. Research Methodology

In this study, two machine learning algorithms are employed and evaluated for their ability to evaluate relation of water-cement (W/c) ratio and compressive strength. The algorithms implemented include Linear Regression and Random Forest. Utilizing the Sklearn library in Python, along with the Split module, the dataset was partitioned into training and testing subsets. The dataset serves as a fundamental element in machine learning, crucial for both training and evaluating models. Subsequently, both algorithms were trained using the dataset and their performance was assessed. The evaluation was conducted using the R2 score metric, with results detailed in the Analysis and Results section of this paper. The quality and comprehensiveness of the dataset significantly influence the accuracy and efficacy of the model. An optimal dataset should be diverse, representative, and sufficiently balanced to cover all potential outcomes, enabling the machine learning model to acquire a comprehensive understanding from the data. Moreover, the size of the dataset is important to ensure it contains adequate information for the model's learning process.
The dataset not only aids in training but also in gauging the performance of the model and pinpointing areas for enhancement, underscoring its pivotal role in the success of machine learning initiatives.

6. Selected Dataset

The selected data set, which contains 1030 samples with details given at Table. 1., has shown the relationship between different components of concrete and compressive strength of concrete. The dataset is commonly known as the "Concrete Compressive Strength Data Set." It was collected to help in predicting the compressive strength of concrete samples based on their mixture components. The dataset is hosted by the UCI Machine Learning Repository and is widely used in both educational settings and research for machine learning experiments. It contains 9 attributes per sample, which are: Cement (kg in a m³ mixture), Blast Furnace Slag (kg in a m³ mixture), Fly Ash (kg in a m³ mixture), water (kg in a m³ mixture), Superplasticizer (kg in a m³ mixture), Coarse Aggregate (kg in a m³ mixture), Fine Aggregate (kg in a m³ mixture), Age (in days), concrete compressive strength (MPa, megapascals). In This dataset also added a new column that show water and cement ratio.

Table. 1. Selected Dataset with 1030 samples.

	Cement (component 1) (kg in a m^3) mixture	Blast FurnaceSlag (component 2) (kg in a m^3) mixture	Fly Ash (component 3) (kg in a m^3) mixture	Water (component 4) (kg in a m^3) mixture	water cement (w/C) Ratio	Super plasticizer (component 5) (kg in a m^3) mixture	Coarse Aggregate (component 6) (kg in a m^3) mixture	Fine Aggregate (component 7) (kg in a m^3 mixture)	Age (day)	Concrete compressive strength (Mpa) (megapascals)
1	540	0	0	162	30	2.5	1040	676	28	79.99
2	540	0	0	162	30	2.5	1055	676	28	61.89
3	332.5	142.5	0	228	68	0	932	594	270	40.27
4	332.5	142.5	0	228	68	0	932	594	365	41.05
5	198.6	132.4	0	192	96	0	978.4	825.5	360	44.3
6	266	114	0	228	85	0	932	670	90	47.03
7	380	95	0	228	60	0	932	594	365	43.7
8	380	95	0	228	60	0	932	594	28	36.45
9	266	114	0	228	85	0	932	670	28	45.85
10	475	0	0	228	48	0	932	594	28	39.29
11	198.6	132.4	0	192	96	0	978.4	825.5	90	38.07
12	198.6	132.4	0	192	96	0	978.4	825.5	28	28.02
13	427.5	47.5	0	228	53	0	932	594	270	43.01
⋮										
1025	159.8	250	0	168.4	105	12.2	1049.3	688.2	28	39.46
1026	166	259.7	0	183.2	110	12.7	858.8	826.8	28	37.92
1027	276.4	116	90.3	179.6	64	8.9	870.1	768.3	28	44.28
1028	322.2	0	115.6	196	60	10.4	817.9	813.4	28	31.18
1029	148.5	139.4	108.6	192.7	129	6.1	892.4	780	28	23.7
1030	159.1	186.7	0	175.6	110	11.3	989.6	788.9	28	32.77
1030	260.9	100.5	78.3	200.6	76	8.6	864.5	761.5	28	32.4

7. Results and Discussion

7.1. Data Correlation Analysis

Correlation is a statistical measure that shows how two or more variables fluctuate in relation to each other. A positive correlation means that the variables increase or decrease in parallel, while a negative correlation means that one variable increases as the other decreases.
A Heat map is a type of plot that shows the amount of data using colors. In this chart, each table cell is marked with a color that shows the value of that cell. The warm colors display cells with higher values, and cold colors show cells with less value. This chart is used to examine patterns and relationships between data. The data correlation for used variables (constituent components of tested concrete samples) is shown in Fig. 18.

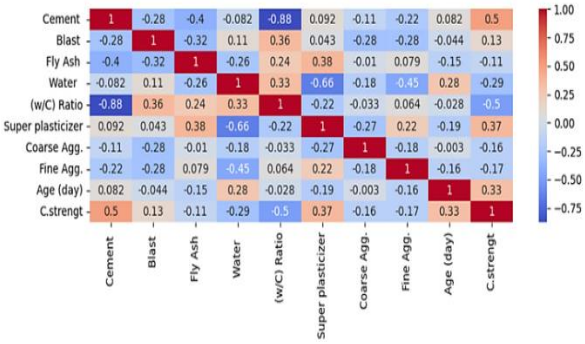

Fig. 18. Correlation of Selected Dataset with 1030 samples.

7.2. Data Relation Analysis

A scatter plot is a type of data visualization that uses dots to represent the values obtained for two different variables. This plot is primarily used to observe and show the relationship between the two variables, helping to determine how closely the two variables are related.

The Kernel Density Estimation Plot (KDE), scatter plot of compressive Strength vs (W/C) Ratio and Cement vs (W/C) Ratio are shown in Fig.19., Fig20.andFig. 21., and also 3D scatter plot of compressive Strength vs (W/C) Ratio and Cement is shown in Fig. 22.

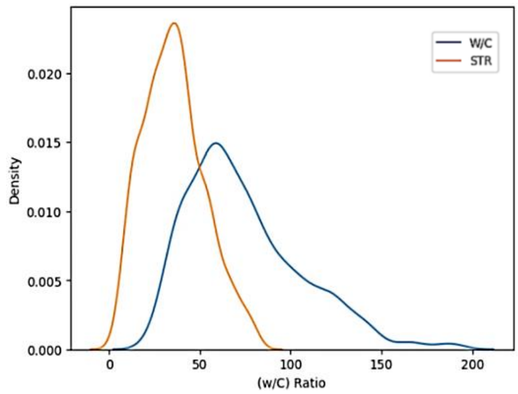

Fig. 19. Kernel Density Estimation Plot (KDE).

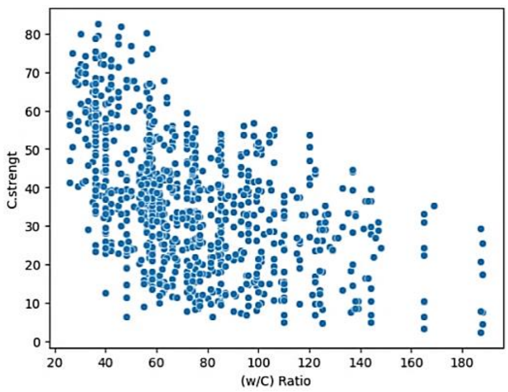

Fig. 20. Compressive Strength vs (W/C) Ratio.

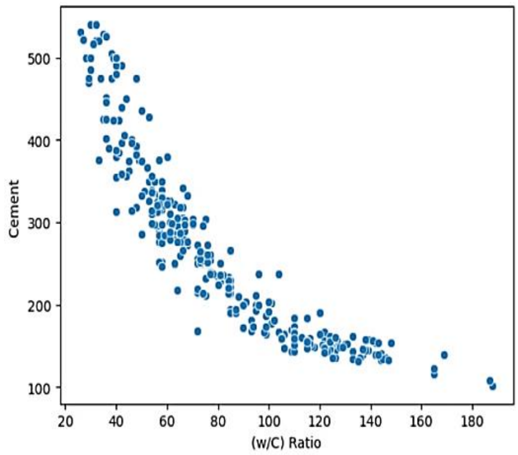

Fig. 21. Cement vs (W/C) Ratio.

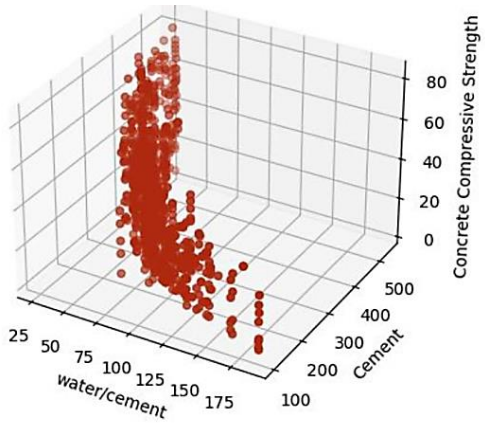

Fig. 22. 3D Scatter plot of Compressive Strength vs (W/C) Ratio and Cement.

7.3. Accuracy (R^2) Analysis

The results of random forest and linear regression are given at Table.2. Random forest algorithm has better performance than linear regression for this dataset to predict concrete compressive strength.

Table. .2. Comparison on Random Forest and Linear Regression

	Random Forest	Linear Regression
R^2 Score	0.886	0.600

7.4. Prediction

The machine learning algorithm used can also be used for prediction. For example, in the data set used, the concrete strength is predicted in relation to the W/C ratio (for concrete with 300 kg/m3 of cement) and the results are shown in Fig. 23. The strength changes (decrease in resistance with increasing water/cement ratio) show a realistic trend.

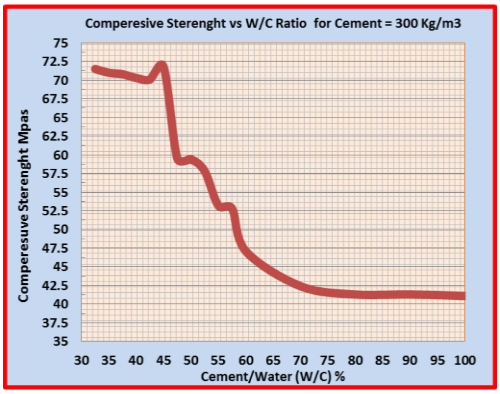

Fig. 23. Compressive Strength vs W/C Ratio Prediction.

8. Conclusion

1. This research demonstrates that artificial intelligence (AI) and machine learning (ML) methodologies are effective in predicting the strength and durability of concrete based on a range of input variables, including the water-to-cement (W/C) ratio, mix design, curing conditions, and environmental influences. Furthermore, machine learning algorithms excel at analyzing extensive datasets to discern patterns and optimize concrete mix designs tailored for specific applications.
Thus, ML proves to be a robust tool for analyzing concrete test datasets and estimating outcomes based on diverse sample characteristics.

2. Among the prevalent algorithms, the Random Forest algorithm stands out due to its high accuracy in both classification and regression tasks. This effectiveness is achieved through the ensemble approach, where multiple deep decision trees, each trained on different subsets of the same data, collectively reduce variance without increasing bias. This algorithm's capability to handle thousands of input variables without the need for variable deletion makes it particularly adept at managing datasets with high dimensionality, thus ensuring good performance without necessitating feature selection.

3. In this study, the Random Forest algorithm was identified as especially suitable for predicting outcomes, evidenced by achieving the highest R^2 score among the evaluated methods. By accurately predicting the optimal water-to-cement ratio, enhanced performance and quality of concrete can be anticipated, thereby contributing to environmental sustainability.

4. Historical data and findings from this research further corroborate that an increase in the water-to-cement ratio significantly reduces the compressive strength of concrete, establishing an inverse relationship between these two parameters.
These conclusions align with empirical realities and are consistent with findings from other scholarly investigations, underscoring the robustness and applicability of the Random Forest algorithm in concrete research.

References

[1] Sun H, Amin MN, Qadir MT, Arifeen SU, Iftikhar B, Althoey F. Investigating the effectiveness of carbon nanotubes for the compressive strength of concrete using AI-aided tools. Case Stud. Constr. Mater. 2024; 1(20):e03083.

[2] Liu Q, Tang H, Chi L, Chen K, Zhang L, Lu C. Evolution of electrochemical impedance spectra characteristics of cementitious materials after capturing carbon dioxide. Sustainability. 2023;15(3):2460.

[3] Park J, Song H, Choi H. Estimation of water-to-cement ratio in cementitious materials using electrochemical impedance spectroscopy and artificial neural networks. Construction and Building Materials. 2022; 3(350):128843.

[4] Zhou G, Moayedi H, Foong LK. Teaching–learning-based metaheuristic scheme for modifying neural computing in appraising energy performance of building. Eng. with Computers. 2021; 37:3037-48.

[5]. Li A, Soltangharaei V, Ziehl P. Evaluation of ASR in concrete using acoustic emission and deep learning. Nucl. Eng. Des. 2021; 380:111328

[6]. Song H, Ahmad A, Farooq F, Aslam F. Predicting the compressive strength of concrete with fly ash admixture using machine learning algorithms. Constr. Build. Mater. 2021; 308:125021

[7]. Pham AD, Ngo NT, Nguyen QT, Truong NS. Hybrid machine learning for predicting strength of sustainable concrete. Soft Computing. 2020; 24(19):14965-80.

[8] Naser MZ. AI-based cognitive framework for evaluating response of concrete structures in extreme conditions. Engineering Applications of Artificial Intelligence. 2019; 1(81):437-449.

[9] Ai L, Soltangharaei V, Ziehl P. Evaluation of ASR in concrete using acoustic emission and deep learning. Nuclear Engineering and Design. 2021; 15(380):111328.

[10] jiao F, Jinxin J, Jichao G, Wang ZH. Study on the Concrete in Chloride Environment Based on Electrochemical Impedance Spectroscopy. Int. J. Pattern Recognit. Artif. Intell. 2019; 34(6):2059017.

[11] Yang L, Qi C, Lin X, Li J, Dong X. Prediction of dynamic increase factor for steel fibre reinforced concrete using a hybrid artificial intelligence model. Eng. Struct. 2019; 15(189):309-18.

[12] Cook R, Lapeyre J, Ma H, Kumar A. Prediction of compressive strength of concrete: critical comparison of performance of a hybrid machine learning model with standalone models. J. Mater. Civ. Eng. 2019; 31(11):04019255.