Available online www.jsaer.com

***Journal of Scientific and Engineering Research*, 2021, 8(7):18-26**

Research Article

ISSN: 2394-2630
CODEN(USA): JSERBR

Concrete Compressive Strength Estimation using Artificial Intelligence Approach

Nanchen Nimyel Caleb*[1], Selfa Johnson Zwalnan[2], Wesley Odeh Odumu[1]

[1]Department of Computer Engineering, Plateau State Polytechnic, Barkin Ladi, Nigeria
[2]Department of Metallurgical Engineering, Plateau State Polytechnic, Barkin Ladi, Nigeria

Abstract As a construction material, concrete is employed to resist compressive stresses. Compressive strength of concrete is one of the most important and useful properties in structures. The classical method of testing the compressive strength by crushing specimens in the compressive testing machine to obtain the compressive strength is quite expensive, destructive, and time-consuming. This study uncovers the possibility of using machine learning techniques on huge dataset of earlier experiments to estimate the concrete compressive strengths with very good results. Linear Regression, K-Nearest Neighbor, Decision Trees and Random Forests were used in the validation and performances. Random Forest Regressor got the lowest root mean squared error of 5.45 and the highest R^2 value of 89%, hence the best algorithm for this estimation. With this approach, the estimation of the compressive strength is done almost instantaneously, and inexpensive. This research opens up the possibility of deriving an inverse regression procedure for cost optimization of the materials when given a target Concrete Compressive Strength.

Keywords Concrete Compressive Strength, Machine Learning, Alternative concrete mixture, Concrete Cost Optimization, Exploratory Data Analysis

1. Introduction

Concrete mixtures can be designed to provide a wide range of mechanical and durability properties to meet the design requirements of a structure. The compressive strength of concrete is the most common performance attribute used by the engineer when designing structures. The strengths of structures in engineering are linked to the inherent strength of the materials used. The use of substandard materials, particularly low-quality concrete, has been identified in literature as the leading cause of structure collapse [1]. It is therefore eminent to know the concrete compressive strength of the mixture at all costs.

It is quite expensive and time consuming, though worthwhile, to genuinely obtain the approximate concrete compressive strength in constructions. The steps for a complete compressive strength test include making the specimen at the site, site-curing the specimen, transporting the specimen to the laboratory, curing the specimen in the laboratory for at least 28 days, preparing the ends of the specimen for test, and finally loading the specimen to failure to determine the compressive strength [2]. If any of these steps are performed incorrectly, the result is incorrect and misleading.

Cement is one of the construction materials widely used around the world in order to develop infrastructure, and the production of this cement consumes a lot of raw materials like limestone, which releases CO_2 into the atmosphere thus leads to global warming. Studies are pointing to alternative material mixtures of cement, blast furnace slag, fly ash, water, superplasticizer, coarse aggregate, fine aggregate, and curing processes (Age) to yield even better concrete compressive strength in infrastructure [3]. The computation of the concrete compressive strength in a multivariate objective problem requires accurate and precise intelligence. With advances in technology, especially in the area of Artificial Intelligence (AI) and powerful multiprocessing technologies, many such complex problems with complicated tasks that require human-like intelligence and intuition could be more accurately and precisely solved by intelligent artifacts [4-5].

Concrete producers and construction companies are interested in improving the sustainability of concrete, including reducing its CO_2 emissions and the costs of materials while maintaining its mechanical properties, workability, and durability. There is every need to establish the concrete compressive strength in the construction of all infrastructures. But running the concrete compressive strength test takes time and a huge cost [6]. From the big data acquired from records of previous experiments, engineers can use the analysis to project the concrete compressive strengths of designs and implementations of structures. More so, proper analysis of concrete compressive strength can help in the optimal estimation of the costs of construction of the infrastructure.

In this study, the researcher aims at developing an Artificial Intelligent System that estimates the concrete compressive strength of any material mixture of cement, blast furnace slag, fly ash, water, superplasticizer, coarse aggregate, fine aggregate, and curing processes (Age).

1.1. Contribution

The study will indicate the following contribution to knowledge: -

Correlations between the concrete compressive strength and the component composition of the concrete.

The organization of this project is as follows. Section 1 covers the introduction in the view of statement of the problem, aim of the study, and the contributions this study proffers. Section 2 overviews the various contributions from other researchers on Concrete Compressive Strength and Machine Learning. Section 3 shows the proposed layout of the solution pattern of the project. Section 4 presents the results of the simulations on implementation. Section 5 discusses the results and subsequently recommendations are made. For more readability, Table 1 provides a summary of the used abbreviations and their definitions.

Table 1: Used abbreviations and acronyms

Abbreviations	Definitions
AI	Artificial Intelligence
ML	Machine Learning
SVM	Support Vector Machine
ANN	Artificial Neural Network
MPa	Megapascal
MSE	Mean Squared Error
CSV	Comma-Separated Values
RMSE	Root Mean Squared Error
MAE	Mean Absolute Error
UPV	Ultrasonic Pulse Velocity
RH	Rebound Hammer
NS	Nano-Silica
MS	Micro-Silica

2. Review of Related Literature

Compressive strength of concrete is essential for concrete structure utilization and is the main feature of its safety and durability. Concrete mix design is a complex and multistage process in which we try to find the best composition of ingredients to create good performing concrete that is economical. Because the high cost and time-consuming process of experimental determination of compressive strength of concrete, machine learning is gaining significant attention and future estimations for this technology are even more promising [7]. And employing machine learning approaches instead of traditional experimental models that requires a high degree of engineering expertise makes it possible to develop better understanding of the compressive strength of concrete [8].

In a comparative analysis of concrete compressive strength using techniques of Decision Tree model, Random Forest model, and Artificial Neural Network, it is inferred that the Artificial Neural Network model predicts with high accuracy for compressive strength of concrete [9].

Many studies have evaluated the effects of additives such as nano-silica (NS), micro-silica (MS) and polymer fibers on optimizing the mechanical properties of concrete, such as compressive strength. Nowadays, with progress in cement industry, it has become possible to produce cement type I with strength classes of 32.5, 42.5, and 52.5 MPa. The microstructure of cement has changed, and modified by NS, MS, and polymers; therefore, it is very important to determine the optimal percentage of each additives for those cement strength classes [10].

According to Ju [11], estimating the concrete compressive strength of high strength concrete (HSC) is an essential investigation for the maintenance of nuclear power plant (NPP) structures. This involves evaluating the compressive strength of HSC using two approaches: non-destructive tests and concrete core strength. For non-destructive tests, samples of HSC were mixed to a specified design strength of 40, 60 and 100 MPa. Based on a dual regression relation between ultrasonic pulse velocity (UPV)and rebound hammer (RH) measurements, an estimation expression is developed, and this can be subsequently used to estimate other combinations.

Cost optimization of concrete structures is the starkest quality in the design of infrastructures. Mahzuz [12] suggested that the designs of concrete structures should be based on cost minimization rather than weight minimization. It was concluded that there is a real demand to perform research on cost optimization, especially for large structures with many members where optimization can result in significant savings. Reduction in the cost of construction is a major desire of most construction companies [13-14]. According to Madhusudanan & Amirtham [15], building materials constitute about 60%-70% of the total cost of construction, and so reducing these costs of building materials will significantly reduce the total cost of construction. The introduction of industrial wastes such as Copper slag, Phospogypsum and Fly ash, when used as supplements for sand and coarse aggregate in building materials, reduces the cost of construction considerably. Our focus is the cost optimization of the mix ratio of the building materials [16], [17].

3. Materials and Methods

3.1. Nature of the Dataset

In this study, the dataset for analysis of obtained in raw form (not scaled) from UCL Machine Learning Repository. https://archive.ics.uci.edu/ml/datasets/Concrete+Compressive+Strength. The Exploratory Data Analysis of the data is considered at the design stage in Section 3 of the project.

Concrete Compressive Strength Dataset

Data Type: multivariate

Concrete is the most important material in civil engineering. The concrete compressive strength is a highly nonlinear function of age and ingredients. These ingredients include cement, blast furnace slag, fly ash, water, superplasticizer, coarse aggregate, and fine aggregate.

Data Characteristics:

The actual concrete compressive strength (MPa) for a given mixture under a specific age (days) was determined from laboratory. Data is in raw form (not scaled).

Summary Statistics:

Number of instances (observations): **1030**

Number of Attributes: **9**

Attribute breakdown: **8** quantitative input variables, and **1** quantitative output variable

Missing Attribute Values: None

Variable Information:

Given is the variable name, variable type, the measurement unit and a brief description. The concrete compressive strength is the regression problem. The order of this listing corresponds to the order of numerals along the rows of the database. Table 2 presents information about the variables as captured in Comma Separated Value (CSV).

Table 2: Variable descriptions

Name	Data type	Measurement Unit	Description
Cement (component 1)	Quantitative	kg in a m^3 mixture	Input Variable
Blast Furnace Slag (component 2)	Quantitative	kg in a m^3 mixture	Input Variable
Fly Ash (component 3)	Quantitative	kg in a m^3 mixture	Input Variable
Water (component 4)	Quantitative	kg in a m^3 mixture	Input Variable
Superplasticizer (component 5)	Quantitative	kg in a m^3 mixture	Input Variable
Coarse Aggregate (component 6)	Quantitative	kg in a m^3 mixture	Input Variable
Fine Aggregate (component 7)	Quantitative	kg in a m^3 mixture	Input Variable
Age	Quantitative	Day (1~365)	Input Variable
Concrete compressive strength	Quantitative	MPa	Output Variable

3.2. Performance Metrics

Among the various metrics used to evaluate the results of the prediction which include Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), R^2 (Coefficient of Determination), and Adjusted R^2; Coefficient of Determination or R^2 best fits the analysis of compressive strength as the metric helps us to compare our current model with a constant baseline from the dataset. R^2 is given by: -

$$R^2 = 1 - \frac{\mathrm{MSE(model)}}{\mathrm{MSE(baseline)}}$$

Where MSE is the Mean Squared Error.

3.3. Procedure

The approach in this study is to use Machine Learning principles of Exploratory Data Analysis, Data Processing, Model Building, Training, Validating, and Testing, and then Prediction.

3.3.1. Exploratory Data Analysis

Visualizing the dataset speaks volumes about the dataset, and also suggests correlations between features and points out missing values in the dataset. As depicted in fig. 1, the probability distribution is largely symmetric and so it does not bring certain philosophical complexity to the very process of estimating a typical-value for the distribution. A summary of the variables in the compressive strength experiments is shown in fig. 2. Each variable (feature) relates with one another as indicated in the fig. 3. In order to gain insight into the dataset and its underlying structure, statistical data description is carried out as shown in fig. 4.

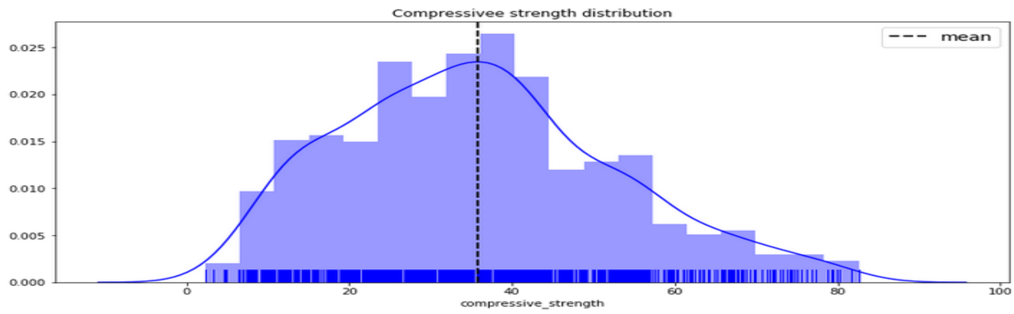

Figure 1: Distribution of Concrete Compressive Strength

Figure 2: Variables summary

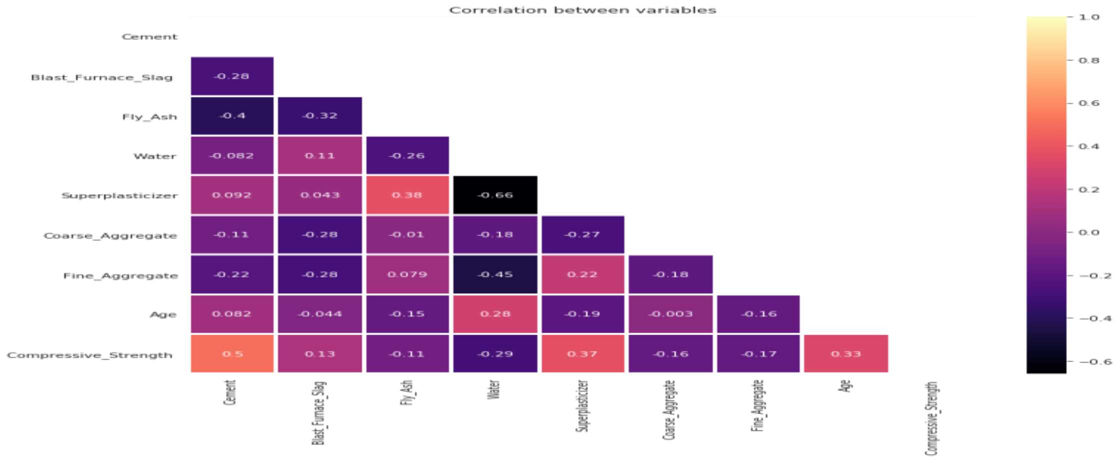

Figure 3: Correlations between variables

```
df.describe()
```

Out[22]:

	Cement	Blast_Furnace_Slag	Fly_Ash	Water	Superplasticizer	Coarse_Aggregate	Fine_Aggregate	Age	Compressive_Strength
count	1030.000000	1030.000000	1030.000000	1030.000000	1030.000000	1030.000000	1030.000000	1030.000000	1030.000000
mean	281.167864	73.895825	54.188350	181.567282	6.204660	972.918932	773.580485	45.662136	35.817961
std	104.506364	86.279342	63.997004	21.354219	5.973841	77.753954	80.175980	63.169912	16.705742
min	102.000000	0.000000	0.000000	121.800000	0.000000	801.000000	594.000000	1.000000	2.330000
25%	192.375000	0.000000	0.000000	164.900000	0.000000	932.000000	730.950000	7.000000	23.710000
50%	272.900000	22.000000	0.000000	185.000000	6.400000	968.000000	779.500000	28.000000	34.445000
75%	350.000000	142.950000	118.300000	192.000000	10.200000	1029.400000	824.000000	56.000000	46.135000
max	540.000000	359.400000	200.100000	247.000000	32.200000	1145.000000	992.600000	365.000000	82.600000

In [25]:

Figure 4: Statistical Data Description

4. Results & Discussion

In this aspect of processing, efforts are put into transforming the raw data into dataset which can be processed by algorithmic training. This involves generating the 80% training set (Input training and their subsequent training labels), the 10% validation sets (Input validation set and their validation labels), and the 10% test set (input testing set and their actual labels). In other words, 80% of the dataset is for training, and 20% for holdout set (Validation and test sets). In summary, there are a total of six (6) variables from the dataset/array which will be used. They are: -

- X_train 8 input features, 80% of full dataset for training of the model
- y_train 1 output feature, 80% of full dataset for training of the model
- X_val 8 input features, 10% of full dataset for validation of the model
- y_val 1 output feature, 10% of full dataset for validation of the model
- X_test 8 input features, 10% of full dataset for testing of the model
- y_test 1 output feature, 10% of full dataset for testing of the model

4.1. Model Building, Training, Validating, and Testing

Machine Learning consists of two basic steps. The first step is to specify a template (an architecture), and the second step is to find the best numbers (optimal values) to fill in the template. From this point, the code here will also follow the same steps. These incremental learning algorithms (Support Vector Regressor, k-Nearest Neighbors, Decision Tree, Random Forest Regressor, and Artificial Neural Network) were built, trained, validated and tested accordingly. The results of these model building, training, validating, and testing with Linear regression model, K-Nearest Neighbors, Decision Trees model, and Random Forest regressor model are respectively shown in fig. 5, fig. 6, fig. 7 and fig. 8.

4.1.1. Building, Training, Validating, and Testing Linear Regression Model

The approaches to solutions of Linear or Multiple Linear Regressions are basically the same. A regression model involving multiple variables can be represented as: -

$$y = b_0 + m_1b_1 + m_2b_2 + m_3b_3 + \ldots + m_nb_n.$$

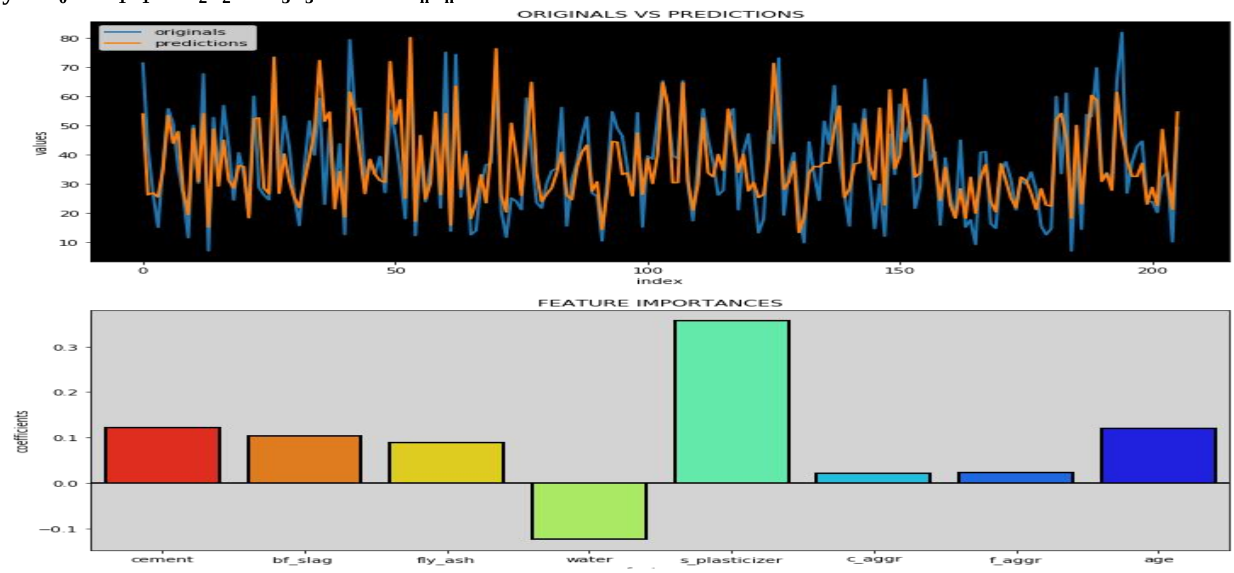

Figure 5: Comparison between Real and predicted Compressive Strengths Using Linear Regression Model

4.1.2. Building, Training, Validating, and Testing K-Nearest Neighbors Model

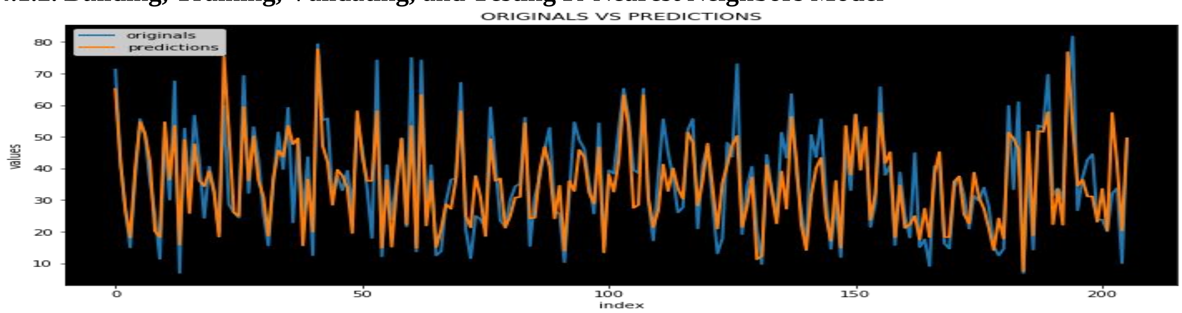

Figure 6: Comparison between Real and predicted Compressive Strengths Using K-Nearest Neighbors Model

4.1.3. Building, Training, Validating, and Testing Decision Tree Model

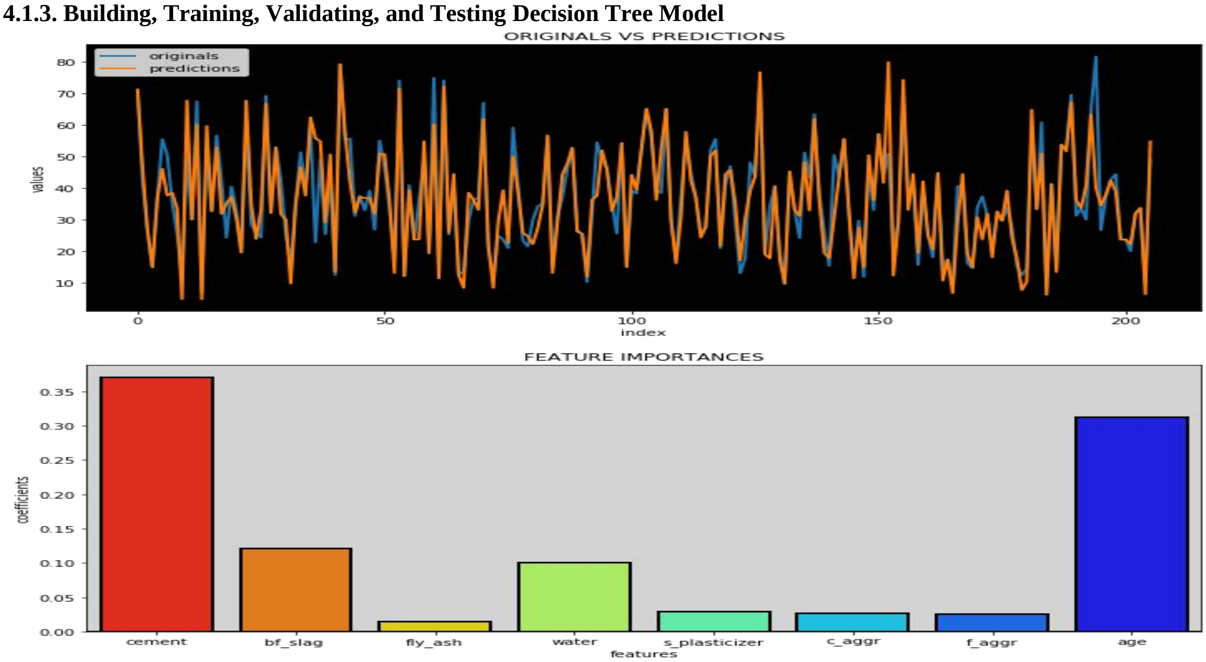

Figure 7: Comparison between Real and predicted Compressive Strengths Using Decision Tree Model

4.1.4. Building, Training, Validating, and TestingRandom Forest Regressor Model

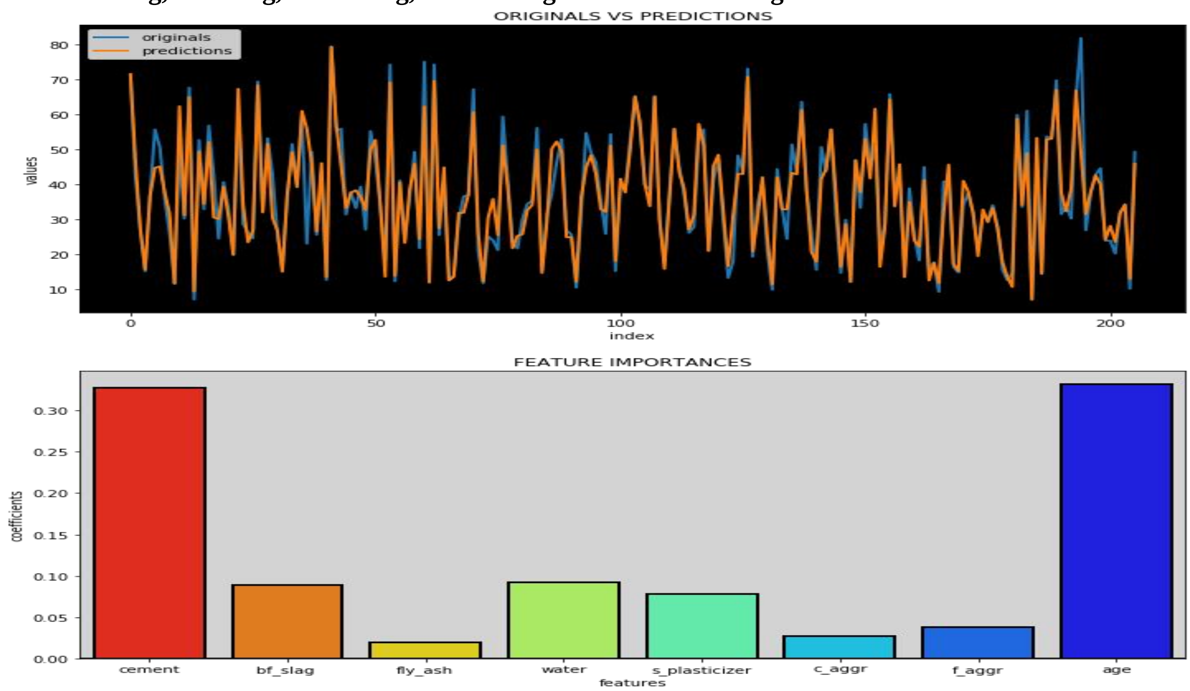

Figure 8: Comparison between Real and predicted Compressive Strengths Using Random Forest Regressor Model

4.2. Evaluations of the Algorithms Used

Comparisons of the four algorithms (Linear Regression, K-Nearest Neighbors, Decision Tree, and Random Forest) used in the analysis of the Concrete Compressive Strength dataset can be plotted and tabulated to show at a glance the best option for future computations as shown in fig. 9, and summarized in Table 3.

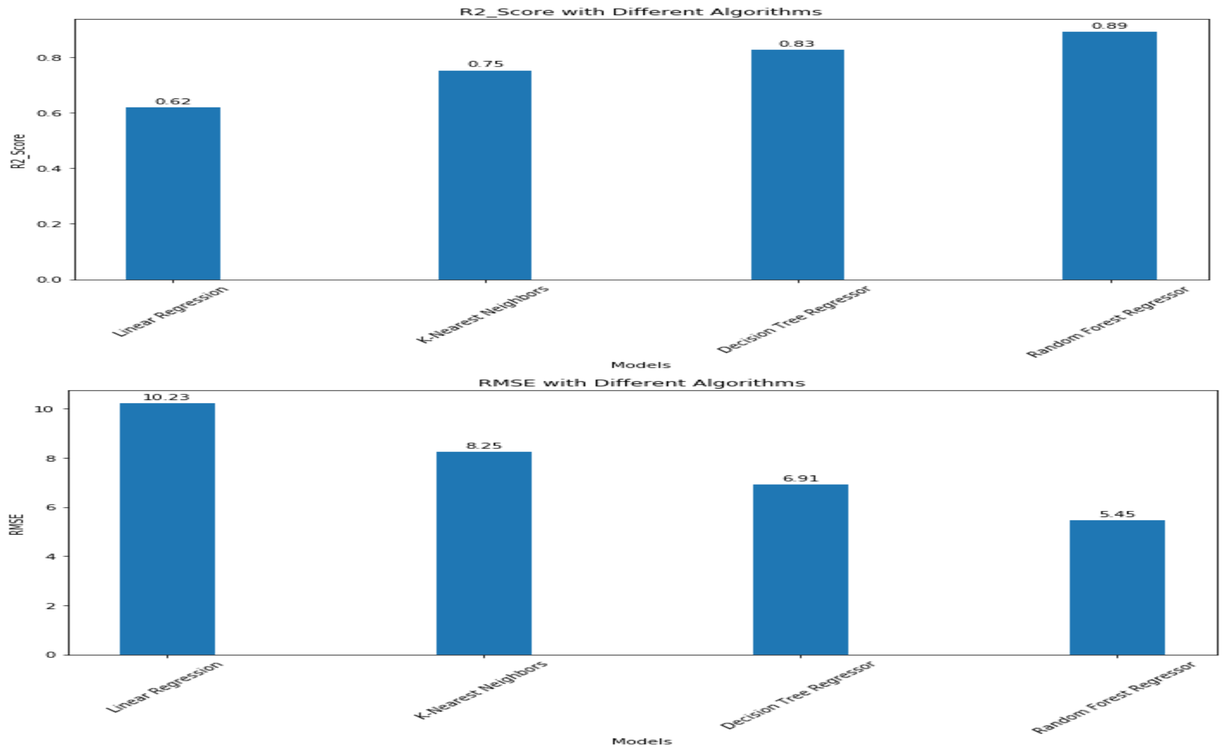

Figure 9: RMSE and R^2values of the Models used

Table 3: Summary of Models Performances

	RMSE	R^2
Linear Regression Model	10.23	0.62
K-Nearest Neighbors	8.25	0.75
Decision Tree Regressor	6.91	0.83
Random Forest Regressor	5.45	0.89

5. Conclusion and Recommendation

Putting all the results together, it is obvious that the four algorithms used perform good at predicting the Concrete Compressive Strength of structures at various error values. We processed the Compressive Strength Data and used Machine Learning approach to estimate the Compressive Strength of Concrete. We have used Multiple Linear Regression, K-Nearest Neighbor, Decision Trees and Random Forests to make predictions and compared their performances. Random Forest Regressor has the lowest RMSE of 5.45 and the highest R^2 value of 89% is the best choice out of the four algorithms for this problem. This research opens up the possibility of deriving an inverse regression procedure for cost optimization of the materials when given a target Concrete Compressive Strength.

Acknowledgment

We appreciate Dr. Sunday Iliya and Mr. Joel Gogwim for their constructive criticisms.

References

[1]. Yeshaswini, N. S. P., & Kameswara, B. R. (2019). Method of mix design for concrete by using GGBS and Silica Fume. International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-7, Issue-6C2, April 2019.

[2]. Wang, XY (2019). Optimal design of the cement, fly ash, and slag mixture in ternary blended concrete based on gene expression programming and the genetic algorithm. Multidisciplinary Digital Publishing Institute (MDPI): Materials (Basel). 2019 Aug: 12(15): 2448

[3]. Bellum, R. R., Nerella, R., Madduru, S. R. C., & Indukuri, C. S. R. (2019). Mix design and mechanical properties of fly ash and GGBFS-synthesized alkali-activated concrete. Infrastructures 2019, 3, 20; doi:10.3390/infrastructures4020020.

[4]. Davenport, T., Guha, A., Grewal, D., & Bressgott, T. (2020). How artificial intelligence will change the future of marketing. *Journal of the Acad. Mark. Sci.***48,** 24–42 (2020). https://doi.org/10.1007/s11747-019-00696-0

[5]. Al-Turjman, F., &Baali, I. (2019). Machine learning for wearable IoT-based applications: A survey. *Transactions on Emerging Telecommunications Technologies, e3635.* doi:10.1002/ett.3635

[6]. Hyun-Oh, S., Doo-Yeol, Y., Joo-Ha, L., Seung-Hoon, L., & Young-Soo, Y. (2019). Optimized mix design for 180 MPa ultra-high-strength concrete. Journal of Materials Research and Technology Volume 8, Issue 5, September–October 2019, Pg 4182-4197. https://doi.org/10.1016/j.jmrt.2019.07.027

[7]. Ziolkowski, P., & Niedostatkiewicz, M. (2019). Machine learning techniques in concrete mix design. *Materials, 12(8), 1256.* doi:10.3390/ma12081256

[8]. Silva, P. F. S., & Moita, G. F. (2019). A comparative study of machine learning methods for compressive strength of concrete. *Proceedings of the 4th World Congress on Civil, Structural, and Environmental Engineering (CSEE'19) Rome, Italy –April 7-9, 2019Paper No. ICSECT136* DOI: 10.11159/icsect19.136

[9]. Chopra, P., Sharma, R. K., Kumar, M., & Chopra, T. (2018). Comparison of machine learning techniques for the prediction of compressive strength of concrete. *Advances in Civil Engineering, 2018, 1–9.* doi:10.1155/2018/5481705

[10]. Zahiri, F., &Eskandari-Naddaf, H. (2019). Optimizing the compressive strength of concrete containing micro-silica, nano-silica, and polypropylene fibers using extreme vertices mixture design. *Front. Struct. Civ. Eng.*13, 821–830. https://doi.org/10.1007/s11709-019-0518-6

[11]. Ju, M., Park, K., & Oh, H. (2017). Estimation of compressive strength of high strength concrete using non-destructive technique and concrete core strength. *Applied Sciences, 7(12), 1249.* doi:10.3390/app7121249

[12]. Mahzuz, H.M.A., Choudhury, M.R., Ahmed, A.R. & Ray, S. (2019). Effect of material strength on the cost of RCC building frames. *Springer Nature Switzerland Applied Sciences* 2, 46 (2020). https://doi.org/10.1007/s42452-019-1830-4 pp2

[13]. Collantes, R., Portilla, J., Rodríguez, S. and García, F. (2020). Time and cost reduction in the production chain of construction projects through collaborative computer systems, 2020 *Congreso Internacional de Innovación y Tendenciasen Ingeniería (CONIITI), Bogota, Colombia, 2020, pp. 1-6,*doi: 10.1109/CONIITI51147.2020.9240406.

[14]. Gapuz, E. O., & Ongpeng, J. M. C. (2017). Optimizing compressed earth blocks mix design incorporating rice straw and cement using artificial neural network. *2017IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM), Manila, Philippines, 2017,* pp. 1-6, doi: 10.1109/HNICEM.2017.8269450.

[15]. Madhusudanan, S., & Amirtham, L. R. (2016). Optimization of construction cost using industrial wastes in alternative building material for walls. *Key Engineering Materials, 692, 1–8.*https://doi.org/10.4028/www.scientific.net/kem.692.1

[16]. Patil, R. & Rajakumara, H. N. (2018). An overview on optimization of concrete mix design. *International Journal of Pure and Applied Mathematics Vol 120 No. 6 2018, 6647-6663 ISSN: 1314-3395 (on-line version)* url: http://www.acadpubl.eu/hub/Special Issue.

[17]. Biswas, S. and Biswas, S. (2017). Inverse linear regression in machine learning. *Global Journal of Computer Science and Technology: Information & Technology. Volume 17 Issue 3 Version 1.0 Year 2017 Online ISSN: 0975-4172 & Print ISSN: 0975-4350.* https://core.ac.uk/download/pdf/231159469.pdf