Artificial Intelligence for Compressive Strength Prediction of Concrete

S. J. Goutham¹ and V. P. Singh²

¹National Institute of Technology Kurukshetra, Haryana, India.
²National Institute of Technology Kurukshetra, Haryana, India.
*Email: gsjsai@gmail.com

Abstract. Structural health monitoring is an indispensable procedure that is to be carried out to evaluate the serviceability of existing structures. Non-destructive testing methods are attaining increasing popularity for the assessment of concrete strength due to the ease of operation and reliability of the results. The application of machine learning in the field of engineering has increased rampantly. In this study, the compressive strength of concrete has been forecasted using support vector regression which is a machine learning technique. Artificial intelligence is nothing but the potential to impersonate human intelligence. The advantage of artificial intelligence over human intelligence is the absence of human errors that can arise due to various factors which might decrease the accuracy of results. Rebound hammer (RBH) , Windsor probe penetration (WPP) and ultra-sonic pulse velocity(USV) are the non-destructive testing that has been employed to assess the concrete compressive strength. The use of multiple non-destructive testing methods than single testing methods has improved the accuracy of prediction model. Further the results of different combination models were compared using coefficient of determination which is a commonly used statistical parameter for prediction accuracy comparison. The prediction model is in accordance with the experimental results obtained. And the accuracy of prediction indicates that support vector regression can be successfully used for predicting the compressive strength of concrete.

1. Introduction

Engineering is occluded with hurdles that cannot be easily solved using ordinary typical computational methods. Artificial intelligence is the solution to such problems. It is a computational method that strives to replicate human comprehension capacity and creates outputs that challenges traditional normal computational techniques. Cognitive computing is another bough of artificial intelligence which was influenced by human mind’s potential. The problem solving nature of cognitive system and human minds are similar viz. thinking and reasoning. The main attributes of such systems are their capacity to decode large data, training dynamically and learning adaptively which helps in establishing a consistent pattern. The use of artificial intelligence in the field of research has grown rapidly in the past decade.

The compressive strength of concrete at 28 days is a decisive parameter in RCC. The conventional procedure to procure the concrete mix design is empirical and performance based. Unambiguous evaluation of concrete strength at 28 days is advisable for better control of quality. But when it comes to the strength prediction of concrete in actual structures, the process is not so simple due to the irregular environmental conditions and other variations that can transpire during mixing, placing, compaction and finishing. Prediction of concrete strength is a multifaceted problem. Hence in-order to contemplate all the uncertainties involved, it is desirable to go for a non-linear mathematical model. The use of machine learning for concrete compressive strength prediction has
been acclaimed widely due to the uncontrollable variations involved in designing of concrete mix. One of the popular machine learning algorithms for concrete strength prediction is artificial neural network (ANN). Though ANNs have given satisfactory values of prediction, it has the tendency of data overfitting for datasets that are smaller. Such an overfitting can lead to powerful performance of the training set but a substandard performance in generalization on new dataset [1]. Support vector regression (SVR) is yet another machine learning tool that has been extensively used for strength prediction. The shortcomings in prediction using ANN has been resolved in SVR. Existing structures have to be evaluated for their structural performance in order to assess its seismic performance and for its retrofitting after a damage or failure. For this evaluation, the mechanical properties of concrete have to be assessed. Estimation through coring has its limitations due to technical difficulty and expense. Non-destructive testing provides a feasibility alternative for assessment of in-situ strength of concrete. Relationship which is empirical in nature is identified between the strength assessed on the concrete core and the NDT result. For quality in estimation of strength, the empirical relationship identified must be error free and valid for any input dataset. Presence of uncontrollable parameters, type of NDT, quality of measurement is few of the many parameters which influences the reliability of estimation [2]. In this study, concrete cubes of various strength were casted and three NDT viz. RBH, WPP and USV were employed on these cubes. The results of NDT were used to generate a prediction model using SVR in python which accurately predict the compressive strength of concrete using these NDT results. NDT results were used individually as well as in combination to evaluate the performance of these NDT techniques. Coefficient of determination (R²), root mean squared error (RMSE), mean absolute percentage error (MAPE) and mean absolute deviation (MAD) were the statistical parameters identified to derive a comparison between various models.

2. Literature Review

Advanced data mining techniques have a huge scope in the field of engineering for problem solving. Important data mining techniques include prediction model (machine learning implementing supervised learning), association and clustering (unsupervised learning) and so on. Problems associated with civil engineering are diverse and complex. It can be attributed to extremely variable factors. In order to address these issues, machine learning algorithms have been developed that can provide satisfactory solution to these complex problems which will be difficult otherwise to solve using conventional techniques. The process of fine tuning the parameters associated with the algorithm is a salient process. A prediction model is nothing but an algorithm that establishes a distinct relationship with the input and output parameters and which can successfully be used for the prediction for further data. For instance, forecasting the properties of any material in a crucial function in material science. Due to the high learning capabilities of support vector regression, it has found a prime place in the field of data prediction [3]. Neural networks and SVR models has been generated to predict the strength of no-slump concrete and although both the models gave sufficient prediction accuracy, SVR model is preferred over ANN due to its various superior features [4]. Research has also verified the use of SVR for computing the strength of high performance concrete. Experimental results have yielded low prediction errors for the same. Such a prediction model can be used efficiently to identify concrete mixtures of substandard strength[5]. NDT methods have evolved as a response to the detection of structural damage and its prevention. Development in sensors and new materials and reduction in size of devices are improving the scope of developing new NDT methods. Assessment of material integrity using NDT is attaining popularity[6]. From literature survey, it has been found out that fuzzy logic can also be used for generating prediction models to forecast the compressive strength of cement mortar. When the fuzzy logic model was compared to that of ANN, it was discovered that the former model could predict data with a lower errors and also the model generated was more user friendly [7]. Machine learning is just not a knowledge generation tool. It can also be efficiently used for analysis of data. It is very crucial to perceive the exact nature of problem. It can be better studied by analysing previous applications. A comprehensive understanding of machine learning process can be used for incorporating the features of the problem into the techniques of machine learning. To achieve this there is constant need to update with the recent trends and findings in the field of machine learning. Machine learning techniques can be enhanced by carrying out
comparison studies. Evaluation of any kind involving single or multiple techniques of machine learning which is qualitative or quantitative requires special attention in each step of interpretation, design and analysis. Without this, the output can be faulty and might be influenced by random variables [8].

The application of machine learning is structural health monitoring is highly beneficial and constructive. Health monitoring of structures involves collection of data through devices, uprooting the damage prone characteristics and interpreting the extracted data. The studies in the area of monitoring of structures can be broadly seen as data and model drive. The later adopts a numerical model and the former requires the functioning of machine learning and pattern recognition. It is interesting to note that, data driven approach is more efficient and useful if data involved is huge, the physical traits of the structure is unspecified and difficult to model and there is a need to reduce the effort in computation. In the field of structural health monitoring, the functions of machine learning includes creation of database from existing data, understanding the parameters of the model and then solely focussing on prediction of data. Within the scope of machine learning, in terms of robustness and effectiveness, support vector regression, artificial neural network and principal component analysis are few of the techniques which have attained popularity for identification of structural damage. The properties of concrete is known to have highly complex non-linear relationships between its constituents [9]. Hence there is a need to develop a reliable model for better understanding of the characteristics of the material in such a way that it saves both time and cost [10]. Machine learning algorithms could successfully model the traits of concrete which is self- compacting. Support vector regression has also been used successfully for forecasting the split tensile strength of concrete. RBF and polynomial kernel functions were used for the accurate prediction of split tensile strength. The model generated is very efficient because it can execute regression of non-linear nature for data set of higher dimensions [11]. SVM based models have also been generated to establish the relationship between results obtained from NDT and the strength of concrete. The service life assessment and durability can be predicted using machine learning technique with high level of accuracy. The complicated relationship between various parameters can be learned by machine learning algorithms and forecast its remaining service life with a high degree of reliability [12].

3. Support Vector Regression in Python

Support vector machines were first discovered by Vapnik and Guyon. Along with the strong basis in statistical learning theory, SVM exhibited great accomplishment in various fields. Along with other soft computing methodologies, SVM has also established itself as a most sought after machine learning technique. It falls under the category of controlled learning technique. Kernel functions which are non-linear in nature is used for classification problem for transforming input data to data of higher dimension. Compared to neural network, the development path of SVM is straight opposite. Though initially SVMs were not widely popular despite having a powerful theoretical background, later it gained momentum and proved itself better than other machine learning techniques [13]. Support vector regression (SVR) which is used for regression problems is a variant of support vector machines that is mainly used for classification problems. The model generated in SVR is controlled by a dataset subset and is termed as support vectors. These support vectors generates errors greater than threshold value during training of SVR model. Support vector regression is in-demand technique that has produced reliable results in the field of data prediction.

The algorithms in SVR employ functions that are mathematical in nature which is referred to as kernels. The purpose of such kernel functions is to absorb the provided data as input and mutate into the desired form. Different algorithms use divergent kernel functions. Linear, radial basis function (RBF), polynomial, sigmoid and non-linear are some of the un-built kernel functions in support vector regression. Among the various kernel functions available, RBF kernel is the most sought after kernel due to the better output it delivers and the overall reduction in RMSE. For the prediction developed, RBF kernel function itself has been used. It was decided on the basis of trial and error. The major objective of adaptive regression methods like support vector regression is to establish underlying structure in the dataset. For a defined set of data for training, the objective is to attain a function that has the highest variation from the absolute target for all the data falling under training and
simultaneously have the maximum flat which means the errors are neglected as long as they are within the specified limits. Flat refers to LIBSVM is a module used for SVR. Its standard use involves training a dataset to obtain a model that has successfully identified the distinct complex relationship between input and output data. The next step involves predicting the testing dataset. Various extensions of LIBSVM is also available. LIBSVM helps in achieving the target of regression, classification and distribution [14]. SVR can be implemented in various softwares. In the present study, it has been executed in python programming language. Python has established itself as one of the most favoured language for scientific computing. Its popularity can be attributed to its interactive user interface and the vast scientific libraries it possess. Scikit-learn utilize this computing friendly interface to contribute various machine learning algorithms. Hence provides a support for analysing the data by amateurs in the field of machine learning and artificial intelligence. This facilitates the use of artificial intelligence in various branches of engineering. Due to its liberal license policy, python is available as open source software package. Both supervised and unsupervised learning algorithms are provided under scikit-learn [15]. Python’s syntax is sophisticated yet simple lets the programmer write in a procedural manner. The interpreter of python is highly interactive which enables live development of code and its simultaneous experimentation. This reduces the time required for code development and also increases the productivity of the programmer. The library modules available in python are plenty. In addition to the inbuilt library, we can download other library functions also and expand the library collection as per our need. The popularity of python as a computational tool can be attributed to its simple and easily understandable syntax. Blocks of code identified through indentation, looping structures that are easy to comprehend are few examples of python’s uncomplicated syntax. Apart from the lucid syntax of python, python promotes codes that are easily maintainable by dividing the codes into modules [16]. While training a model in python, two parameters are considered viz. gamma and C. The proper choice of gamma and C is crucial for desirable performance of SVR. For different combinations of NDT, different values of gamma and C have been used. It has been arrived on the basis of trial and error method. Table 1 lists out the combination of C, gamma and kernel function for different NDT combination.

| NDT Technique | Kernel | C   | Gamma  |
|---------------|--------|-----|--------|
| USPV          | rbf    | 9800| 0.23   |
| PP            | rbf    | 9800| 0.23   |
| RH            | rbf    | 17000| 0.067 |
| USPV-PP       | rbf    | 930 | 0.00032|
| PP-RH         | rbf    | 930 | 0.00032|
| RH-USPV       | rbf    | 930 | 0.00032|
| USPV-PP-RH    | rbf    | 1320| 0.0011 |

4. Experimental Work

350 concrete cubes were casted as per the recommendation of IS 10262:2009. Moulds of size 15x15x15 cm were used. Major steps involved in the cube preparation was sieving of coarse and fine aggregates, uniform mixing of the raw materials, placing of concrete, compaction and curing. The fractions of raw materials were varied to achieve samples of various strengths. It was done in order to have a dataset possessing wide range of concrete strength. The minimum and the maximum compressive strength obtained from dataset of 350 concrete cubes were 5 N/mm² and 55 N/mm² respectively. Initially non-destructive testing was carried out on the cubes followed by subjecting the specimen to compressive loading upto failure in compression testing machine (CTM). Schmidt RBH was used to determine the hardness of concrete cubes by measuring the RBH number. Pundit USV instrument was used to record the velocity of propagation (km/s) of pulses through the test specimen. The velocity of propagation indicates the quality of concrete. Finally Windsor penetrometer was employed to find the depth of penetration of the probe. The depth of penetration indicates the strength of in-situ concrete.
5. Results & Discussion

After the completion of experiment, we have four types of data available viz. depth of penetration (mm), pulse velocity (km/s), RBH number which has been obtained by performing NDT on the concrete cubes. The compressive strength of concrete is also available which has been found out by subjecting the cubes to compressive load upto failure in compression testing machine (CTM). The dataset associated with the SVR model is classified into input and output. The NDT results form the input data set and the compressive strength of concrete constitutes the output dataset. $R^2$ formed the basis of comparison of various models generated. The range of $R^2$ is 0-1 where 1 indicates 100% accuracy in prediction. Initially prediction models using single NDT results were generated. Later, 2 NDT input combination prediction models were produced. 3 input prediction models were also created. Along with SVR, predictions using ANN on single NDT techniques were also carried out [17].

From the results, it was concluded that the accuracy of prediction of SVR is higher than that of ANN. Table 2 shows the value of $R^2$ for ANN and SVR.

| NDT Technique | $R^2$ of ANN | $R^2$ of SVR |
|---------------|-------------|-------------|
| RBH           | 0.857       | 0.911       |
| WPP           | 0.786       | 0.8299      |
| USV           | 0.468       | 0.769       |

Figure 1: Graph depicting prediction results using 1 NDT technique.

Figure 1 shows the prediction result when single NDT result is used as input for the SVR model. From the graph we can observe the points are scattered from away prediction curve for USV and that of RBH is nearest. This implies that the accuracy of prediction of RBH is the best among the 3 NDT employed. The prediction result when 2 NDT results are used in combination for the SVR model is showed in figure 2. It can be observed that the extent of scatter is maximum for WPP-ultrasonic pulse velocity combination and minimum for RBH-WPP combination making it the most reliable prediction model combination when 2 NDT results are used in combination for concrete compressive strength prediction.
Figure 2: Graph depicting prediction results using 2 NDT technique combinations.

Figure 3: Graph depicting prediction results using 3 NDT technique combinations.

Figure 3 shows the plot between the actual and the predicted compressive strength of concrete when 3 NDT results are used as input for the SVR prediction model. Ideally when the input for the prediction model is more, higher accuracy must be obtained. But such a pattern is not observed here because of lower quality of results observed in USV technique. Table 3 illustrates the statistical parameters that were defined for various prediction models.

Table 3: Statistical parameters of various NDT techniques

| NDT Technique | R²   | MAD  | RMSE | MAPE  |
|---------------|------|------|------|-------|
| USV           | 0.769| 5.117| 7.291| 14.862|
| RBH           | 0.911| 3.518| 4.284| 10.084|
| WPP           | 0.8299| 4.42 | 6.063| 12.529|
| USV-WPP       | 0.832| 4.314| 5.121| 14.586|
| WPP-RBH       | 0.91  | 3.605| 4.382| 9.891 |
| USV-RBH       | 0.896| 3.735| 4.487| 10.489|
| USV-WPP-RBH   | 0.899| 3.36 | 4.309| 8.488 |

From observing the results, it can be concluded that the use of USV is not feasible due to inferior accuracy of prediction. To have a superior accuracy of prediction using USV, we should have a larger input data set of USV.
6. Conclusion

In the study carried out, a prediction model was generated to predict the compressive strength of concrete using support vector regression which is a machine learning technique. The prediction modelling was executed in python language. It has been found out that RBH produces the best prediction output in comparison to other NDT techniques. The values of MAPE, MAD, RMSE and $R^2$ also support the findings. Furthermore the output of SVR was compared with ANN and SVR was found to be superior. Hence it can be used for accurate prediction.

Acknowledgement

The authors thank National Institute of Technology Kurukshetra for providing technical support for successfully conducting the research work.

References

[1] DeRousseau M A, Laftchiev E, Kasprzyk J R, Rajagopalan B and Srubar W V 2019 A comparison of machine learning methods for predicting the compressive strength of field placed concrete Constr. Build. Mater. 228, 116661
[2] Ali-Benyahia K, Sbartaï Z M, Breyssse D, Kenai S and Ghrici M 2017 Analysis of the single and combined non-destructive test approaches for on-site concrete strength assessment: General statements based on a real case-study Case Stud. Constr. Mater. 6, 109–19
[3] Chou J S and Pham A D 2015 Smart Artificial Firefly Colony Algorithm-Based Support Vector Regression for Enhanced Forecasting in Civil Engineering Comput. Civ. Infrastruct. Eng. 30, 715–32
[4] Sobhani J, Khanzadi M and Movahedian A H 2013 Support vector machine for prediction of the compressive strength of no-slump concrete Comput. Concr. 11, 337–50
[5] Pham A D, Hoang N D and Nguyen Q T 2016 Predicting Compressive Strength of High-Performance Concrete Using Metaheuristic-Optimized Least Squares Support Vector Regression J. Comput. Civ. Eng. 30, 28–31
[6] Helal J, Sofi M and Mendis P 2015 Non-destructive testing of concrete: A review of methods Electron. J. Struct. Eng. 14, 97–105
[7] Akkurt S, Tayfur G and Can S 2004 Fuzzy logic model for the prediction of cement compressive strength Concr. Res. 34, 1429–33
[8] Reich Y 1997 Machine learning techniques for civil engineering problems Comput. Civ. Infrastruct. Eng. 12, 295–310
[9] Demir F and Armagan Korkmaz K 2008 Prediction of lower and upper bounds of elastic modulus of high strength concrete Constr. Build. Mater. 22, 1385–93
[10] Salehi H and Burgueño R 2018 Emerging artificial intelligence methods in structural engineering Eng. Struct. 171, 170–89.

[11] Yan K, Xu H, Shen G and Liu P 2013 Prediction of splitting tensile strength from cylinder compressive strength of concrete by support vector machine Adv. Mater. Sci. Eng. 2013.
[12] Taffese W and Sistonen E 2017 Machine learning for durability and service-life assessment of reinforced concrete structures: Recent advances and future directions Autom. Constr. 77, 1–14.
[13] Çevik A, KURTOĞLU A E, Bilgehan M, Gülşan M E and Albegnprl H M 2015 Support vector machines in structural engineering: A review J. Civ. Eng. Manag. 21, 261–81
[14] Chang C and Lin C 2013 LIBSVM : A Library for Support Vector Machines 1–39
[15] Pedregosa F, Grisel O, Weiss R, Passos A and Brunner M 2011 Scikit-learn: Machine Learning in Python 12, 2825–30
[16] Oliphant T E 2007 Python for scientific computing Comput. Sci. Eng. 9, 10–20
[17] Singh V P and Kotiyal Y C 2013 Prediction of compressive strength of concrete using artificial neural networks Res. J. Pharm. Technol. 10, 35–40