Concrete Classification Using Machine Learning Techniques

Narayanan Natarajan[1](✉), Murugesan Shyam Sundar[1], Mohanadhas Berlin[2], and Mangottiri Vasudevan[3]

[1] Department of Civil Engineering, Dr. Mahalingam College of Engineering and Technology, Pollachi 642003, Tamil Nadu, India
itsrajan2002@yahoo.co.in

[2] Department of Civil Engineering, National Institute of Technology, Arunachal Pradesh, Yupia, Papum Pare 791112, Arunachal Pradesh, India

[3] Department of Civil Engineering, Bannari Amman Institute of Technology, Sathyamangalam 638401, Tamil Nadu, India

Abstract. Present study aims at evaluating the performance of versatile machine learning techniques such as Support Vector Machine, Bagged trees, Boosted trees and RUSBoosted trees in classification of the strength of concrete blocks into three classes: low, medium and high. Ten-fold cross validation was performed on the dataset to classify the concrete strength. The algorithms were compared in terms of ROC (area under the curve) and accuracy of classification. The bagging ensemble method outperformed all other methods. Among the SVM methods, the cubic SVM performed better when compared to quadratic and medium Gaussian methods. This study will be highly beneficial in classification of concrete under limited input dataset without going through the drudgery of performing the physical tests on the concrete blocks.

Keywords: Machine learning · Concrete · SVM · Classification · Compressive strength · Bagging algorithm · Boosting algorithm

1 Introduction

The combination of aggregate, cement, and water is a composite construction material which is popularly known as concrete. It is a commonly used manufactured item worldwide for construction purpose which has varied properties (Lomborg 2001). Depending on the need of the work the strength and appearance of concrete will be attained by the selection of concrete mix and the same has been decided by local legislation and building codes. The main considerations while designing any concrete mix layout are (i) the required strength, and (ii) the exposed weather condition while in service (Castelli et al. 2013). Particularly for concrete, the compressive strength is deliberated as one of the most significant characteristics remarkable for designing the proper mix of ingredients as well as for ensuring quality control. Even more, forecasting the strength of concrete is crucial in any concrete-based building construction. In order to envisage

J. A. Fonseca de Oliveira Correia et al. (Eds.): ASMA 2021, STIN 19, pp. 179–193, 2022.
https://doi.org/10.1007/978-3-030-98335-2_12

the required strength of concrete, many strategies are attempted by various researchers. Abram's exponential mathematical model is a well-known methodology to examine the concrete strength. It is additionally useful for concrete mix design and also quality assurance of concrete. This model connects the volumetric ratio of water to cement, to assess the strength of concrete in a typical 28 days period for testing (Gupta et al. 2006). The continuous concrete strength improvement can be checked using maturity meter and this examination gives a reliable outcome on keeping track the strength of concrete (Luke et al. 2002). Further, researchers have suggested different techniques for forecasting the concrete strength (Hwang et al. 2004; Kim et al. 2004). The strength-porosity design suggested by Balshin and gel-space ratio equations proposed by Power are widely employed to approximate the compressive strength of foam concrete in its various forms such as paste, mortar and regular mix (Nambiar and Ramamurthy 2008). Nonetheless, these techniques have considered only a few criteria such as maturation of concrete as well as concrete mix proportions. Normally, these standard approaches have actually been developed with restricted data.

As one of the versatile predictive tools in machine learning, Artificial Neural Network (ANN) excels in several engineering applications. Researchers have accomplished a sensible level of success making use of ANN in civil engineering applications. ANNs applications have been anticipate to predict the settlement of superficial foundations in geotechnical problems (Shigidi et al. 2003), structural engineering (Rogers et al. 1994), as well as forecast of concrete strength with mix design (Kasperkiewicz et al. 1995; Oh et al. 1999). Concrete strength prediction has been obtained by Lai and Serra (1997) using neural networks. Guang and Zong (2000) proposed multi-layer feed-forward neural network (MFNN) for its application in predicting compressive strength of concrete based on 28-day test results. Sanad and Saka (2001) have extended the application of neural networks to deep beams with reinforced concrete for predicting the ultimate shear strength. Lee (2003) has introduced another smart application called I-PreConS (Intelligent PREdiction system of Concrete Strength) using ANN. Kim et al. (2005) utilized probabilistic neural networks for a range of concrete mix proportions to forecast the compressive strength. Gupta et al. (2006) have introduced a neural expert system to forecast the concrete strength which takes inputs such as proportioning of concrete, adopted methodology and duration for curing, physical dimensions of the specimen (size and shape), and the prevailing environmental conditions (relative humidity, temperature, wind speed etc.). The compressive strength data from ultrasonic pulse velocity measurements has been analysed with multiple regression analysis as well as artificial neural networks (Kewalramani and Gupta 2006). ANN has been efficiently used in modelling the correlation between these velocity measurements and strength of concrete (Topcu et al. 2008). Sobhani et al. (Sobhani et al. 2010) have reported that the prediction of 28-day-based compressive strength with neural network models and ANFIS models were a lot more possible than traditional regression models.

Chou et al. (2011) have compared various data-mining techniques such as ANN, support vector machines (SVM), multiple regression, multiple additive regression trees and bagging regression trees to optimize the prediction accuracy of compressive strength of concrete. These approaches were also found to be effective for predicting concrete strength when prepared with foreign materials such as fly ash (Yu et al. 2018) and silica fume (Ozcan et al. 2009) as additives as well as for recycled aggregate concrete (Duan et al. 2013). Further, the elastic modulus of different types of concrete (normal as well as high strength) was also successfully predicted by ANN (Demir 2008) and fuzzy logic (Demir 2005). ANN has been further extended to apply in mineral admixture concrete to predict the strength (Atici 2011). The result exposes that prediction of compressive strength using ANN is better than multiple regression analysis. However, Chou et al. (2011) found that multiple additive regression trees performed better in compressive strength estimation especially when the variations in admixture composition are high.

The research has been further diversified into development of more specific machine learning techniques for the prediction of compressive strength of concrete. A genetic programming (GP) method with geometric semantic genetic operators was found to be ideal for prediction of compressive strength of high performance concrete (Castelli et al. 2013). Yuan et al. (2014) have used hybrid models with genetic based algorithms and ANFIS for the same purpose. The combination of artificial neural network and evolutionary search protocols such as genetic algorithms, together called evolutionary artificial neural networks (EANNs), are successfully implemented by Nikoo et al. (2015) and Chopra et al. (2016) for predicting the strength of concrete. The recent trends in employing hybrid approach for predicting structural performance of special types of concrete can be summarised as: optimized self-learning method for high performance concrete (Yu et al. 2018); variable analytical approach (a combination of ANN and GP) for soil-fly ash geo-polymer (Leong et al. 2018); support vector machine for cement-based materials exposed to sulphate attack (Chen et al. 2018); ANN and SVM for bentonite/sepiolite concrete (Ghanizadeh et al. 2018); ANN for concrete based on agricultural and building construction wastes (Getahun et al. 2018); relevance vector machine and emotional neural network for concrete (Biswas et al. 2019); hybrid ultrasonic-neural prediction method for eco-friendly concrete screeds with high volume of waste quartz mineral dust (Sadowski et al. 2019); and hybrid machine learning model with standalone models for normal concrete (Cook et al. 2019). Though it is evident that several studies have been conducted on prediction of compressive strength of concrete using various machine learning techniques, but none of the studies have explored the possibility of classification of concrete based on the predicted compressive strength.

The determination of the compressive strength of concrete is a time taking cumbersome process. If we can build a classifier which can classify the concrete into low, medium or high strength only on the basis of selected input parameters, it will be very helpful for the construction engineers to choose the correct type of concrete for the intended strength requirements without actually testing the concrete which can save huge amount of time, energy and material resources. In this study, an attempt has been made for the first time to apply the classification algorithm on the compressive strength of concrete to attribute as low, medium and high strength. The main focus of this study is to determine the best classification algorithm which can group the concrete based on the various input parameters such as presence of coarse aggregate, fine aggregate such as cement, additives such as blast furnace slag, super-plasticizer and water; time of concrete testing and manual classifiers such as low, medium and high. It should be noted that the machine learning experiments were carried out only for the classified data and not the actual compressive strength data.

2 Machine Learning Methods

The dataset containing 1005 records of compressive strength of concrete was collected from the UCI repository. Based on the range of the reported compressive strength values, it was sorted into three classes as low (<20 MPa), medium (20 to 50 MPa) and high (>50 MPa). After the manual classification, 196 records belonged to low strength, 616 records belonged to medium strength and 193 belonged to high strength concrete. A brief description of the selected machine learning methods provides closer insights for comparing their performance.

2.1 Support Vector Machine

Support Vector Machines is a learning method for prediction and classification of data where a dataset point is conceived as an n-dimensional vector (Chou et al. 2011). The basic concept of SVM is to determine a hyperplane that divides the n-dimensional dataset perfectly into two classes. A brief description of the SVM classifiers used in this study is given below in Table 1.

Table 1. Description of SVM classifiers

Classifier type	Prediction speed	Memory usage	Interpretability	Model flexibility
Quadratic SVM	Binary: Fast Multiclass: Slow	Binary: Medium Multiclass: Large	Hard	Medium
Cubic SVM	Binary: Fast Multiclass: Slow	Binary: Medium Multiclass: Large	Hard	Medium
Medium Gaussian SVM	Binary: Fast Multiclass: Slow	Binary: Medium Multiclass: Large	Hard	Medium

2.2 Bagging Algorithm

The bagging algorithm comprises of classifiers produced by various bootstrap (statistical) samples (Rogers 1994). Given a standard training set, the bootstrap sample can be generated by sampling m instances uniformly with replacement (Oh et al. 1999). After generating a series of T bootstrap samples (B_1, B_2… B_T), individual classifier function (C_i) can be built from each B_i. This can be further evaluated in combination (C_1, C_2 …C_T) to identify the final classifier (C^*) which is used to define the output class in terms of its sub-classifiers.

2.3 Boosting Algorithm

Unlike the Bagging algorithm which defines the classifiers parallel, the adaptive boosting (AdaBoost) algorithm sequentially produces classifiers. The AdaBoost algorithm essentially tweaks the subsequent weak learners by modifying the input weights of the training instances, thus adapting in favour of the instances where the data were misclassified by the previously constructed classifiers. This enforces the inducer to converge to a strong learner by reducing much of the expected and repeating errors among the various input distributions. The construction of AdaBoost algorithm consists of producing a series of T weighted training sets ($S_1, S_2, \ldots S_T$) and T classifiers ($C_1, C_2, C_{3\ldots..}C_T$) in sequence where T is an integer indicating the number of trials. The final classifier (C^*) is defined as an output of a weighted voting scheme where a certain weight is assigned to each sample at every iteration. Since the assigned weight of each classifier indicates the associated error at that stage, the construction of the decision tree can be made in favour of set of samples with higher weights which indicates better performance.

2.4 RUSboosted Method

RUSBoost is a combination of random under-sampling (RUS) and AdaBoost algorithm which can handle class imbalance problems in discrete datasets by removing samples from majority class. At the end of each iteration, the weights are reassigned based on the estimated errors, and the best class will be identified based on weighted majority. In order to accurately model the minority classes, the model is allowed to under-sample the over-represented sample during each iteration.

3 Results and Discussion

In this study, we have employed SVM, and Boosting and Bagging classifiers to classify the concrete according to its compressive strength as low, medium and high. The outcomes from each classifier analysis were verified using a tenfold cross-validation technique.

Quadratic SVM

The quadratic SVM method has achieved 86.4% accuracy in classifying the concrete strength into low, medium and high. It is observed in Fig. 1 that 136 out of the 193 high strength samples have been accurately classified. Similarly, 165 out of the 196 low strength samples, and 567 out of the 626 medium strength samples have been accurately classified. Figure 2 shows that 96% of the area is under the curve.

Cubic SVM

The cubic SVM method has achieved 87.8% accuracy in classifying the concrete strength into low, medium and high. It is observed in Fig. 3 that 154 out of the 193 high strength samples have been accurately classified. Similarly, 169 out of the 196 low

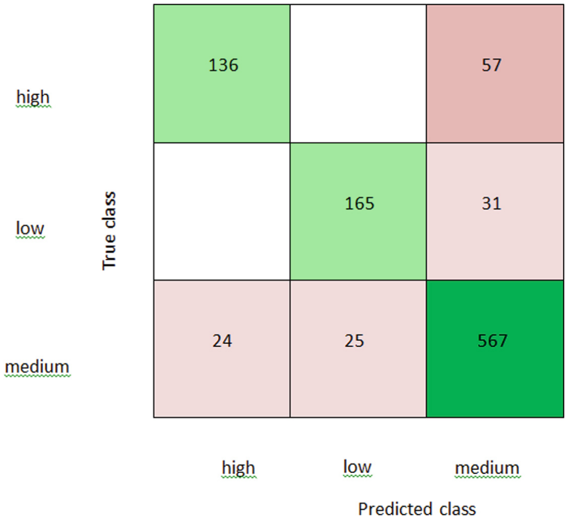

Fig. 1. Confusion matrix obtained using quadratic SVM.

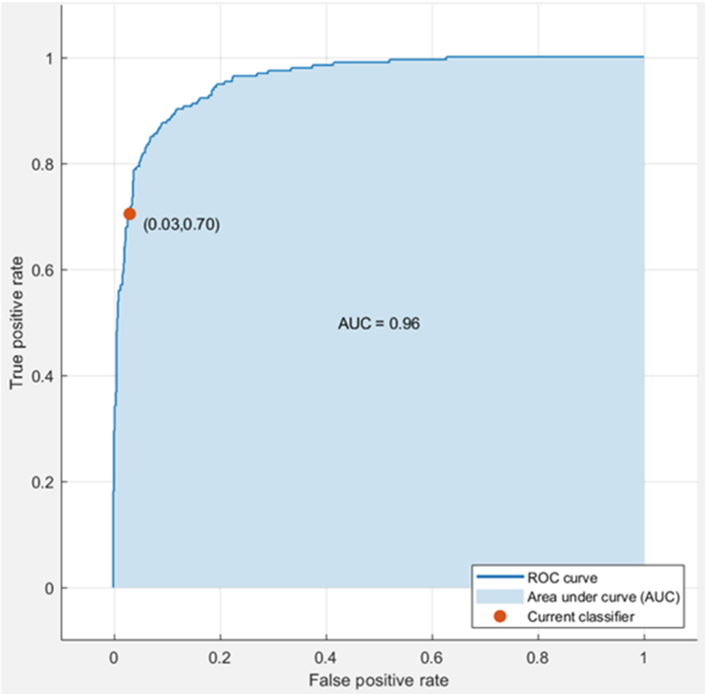

Fig. 2. Receiver operating characteristic curve (ROC) obtained using quadratic SVM.

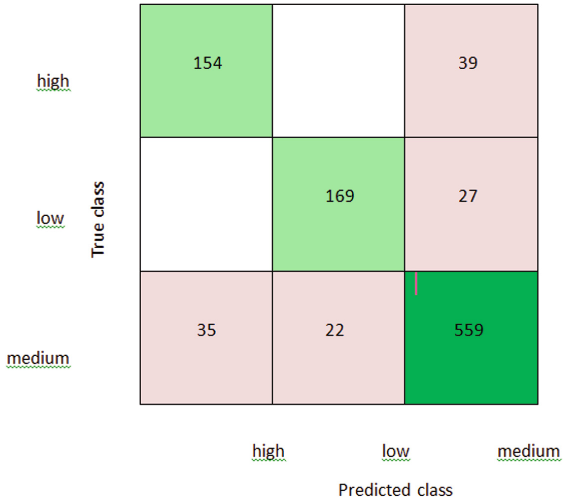

Fig. 3. Confusion matrix obtained using cubic SVM.

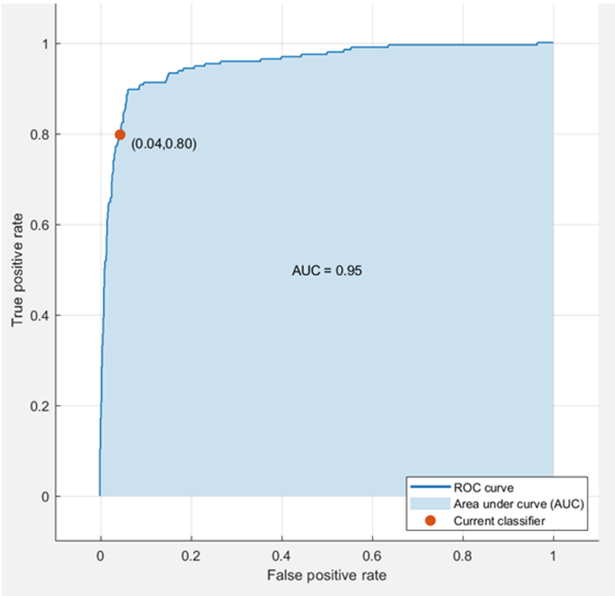

Fig. 4. Receiver operating characteristic curve (ROC) obtained using cubic SVM.

strength samples, and 559 out of the 626 medium strength samples have been accurately classified. It is interesting to note that cubic SVM has classified high strength and low strength concrete samples more accurately compared to the medium strength concrete samples. Figure 4 shows that the area under the curve is 95%. Thus cubic SVM has outperformed quadratic SVM.

Medium Gaussian SVM

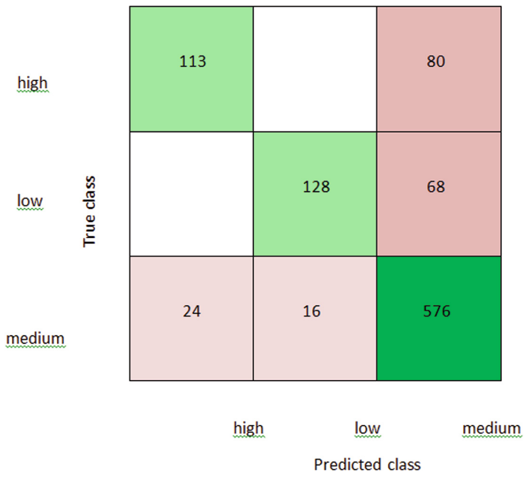

Fig. 5. Confusion matrix obtained using medium Gaussian SVM.

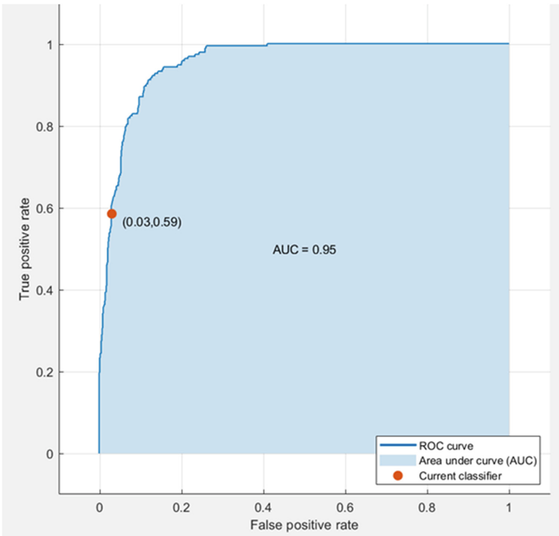

Fig. 6. Receiver operating characteristic curve (ROC) obtained using medium Gaussian SVM.

The accuracy of this method is only 81.3% which is much lower compared to the other two SVM methods. The same is evident from Fig. 5 where the number of samples classified as low, medium and high is lower in comparison with quadratic and cubic SVM. Also, the ROC shows that only 95% of the area is under the curve (Fig. 6). Thus, it can be concluded that among the SVM methods, cubic SVM performance is the best.

Bagging Method

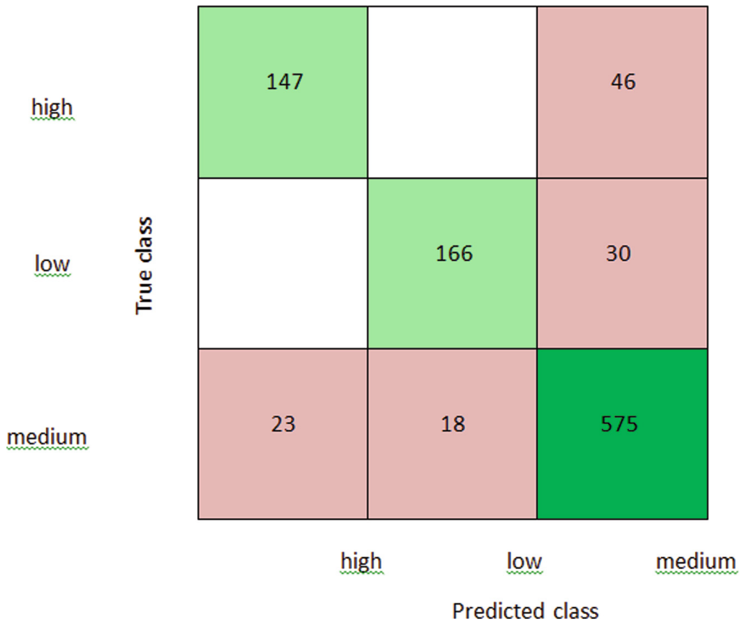

Fig. 7. Confusion matrix obtained using bagging method.

The accuracy of this method is 88.4% which is marginally higher when compared to cubic SVM method. It is observed in Fig. 7 that 147 out of the 193 high strength samples have been accurately classified. Similarly, 166 out of the 196 low strength samples, and 575 out of the 626 medium strength samples have been accurately classified. Figure 8 shows that the area under the curve is 97%. This method has achieved the highest accuracy of 88.4% when compared to other classification methods.

Boosting Method

The accuracy of this method is only 86.7% which is marginally lower when compared to bagging method. It is observed in Fig. 9 that 142 out of the 193 high strength samples have been accurately classified. Similarly, 157 out of the 196 low strength samples, and 572 out of the 626 medium strength samples have been accurately classified. Figure 10

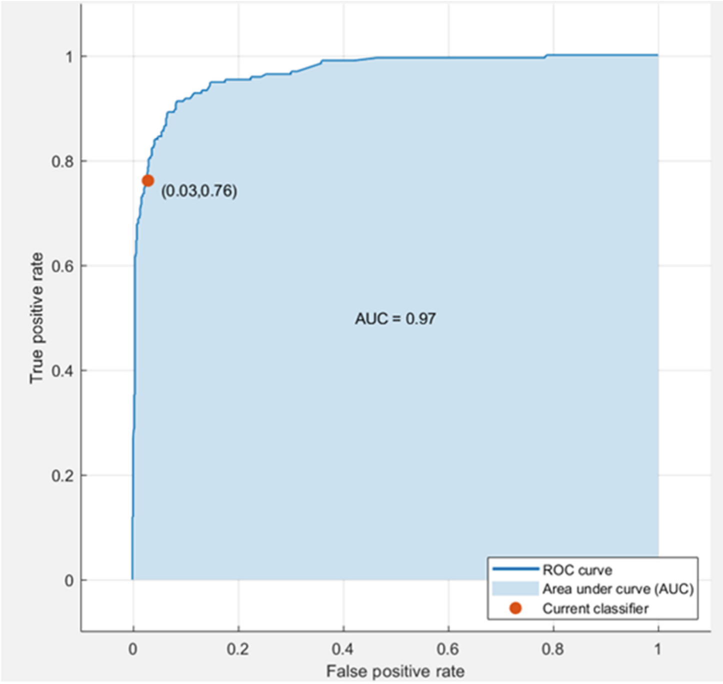

Fig. 8. Receiver operating characteristic curve (ROC) obtained using bagging method.

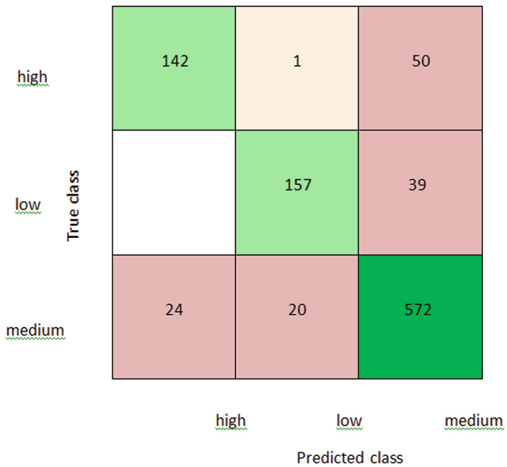

Fig. 9. Confusion matrix obtained using boosting method.

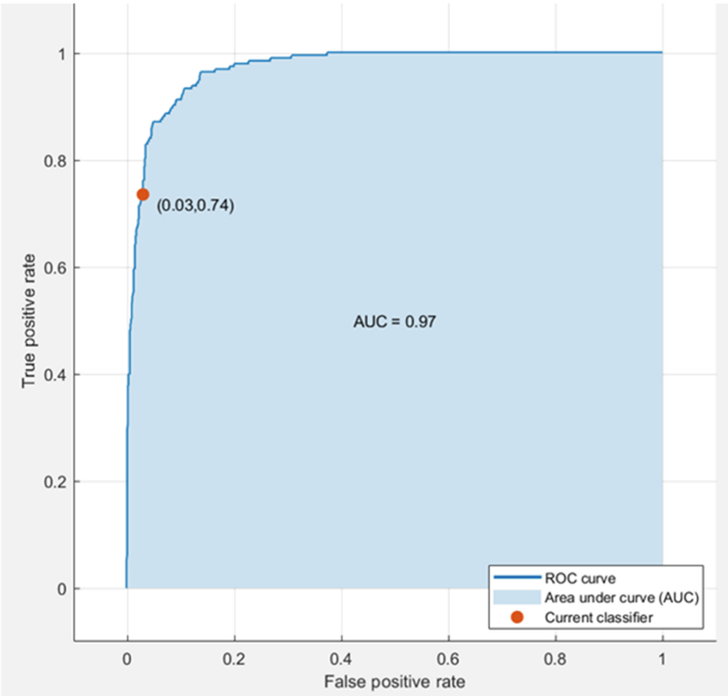

Fig. 10. Receiver operating characteristic curve (ROC) obtained using boosting method.

shows that the area under the curve is 97%. Although the ROC curve shows that 97% of the area is under the curve the accuracy is lower when compared to bagging method.

RUSBoosted Method

This method shows an accuracy of 81.5% which is lowest among all the methods used in this study to classify the concrete as low, medium and high strength. From Fig. 11, it can be observed that 160 out of the 193 high strength samples have been accurately classified. Similarly, 169 out of the 196 low strength samples, and 490 out of the 626 medium strength samples have been accurately classified. Figure 12 shows that the area under the curve is 97%. The ROC curve shows that only 95% of the area is under the curve, which is very much similar to the medium Gaussian ROC curve.

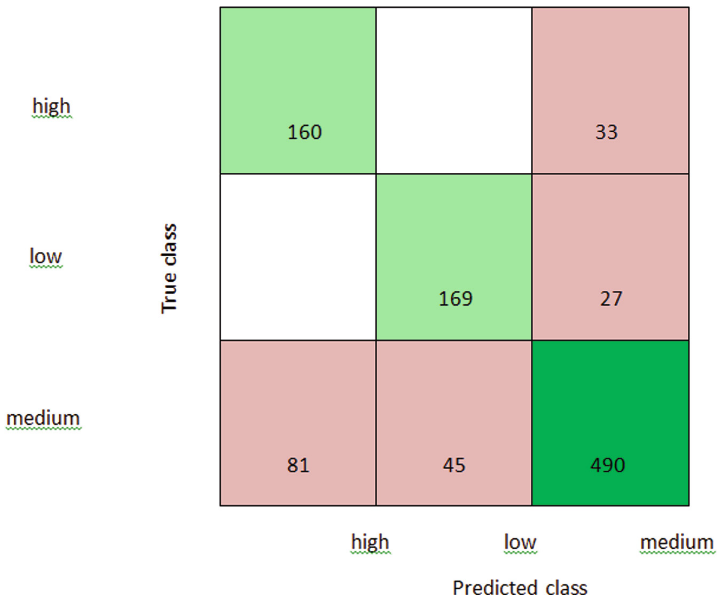

Fig. 11. Confusion matrix obtained using RUSboosted method.

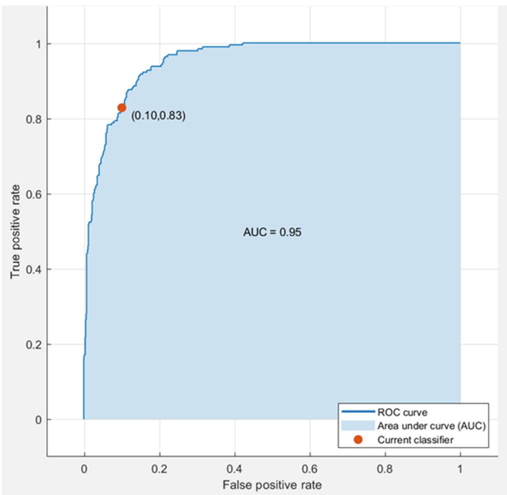

Fig. 12. Receiver operating characteristic curve (ROC) obtained using RUSboosted method.

The comparison of the classification of the different methods adopted in this study is shown in Table 2 below.

Table 2. Accuracy of classification of various methods

S. No.	Classification method	Accuracy (%)
1	Quadratic SVM	86.4
2	Cubic SVM	87.8
3	Medium Gaussian SVM	81.3
4	Bagging method	88.4
5	Boosting method	86.7
6	RUSboosted method	81.5

It can be concluded from Table 2 that among the SVM methods, the performance of Cubic SVM is the best. Among the ensemble methods, bagging method has outperformed boosting and RUSboosted methods.

4 Conclusion

In this study, we classified the concrete as high, medium and low strength using various classification algorithms. The output from the algorithms was compared with each other using the classification accuracy. Based on the accuracy, the bagging ensemble method has outperformed all other classification methods. Among the SVM methods, the cubic SVM performs better than quadratic and medium Gaussian SVM methods. This classification study can be used to classify the concrete strength based on the input parameters without actually performing the laboratory experiment.

References

Atici, U.: Prediction of the strength of mineral admixture concrete using multivariable regression analysis and an artificial neural network. Expert Syst. Appl. **38**, 9609–9618 (2011)

Biswas, R., Samui, P., Rai, B.: Determination of compressive strength using relevance vector machine and emotional neural network. Asian J. Civ. Eng. **20**(8), 1109–1118 (2019). https://doi.org/10.1007/s42107-019-00171-9

Cook, R., Lapeyre, R., 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. **31**(11), 04019255 (2019)

Chou, J.S., Chiu, C.K., Farfoura, M., Taharwa, I.A.: Optimizing the prediction accuracy of concrete compressive strength based on a comparison of data-mining techniques. J. Comput. Civ. Eng. **25**, 242–253 (2011)

Castelli, M., Vanneschi, L., Silva, S.: Prediction of high performance concrete strength using genetic programming with geometric semantic genetic operators. Expert Syst. Appl. **40**, 6856–6862 (2013)

Chen, H., Qian, C., Liang, C., Kang, W.: An approach for predicting the compressive strength of cement-based materials exposed to sulfate attack. PLoS ONE **13**(1), e0191370 (2018)
Chopra, P., Sharma, R.K., Kumar, M.: Prediction of compressive strength of concrete using artificial neural network and genetic programming. Adv. Mater. Sci. Eng. **2016**, 7648467 (2016)
Demir, F.: Prediction of elastic modulus of normal and high strength concrete by artificial neural networks. Constr. Build. Mater. **22**, 1428–1435 (2008)
Demir, F.: A new way of prediction elastic modulus of normal and high strength concrete—fuzzy logic. Cem. Concr. Res. **35**, 1531–1538 (2005)
Duan, Z.H., Kou, S.C., Poon, C.S.: Prediction of compressive strength of recycled aggregate concrete using artificial neural networks. Constr. Build. Mater. **40**, 1200–1206 (2013)
Gupta, R., Kewalramani, M.A., Goel, A.: Prediction of concrete strength using neural-expert system. J. Mater. Civ. Eng. **18**, 462–466 (2006)
Guang, N.H., Zong, W.J.: Prediction of compressive strength of concrete by neural networks. Cem. Concr. Res. **30**, 1245–1250 (2000)
Getahun, M.A., Shitote, S.M., Gariy, Z.C.A.: Artificial neural network based modelling approach for strength prediction of concrete incorporating agricultural and construction wastes. Constr. Build. Mater. **190**, 517–525 (2018)
Ghanizadeh, A.R., Abbaslou, H., Amlashi, A.T., Alidoust, P.: Modeling of bentonite/sepiolite plastic concrete compressive strength using artificial neural network and support vector machine. Front. Struct. Civ. Eng. **13**(1), 215–239 (2018). https://doi.org/10.1007/s11709-018-0489-z
Hwang, K., Noguchi, T., Tomosawa, F.: Prediction model of compressive strength development of fly-ash concrete. Cem. Concr. Res. **34**, 2269–2276 (2004)
Kasperkiewicz, J., Racz, J., Dubrawski, A.: HPC strength prediction using artificial neural network. J. Comput. Civ. Eng. **9**(4), 279–284 (1995)
Kim, J.-I., Kim, D.K., Feng, M.Q., Yazdani, F.: Application of neural networks for estimation of concrete strength. J. Mater. Civ. Eng. **16**(3), 257–264 (2004)
Kim, D.K., Lee, J.J., Lee, J.H., Chan, S.K.: Application of probabilistic neural networks for prediction of concrete strength. J. Mater. Civ. Eng. **17**, 353–362 (2005)
Kewalramani, M.A., Gupta, R.: Concrete compressive strength prediction using ultrasonic pulse velocity through artificial neural networks. Autom. Constr. **15**, 374–379 (2006)
Lai, S., Serra, M.: Concrete strength prediction neural network. Constr. Build. Mater. **11**(2), 93–98 (1997)
Leong, H.Y., Ong, D.E.L., Sanjayan, J.G., Nazari, A., Kueh, S.M.: Effects of significant variables on compressive strength of soil-fly ash geopolymer: variable analytical approach based on neural networks and genetic programming. J. Mater. Civ. Eng. **30**(7), 04018129 (2018)
Lomborg, B.: The Skeptical Environmentalist: Measuring the Real State of the World. Cambridge University Press (2001)
Lee, S.C.: Prediction of concrete strength using artificial neural networks. Eng. Struct. **25**, 849–857 (2003)
Luke, A., Hsu, C.T.T., Punurai, S.: Implementation of concrete maturity meters. Department of Transportation Division of Research and Technology, and U.S. Department of Transportation Federal Highway Administration, N.J (2002)
Nambiar, E.K.K., Ramamurthy, K.: Models for strength prediction of foam concrete. Mater. Struct. **41**, 247–254 (2008)
Nikoo, M., Moghadam, F.T., Sadowski, A.: Prediction of concrete compressive strength by evolutionary artificial neural networks. Adv. Mater. Sci. Eng. **849126**, 1–8 (2015)
Ozcan, F., Atis, C.D., Karahan, O., Uncuoglu, E., Tanyildizi, H.: Comparison of artificial neural network and fuzzy logic models for prediction of long-term compressive strength of silica fume concrete. Adv. Eng. Softw. **40**, 856–863 (2009)
Oh, J.W., Lee, I.W., Kim, J.T., Lee, G.W.: Application of neural networks for proportioning of concrete mixes. ACI Mater. J. **96**(1), 61–67 (1999)

Rogers, J.L.: Simulating structural analysis with neural network. J. Comput. Civ. Eng. **8**(2), 252–265 (1994)

Sadowski, L., Mielnik, M.P., Widziszowski, T., Gardynik, A., Mackiewicz, S.: Hybrid ultrasonic-neural prediction of the compressive strength of environmentally friendly concrete screeds with high volume of waste quartz mineral dust. J. Clean. Prod. **212**, 727–740 (2019)

Shigidi, A., Garcia, L.A.: Parameter estimation in groundwater hydrology using artificial neural networks. J. Comput. Civ. Eng. **17**(4), 281–289 (2003)

Sanad, A., Saka, M.P.: Prediction of ultimate shear strength of reinforced concrete deep beams using neural networks. J. Struct. Eng. **127**(7), 818–828 (2001)

Sobhani, J., Najimi, M., Pourkhorshidi, A.R., Parhizkar, T.: Prediction of the compressive strength of no-slump concrete: a comparative study of regression, neural network and ANFIS models. Constr. Build. Mater. **24**, 709–718 (2010)

Topcu, I.B., Saridemir, M.: Prediction of compressive strength of concrete containing fly ash using artificial neural networks and fuzzy logic. Comput. Mater. Sci. **41**, 305–311 (2008)

Yu, Y., Li, W., Li, J., Nguyen, T.N.: A novel optimised self-learning method for compressive strength prediction of high performance concrete. Constr. Build. Mater. **184**, 229–247 (2018)

Yuan, Z., Wang, L.N., Ji, X.: Prediction of concrete compressive strength: research on hybrid models genetic based algorithms and ANFIS. Adv. Eng. Softw. **67**, 156–163 (2014)