



Developing a boosted decision tree regression prediction model as a sustainable tool for compressive strength of environmentally friendly concrete

Sarmad Dashti Latif¹

Received: 6 February 2021 / Accepted: 22 July 2021 / Published online: 29 July 2021

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract

One of the most significant parameters in concrete design is compressive strength. Time and money could be saved if the compressive strength of concrete is accurately measured. In this study, two machine learning models, namely, boosted decision tree regression (BDTR) and support vector machine (SVM), were developed to predict concrete compressive strength (CCS) using a complete dataset through the previous scientific studies. Eight concrete mixture parameters were used as the input dataset. Four statistical indices, namely the coefficient of determination (R^2) and root mean square error (RMSE), mean absolute error (MAE), and RMSE-Standard Deviation Ratio (RSR), were used to illustrate the efficiency of the proposed models. The results show that the BDTR model outperformed SVM model with the overall result of $R^2=0.86$ and RMSE=6.19 and MAE=4.91 and RSR=0.37, respectively. The results of this study suggest that the compressive strength of high-performance concrete (HPC) can be accurately calculated using the proposed BDTR model.

Keywords Concrete strength prediction · Machine learning model · Boosted decision tree regression (BDTR)

Introduction

In the construction industry, the sustainability of construction materials is now a critical issue. There has been a remarkable momentum in the last decade to use industrial by-products in the construction sector to achieve sustainability. Time and cost are the most critical considerations to be included in the preparation of each project during the construction project (Aprianti et al. 2015; Zhong and Wu 2015; Kylili and Fokaides 2017; Mirzahosseini et al. 2019; Latif et al. 2020c). In traditional and industrial constructions, conventional concrete and high-performance concrete (HPC) have been commonly used, respectively. The HPC is a composite material that has been used to withstand environmental conditions in the manufacture of high-strength

concrete used in bridges, tall buildings, tunnels, and pavement structures. In addition, in structural applications, the durability and workability of the HPC are essential considerations. A complex method has been used to develop the technological characteristics of concrete to achieve the necessary characteristics of the HPC (Kasperkiewicz et al. 1995; Chou et al. 2011; Zhong and Wille 2015; Gonzalez-Corominas and Etxeberria 2016; Hoang et al. 2016; Wang et al. 2019; Kaloop et al. 2020).

Nowadays, artificial intelligence has been widely used for prediction and other purposes in engineering fields (Borhana et al. 2020; Ehteram et al. 2020a; Latif et al. 2021b, c; Ehteram et al. 2020b, c; Lai et al. 2020; Latif et al. 2020a, b, 2021a; Najah et al. 2021; Parsaie et al. 2021; Jumin et al. 2021; Latif and Ahmed 2021). To measure the HPC and CCS and test the input variables, many of the previous studies used soft computing techniques (Gupta et al. 2006; Chopra et al. 2018; Dutta et al. 2018; Yaseen et al. 2018; Al-Shamiri et al. 2019, 2020; Vakharia and Gujar 2019; Young et al. 2019; Feng et al. 2020; Latif 2021). For instance, Abuodeh et al. (2020) employed two deep machine learning techniques, namely sequential feature

Responsible Editor: Philippe Garrigues

✉ Sarmad Dashti Latif
sarmad.latif@komar.edu.iq

¹ Civil Engineering Department, College of Engineering, Komar University of Science and Technology, Sulaimany 46001, Kurdistan Region, Iraq

selection (SFS) and neural interpretation diagram (NID), to classify the essential material constituents that affect the artificial neural network (ANN). Based on the material quantities, they used 110 ultra-high performance concrete (UHPC) compressive strength tests to train the ANN. As a result, four material components were chosen, primarily cement, fly ash, silica fume, and water, and then used in the ANN to determine more precise predictions than the model with all eight material components. Finally, based on the four chosen material constituents, they have developed a nonlinear regression model. Their findings show that the use of ANN with SFS and NID greatly enhanced the model's accuracy and offered useful insights into the predictions of ANN compressive intensity for various UHPC mixes.

Ling et al. (2019) optimized support vector machine (SVM) model by K-Fold cross validation for predicting and evaluating the degradation of CCS in a complicated marine environment. They also built ANN and decision tree (DT) to compare the prediction precision with the SVM model. Their results showed that the SVM model had the best prediction performance.

Furthermore, Shaqadan (2020) developed SVM and neural network models to predict the CCS using five input variables, including silica additive fraction. He used a 90 samples data set and measured compressive strength after 3 and 28 days for different levels of milling time. According to his result, both SVM and ANN showed a good correlation coefficient of 0.929 and 0.986, respectively.

Another study were proposed by Naderpour et al. (2018) to predict recycled aggregate concrete compressive strength using ANN. In their analysis, they used 139 existing sets of data derived from 14 published literature sources to create training and testing data for ANN model creation. Their findings suggest that the ANN is a useful method for predicting the compressive strength of recycled aggregate concrete, which is made up of various forms and sources of recycled aggregates.

Moreover, Deng et al. (2018) proposed a study using a deep learning model to predict the compressive strength of recycled concrete. In their study, softmax regression was used to construct their proposed model. Their findings revealed that the prediction model based on deep learning outperforms the conventional neural network model in terms of precision, performance, and generalization ability, and could be considered a new approach for calculating the strength of recycled concrete.

Furthermore, Latif (2021) conducted a research on developing deep learning model for predicting CCS. He developed long short-term memory (LSTM) model and applied SVM as a conventional machine learning model in order to compare the accuracy of both models. According to his finding, LSTM outperformed SVM with $R^2=0.98$,

$R^2=0.78$, $MAE=1.861$, $MAE=6.152$, and $RMSE=2.36$, $RMSE=7.93$, respectively.

In addition, Van Dao et al. (2019) applied an adaptive neuro-fuzzy inference system (ANFIS) and ANN for compressive strength prediction of geopolymer concrete (GPC). They have prepared concrete mixtures as input parameters. Their findings revealed that ANN and ANFIS models were successful, but ANFIS was better than ANN.

In this paper, BDTR and SVM have been developed for predicting CCS.

Materials and methods

Data set

Nine multivariate variables already reported have been used as input parameters in order to predict the concrete compressive strength. The basic compressive intensity of the output variable is (megapascal-MPa). The number of instances is 1030, and no missing data is available. The selection of dataset components is represented in Table 1.

BDTR model

Boosted tree regression is a hybrid model combining statistical and soft computing techniques. Unlike traditional approaches, whether regression or non-regression, BRT combines regression prediction at various trees for regression to build the best regression tree. Furthermore, when input parameters do not need to be removed from the output, boosted tree regression can help to highlight the nonlinear relationship between the input and output parameters. In boosted tree regression, two techniques are used: regression tree and boosting. The usage of decision tree consequences is one of the key advantages of the regression tree approach. In terms of predictor parameters, the regression trees' technique is unforgiving on outliers and harsh on missing data. To improve model accuracy, numerous decision trees are incorporated into the boosting method (Jumin et al. 2020). The BDTR algorithm is as follow:

$$\hat{y}(x) = \sum_t w_t h_t(x) \quad (1)$$

$$O(x) = \sum_i l(\hat{y}_i, y_i) + \sum_t \Omega(f_t) \quad (2)$$

where $h(x)$ is the tree's output, w is the weight, $l(\hat{y}_i, y_i)$ is the loss function, distance between the truth and the prediction in i th sample, and $\Omega(f_t)$ is the regularization function. Fig. 1 shows the structure of the BDTR model.

Table 1 Statistical components of the utilized datasets

Component	Minimum (kg/m ³)	Maximum (kg/m ³)	Average (kg/m ³)	Standard deviation (SD)
Cement	102.0	540.0	281.20	104.50
Blast furnace slag	0.0	359.4	73.90	86.28
Fly ash	0.0	200.1	54.20	63.99
Water	121.8	247.0	181.60	21.35
Superplasticizer	0.0	32.30	6.20	5.97
Coarse aggregate	801.0	1145.0	972.90	77.75
Fine aggregate	594.0	992.60	773.60	80.17
Age	1.0	365.0	45.70	63.17
CCS	2.3	82.60	35.80	16.71

Support vector machine

SVM is a progressive type of machine learning that highlights statistical learning rules under minor trials in statistical learning theory. Using the standard of structural risk minimization, SVM solves many functional troubles to increase simplified proficiencies, for instance, a limited sample, non-linear, high dimensional number, and global minimum points (Ben-Hur and Weston 2010). Fig. 2 shows the general structure of SVM.

Many typical neural network models can be shown to decrease the error in the training by using the empirical risk minimization principle. SVM, on the other hand, uses the structural risk minimization principle to decrease the upper limit of the simplification error by finding the right balance between the error of the training part and the system’s capability (Latif 2021).

In this study, a sigmoid kernel was used for SVM method.

Sigmoid Kernel : $K(x_i, x_j) = \tanh(x_i^T \times x_j + r)$ (3)

where $K(x_i, x_j)$ is defined as the kernel function. γ , r , and d are kernel parameters.

There are two forms of SVM regression. Form 1 or Epsilon is regarded as the first phase of SVM regression which was used in this study.

Performance indices

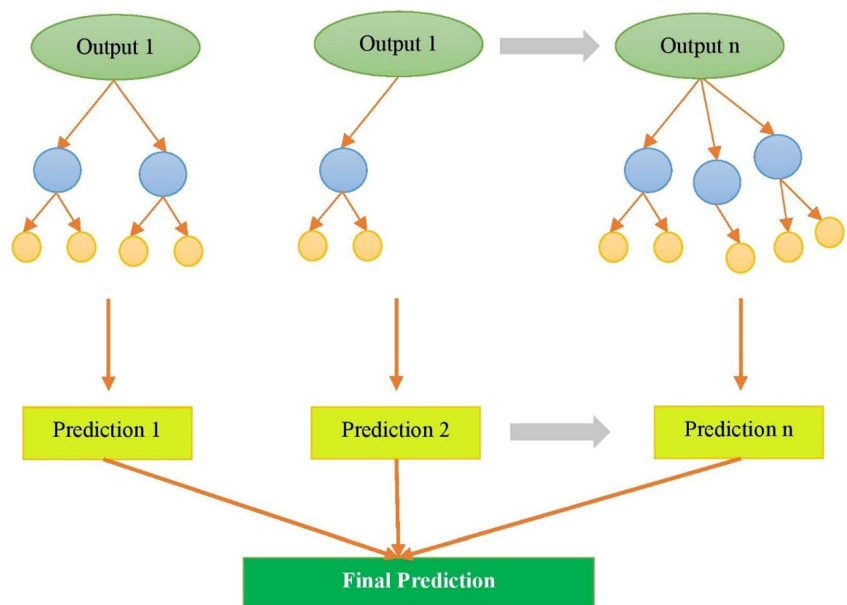
Generally, it is important to evaluate the achievement of the model when evaluating the fulfillment of predictive models, using a wide range of measurement indices to decide the best model. This study suggests unique statistical indices.

R²

R² is the primary regression analysis performance.

$$R^2 = \left\{ (1/N) * \sum [(x_i - X) * (y_i - Y) / (\sigma_x * \sigma_y)] \right\}^2 \quad (4)$$

Fig. 1 The structure of a typical BDTR (Lai et al. 2019)



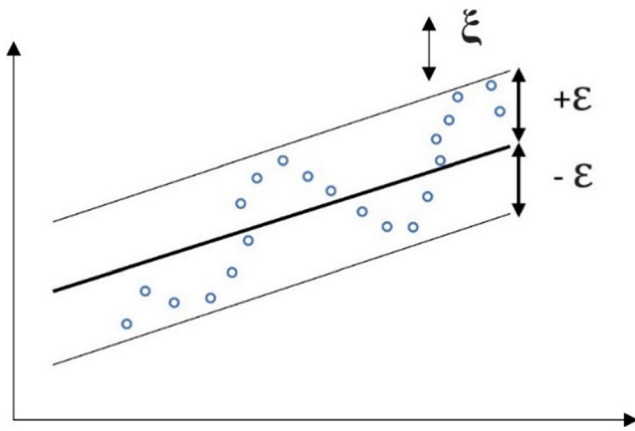


Fig. 2 The architecture of SVM (Latif 2021)

RMSE

RMSE is a typical variance of residuals (predictive errors).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (C_o - C_p)^2} \tag{5}$$

where C_o and C_p are observed and predicted CCS values.

MAE

MAE is a metric for comparing errors between paired observations that express the same concept. A disparity between the true value and the expected value from the example is an increasing prediction error.

Table 2 RSR range and the corresponding performance rate

Performance rating	Unsatisfactory	Satisfactory	Good	Very good
RSR value	$RSR > 0.7$	$0.6 \leq RSR \leq 0.7$	$0.5 \leq RSR \leq 0.6$	$0.00 \leq RSR \leq 0.5$

Fig. 3 The presentation of BDTR model a actual versus prediction; b scatterplot for the overall data

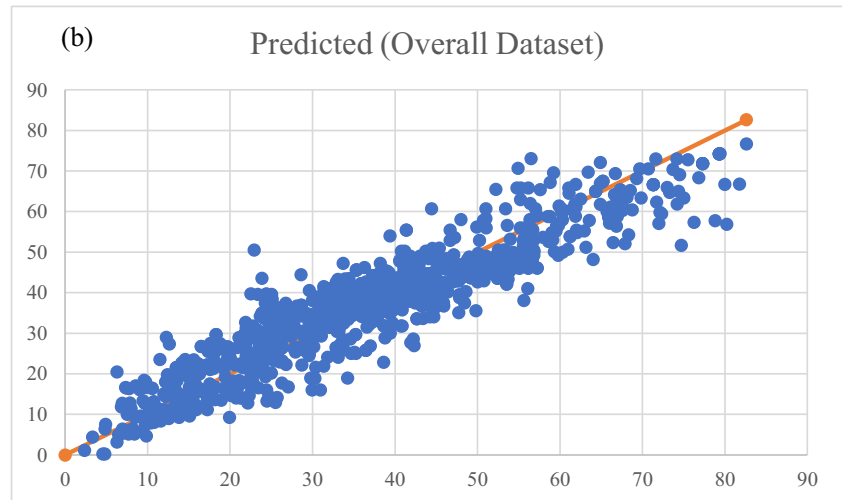
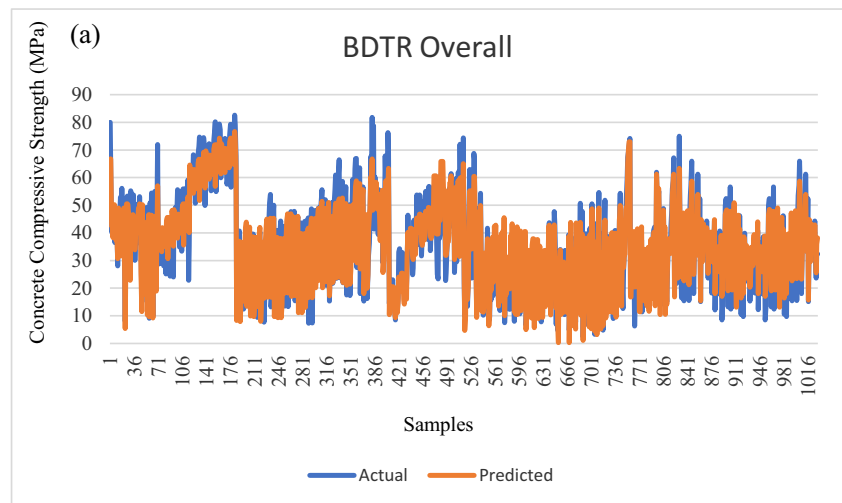
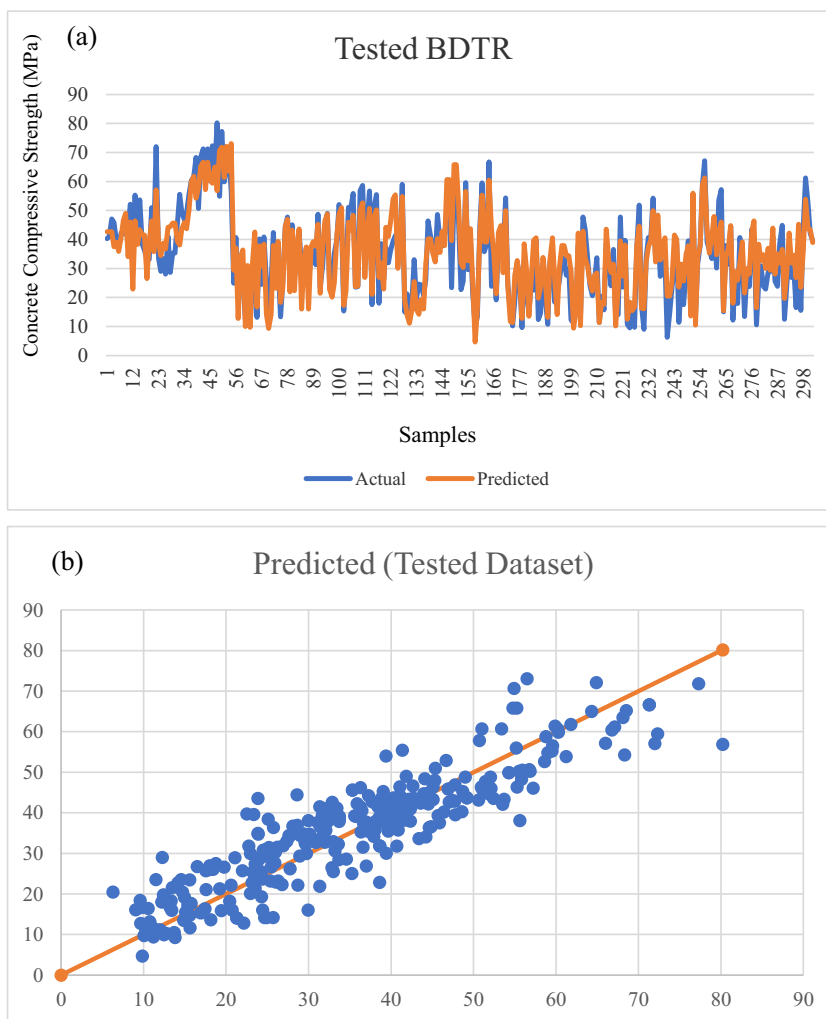


Fig. 4 The presentation of BDTR model **a** actual versus prediction; **b** scatterplot for the tested data



$$MAE = \frac{\sum_{i=1}^n ABS(y_i - \lambda(x_i))}{n} \tag{6}$$

RSR

RMSE-standard deviation ratio (RSR) is used in this research to compare the best model that can be implemented to predict sediment from rivers. RSR is a valuable index for testing the computational models:

$$RSR = \frac{\sqrt{\sum_{i=1}^N (Y_i^{obs} - Y_i^{sim})^2}}{\sqrt{\sum_{i=1}^N (Y_i^{obs} - Y_i^{mean})^2}} \tag{7}$$

Table 3 Performance of BDTR model for overall and testing dataset

Dataset	R ²	RMSE	MAE	RSR
Overall	0.86	6.19	4.91	0.37
Testing	0.81	6.71	5.38	0.44

where Y_i^{mean} is the average of the actual data throughout the monitoring process (Ehteram et al. 2020b). In Table 2, the presentation of RSR index ranges regarding performance rate and class.

Results and discussion

Using a machine learning algorithm, 80% of the randomly selected independent variable data can go through intense training. The remaining untrained data would then be used to assess the output of the model by using a qualified dataset. For the second data partition, the same procedure is applied. In

Table 4 Performance of SVM model for overall and testing dataset

Dataset	R ²	RMSE	MAE	RSR
Overall	0.48	12.70	9.53	0.76
Testing	0.45	13.24	9.52	0.81

this research, two machine learning algorithms were developed and compared, namely, BDTR and SVM, in order to check their accuracy in predicting compressive strength of concrete.

Two different approaches apply to these training datasets. The first approach is a traditional method by which the model is configured by changing the learning rate or a number of algorithm trees manually. The second solution is to add the hyperparameter module of the tuning model to the model.

For the BDTR model, a mixture of two techniques, which are decision tree algorithms and boosting techniques, are Boosted Regression Tree (BRT) models. BRTs match several decision trees repeatedly to improve the model's accuracy since the BDTR is an algorithm used by the multiple additive regression trees (MART) gradient boosting algorithm to train the model. In a stage-wise manner, boosting constructs a set of trees, and each tree depends on prior trees. Therefore, the error in the previous tree is calculated and corrected in the next tree using a predefined loss function. This implies that the prediction is a composite of several weaker prediction models that has resulted in a reliable prediction model.

For the SVM model, sigmoid kernel was used with two kernel parameters, namely, $\gamma=0.10$ and $\text{coefficient}=0.00$. The cross validation was applied in order to enhance the accuracy of the proposed model. The SVM regression type 1 was used with the training and testing partition of 80 and 20%, respectively.

The purpose of this study is to test and compare the ability to predict the CCS of the BDTR and SVM model. Therefore, to ease the performance review, the real and expected outputs were tabulated and plotted. Based on statistical indices such as R^2 , RMSE, MAE, and RSR, the extracted outputs were evaluated as the performance parameters for evaluating the proposed models.

BDTR model successfully applied on CCS, where $R^2 = 0.86$, RMSE = 6.19, MAE=4.91, and RSR=0.37 for overall dataset. Fig. 3 shows the attained prediction value for the desired output CCS by BDTR versus actual data and scatter plot for the overall dataset.

For the testing part, the model performed the lowest accuracy, where $R^2=0.81$, RMSE=6.71, MAE=5.38, and RSR=0.44. Fig. 4 shows the attained prediction value for the

Fig. 5 The presentation of SVM model **a** actual versus prediction; **b** scatterplot for the overall data

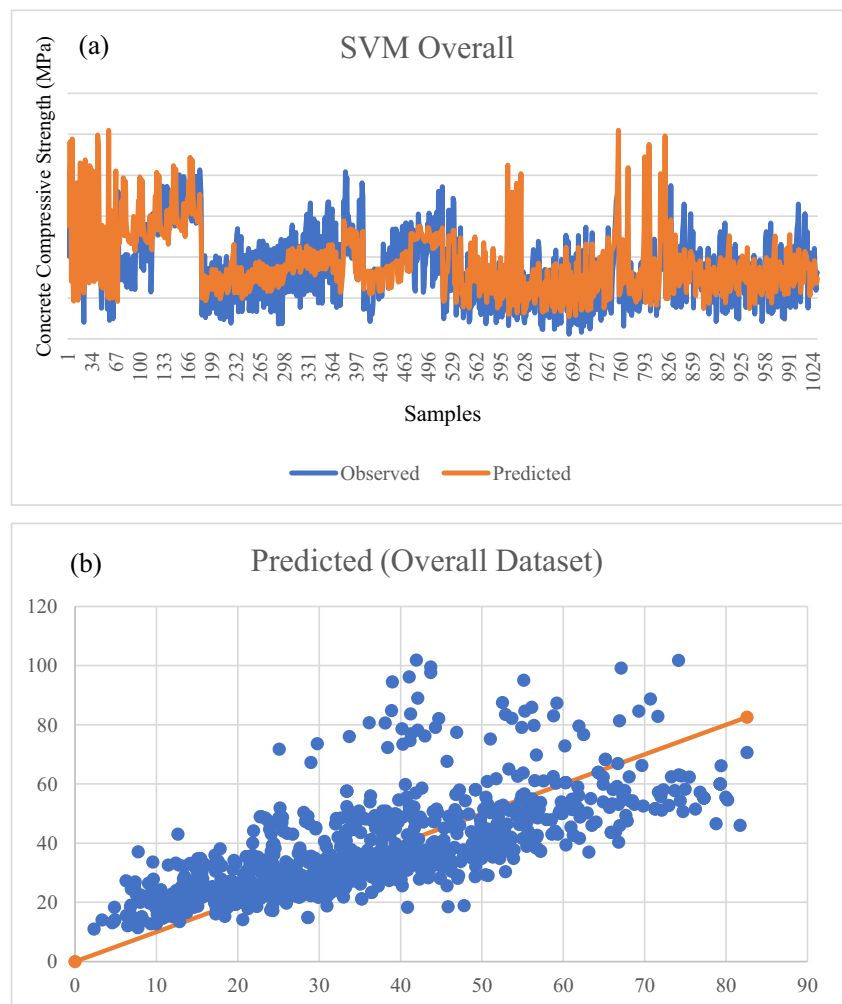
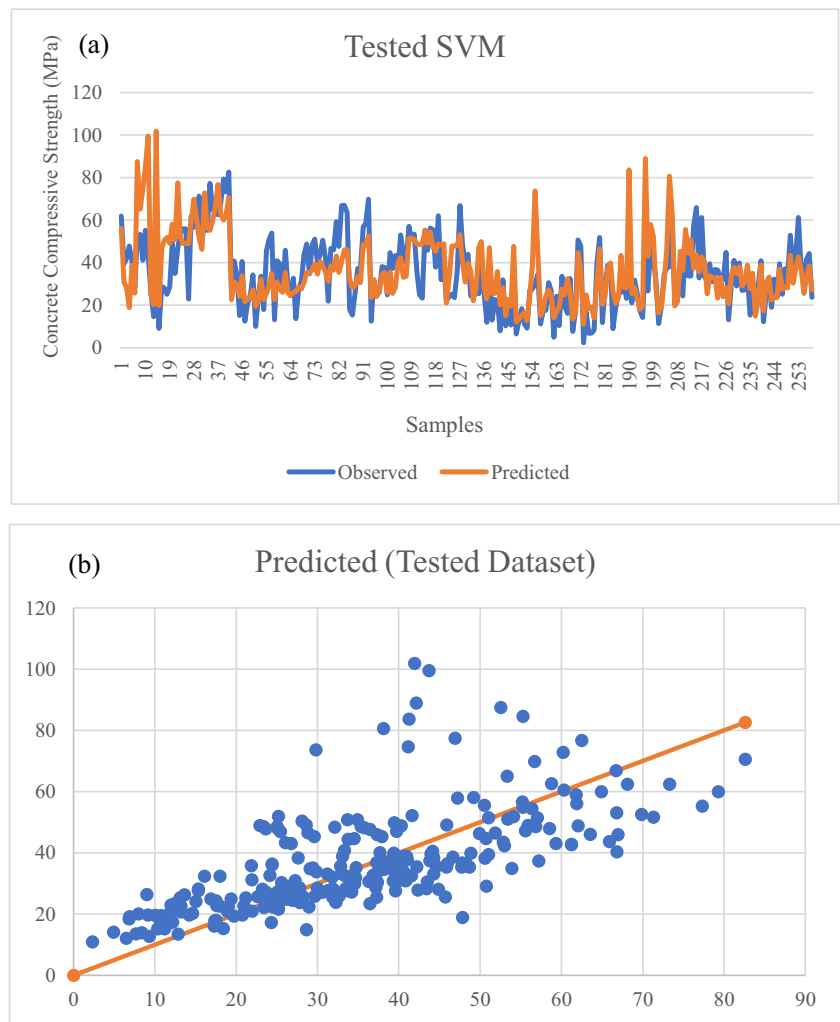


Fig. 6 The presentation of SVM model **a** actual versus prediction; **b** scatterplot for the tested data



desired output CCS by BDTR versus actual data and scatter plot for the tested dataset.

Table 3 shows the performance of BDTR model for the overall and tested dataset.

Regarding the SVM model, the results showed the lower accuracy than BDTR model. For the overall SVM dataset result, $R^2=0.48$, $RMSE=12.70$, $MAE=9.53$, and $RSR=0.76$. For the tested dataset, $R^2=0.45$, $RMSE=13.24$, $MAE=9.52$, and $RSR=0.81$. Table 4 shows the performance of the proposed SVM model for the overall and tested dataset.

Table 5 Comparison of BDTR and SVM results for overall and testing dataset

Models	Dataset	R^2	RMSE	MAE	RSR
BDTR	Overall	0.86	6.19	4.91	0.37
	Testing	0.81	6.71	5.38	0.44
SVM	Overall	0.48	12.70	9.53	0.76
	Testing	0.45	13.24	9.52	0.81

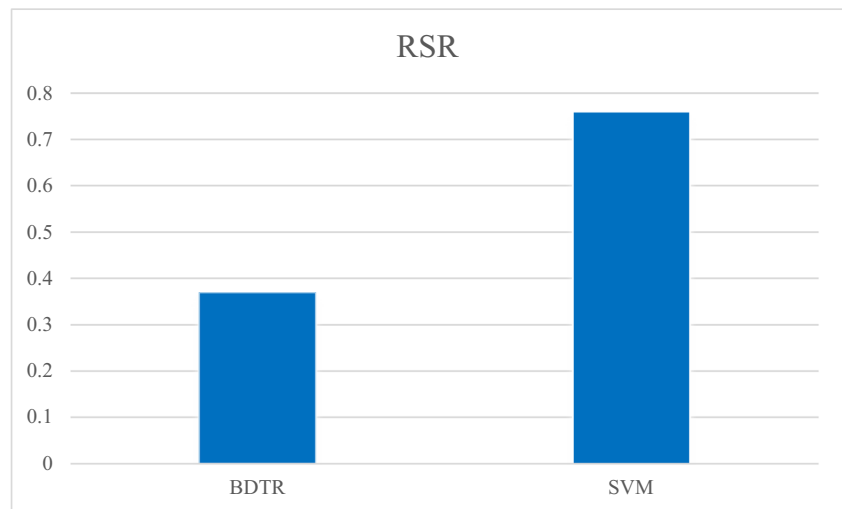
Figs. 5 and 6 show the attained prediction value for the desired output CCS by SVM versus actual data and scatter plot for the overall and tested dataset, respectively.

According to the results, BDTR outperformed SVM model. Table 5 shows the comparison of both BDTR and SVM model for the testing and overall dataset.

The RSR value for the BDTR model is in the range of very good, while it is in the range of unsatisfactory for the SVM model. Fig. 7 shows the RSR of BDTR and SVM for the overall dataset.

According to the results, BDTR outperformed SVM with a significant different. The results of the current study were similar to the previous studies in literature. Ling et al. (2019) optimized SVM and built ANN and DT models to compare the prediction precision with the SVM model. According to their results, SVM outperformed ANN and DT models. On the other hand, Shaqadan (2020) developed SVM and ANN models to predict the CCS. According to his result, both SVM and ANN showed a good correlation coefficient of 0.929 and 0.986, respectively. Furthermore, Latif (2021) developed LSTM and SVM models for predicting CCS. According to

Fig. 7 RSR value for the computed models



his finding, LSTM outperformed SVM with $R^2=0.98$ and $R^2=0.78$, $MAE=1.861$ and $MAE=6.152$, and $RMSE=2.36$ and $RMSE=7.93$, respectively. Therefore, it can be concluded that the results of the current study was similar to the other studies in the field. BDTR can be a reliable model for predicting CCS.

Conclusion

In order to precisely build a suitable model, this study focuses on predicting CCS using two machine learning models namely, BDTR and SVM. As the input of the models, eight distinct concrete components were used, and the observed CCS was used as the model output. In predicting concrete compressive strength, BDTR outperformed SVM. It can be assumed that BDTR is an effective method to predict CCS, but it can depend on the input appropriateness of the datasets. In determining the weakness of concrete compressive power, the outcome of this study may be very significant. Future experiments can be carried out using a different dataset to check the consistency of the proposed model.

Abbreviations *BDTR*, Boosted decision tree regression; R^2 , Coefficient of determination; *RMSE*, Root mean square error; *HPC*, High-performance concrete; *CCS*, Concrete compressive strength; *SFS*, Sequential feature selection; *NID*, Neural interpretation diagram; *ANN*, Artificial neural network; *UHPC*, Ultra-high performance concrete; *SVM*, Support vector machine; *DT*, Decision tree; *ANFIS*, Adaptive neuro-fuzzy inference system; *GPC*, Geopolymer concrete; *AI*, Artificial intelligence; *BRT*, Boosted regression trees; *MART*, Multiple additive regression

Acknowledgements The author would like to thank Prof. I-Cheng Yeh for providing the data set online.

Author contribution Writing original draft, methodology, and analysis: Sarmad Dashti Latif.

Data availability Not applicable.

Code availability Not applicable.

Declarations

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Conflict of interest The author declares no competing interests.

References

- Abuodeh OR, Abdalla JA, Hawileh RA (2020) Assessment of compressive strength of Ultra-high Performance Concrete using deep machine learning techniques. *Appl Soft Comput J* 95:106552. <https://doi.org/10.1016/j.asoc.2020.106552>
- Al-Shamiri AK, Kim JH, Yuan TF, Yoon YS (2019) Modeling the compressive strength of high-strength concrete: an extreme learning approach. *Constr Build Mater* 208:204–219. <https://doi.org/10.1016/j.conbuildmat.2019.02.165>
- Al-Shamiri AK, Yuan TF, Kim JH (2020) Non-tuned machine learning approach for predicting the compressive strength of high-performance concrete. *Materials (Basel)* 13:1–15. <https://doi.org/10.3390/ma13051023>
- Aprianti E, Shafiqh P, Bahri S, Farahani JN (2015) Supplementary cementitious materials origin from agricultural wastes - a review. *Constr Build Mater* 74:176–187
- Ben-Hur A, Weston J (2010) A user's guide to support vector machines. *Methods Mol Biol*. https://doi.org/10.1007/978-1-60327-241-4_13
- Borhana AA, Kamal DDBM, Latif SD, et al (2020) Fault detection of bearing using support vector machine-SVM. In: 2020 8th International Conference on Information Technology and Multimedia (ICIMU). pp 309–315
- Chopra P, Sharma RK, Kumar M, Chopra T (2018) Comparison of machine learning techniques for the prediction of compressive strength of concrete. *Adv Civ Eng* 2018:1–9. <https://doi.org/10.1155/2018/5481705>

- Chou J-S, Chiu C-K, Farfoura M, Al-Taharwa I (2011) Optimizing the prediction accuracy of concrete compressive strength based on a comparison of data-mining techniques. *J Comput Civ Eng* 25: 242–253. [https://doi.org/10.1061/\(asce\)cp.1943-5487.0000088](https://doi.org/10.1061/(asce)cp.1943-5487.0000088)
- Deng F, He Y, Zhou S, Yu Y, Cheng H, Wu X (2018) Compressive strength prediction of recycled concrete based on deep learning. *Constr Build Mater* 175:562–569. <https://doi.org/10.1016/j.conbuildmat.2018.04.169>
- Dutta S, Samui P, Kim D (2018) Comparison of machine learning techniques to predict compressive strength of concrete. *Comput Concr*. <https://doi.org/10.12989/cac.2018.21.4.463>
- Ehteram M, Ahmed AN, Latif SD, Huang YF, Alizamir M, Kisi O, Mert C, el-Shafie A (2020a) Design of a hybrid ANN multi-objective whale algorithm for suspended sediment load prediction. *Environ Sci Pollut Res* 28:1596–1611. <https://doi.org/10.1007/s11356-020-10421-y>
- Ehteram M, Ahmed AN, Ling L, Fai CM, Latif SD, Afan HA, Banadkooki FB, el-Shafie A (2020b) Pipeline scour rates prediction-based model utilizing a multilayer perceptron-colliding body algorithm. *Water (Switzerland)* 12. <https://doi.org/10.3390/w12030902>
- Ehteram M, Yenn F, Najah A et al (2020c) Performance improvement for infiltration rate prediction using hybridized Adaptive Neuro-Fuzzy Inferences System (ANFIS) with optimization algorithms. *Ain Shams Eng J* 11:12–1676. <https://doi.org/10.1016/j.asej.2020.08.019>
- Feng DC, Liu ZT, Wang XD, Chen Y, Chang JQ, Wei DF, Jiang ZM (2020) Machine learning-based compressive strength prediction for concrete: an adaptive boosting approach. *Constr Build Mater* 230: 117000. <https://doi.org/10.1016/j.conbuildmat.2019.117000>
- Gonzalez-Corominas A, Etxeberria M (2016) Effects of using recycled concrete aggregates on the shrinkage of high performance concrete. *Constr Build Mater* 115:32–41. <https://doi.org/10.1016/j.conbuildmat.2016.04.031>
- Gupta R, Kewalramani MA, Goel A (2006) Prediction of concrete strength using neural-expert system. *J Mater Civ Eng* 18:462–466. [https://doi.org/10.1061/\(asce\)0899-1561\(2006\)18:3\(462](https://doi.org/10.1061/(asce)0899-1561(2006)18:3(462)
- Hoang ND, Pham AD, Nguyen QL, Pham QN (2016) Estimating compressive strength of high performance concrete with Gaussian process regression model. *Adv Civ Eng* 2016:1–8. <https://doi.org/10.1155/2016/2861380>
- Jumin E, Zaini N, Ahmed AN, Abdullah S, Ismail M, Sherif M, Sefelnasr A, el-Shafie A (2020) Machine learning versus linear regression modelling approach for accurate ozone concentrations prediction. *Eng Appl Comput Fluid Mech* 14:713–725. <https://doi.org/10.1080/19942060.2020.1758792>
- Jumin E, Basaruddin FB, Yusoff YBM, Latif SD, Ahmed AN (2021) Solar radiation prediction using boosted decision tree regression model: a case study in Malaysia. *Environ Sci Pollut Res* 28: 26571–26583. <https://doi.org/10.1007/s11356-021-12435-6>
- Kalooop MR, Kumar D, Samui P, Hu JW, Kim D (2020) Compressive strength prediction of high-performance concrete using gradient tree boosting machine. *Constr Build Mater* 264:120198. <https://doi.org/10.1016/j.conbuildmat.2020.120198>
- Kasperkiewicz J, Raczyński J, Dubrawski A (1995) HPC strength prediction using artificial neural network. *J Comput Civ Eng* 9:279–284. [https://doi.org/10.1061/\(asce\)0887-3801\(1995\)9:4\(279](https://doi.org/10.1061/(asce)0887-3801(1995)9:4(279)
- Kylyli A, Fokaides PA (2017) Policy trends for the sustainability assessment of construction materials: A review. *Sustain Cities Soc*. <https://doi.org/10.1016/j.scs.2017.08.013>
- Lai V, Ahmed AN, Malek MA, Abdulmohsin Afan H, Ibrahim RK, el-Shafie A, el-Shafie A (2019) Modeling the nonlinearity of sea level oscillations in the Malaysian coastal areas using machine learning algorithms. *Sustain* 11. <https://doi.org/10.3390/su11174643>
- Lai V, Malek MA, Abdullah S et al (2020) Time-series prediction of sea level change in the east coast of Peninsular Malaysia from the supervised learning approach. *Int J Des Nat Ecodynamics*. <https://doi.org/10.18280/ijdne.150314>
- Latif SD (2021) Concrete compressive strength prediction modeling utilizing deep learning long short-term memory algorithm for a sustainable environment. *Environ Sci Pollut Res* 28:30294–30302. <https://doi.org/10.1007/s11356-021-12877-y>
- Latif SD, Ahmed AN (2021) Application of deep learning method for daily streamflow time-series prediction : a case study of the Kowmung River at Cedar Ford , Australia. *Int J Sustain Dev Plan* 16:497–501. <https://doi.org/10.18280/ijdsdp.160310>
- Latif SD, Ahmed AN, Sherif M et al (2020a) Reservoir water balance simulation model utilizing machine learning algorithm. *Alex Eng J* 60:1365–1378. <https://doi.org/10.1016/j.aej.2020.10.057>
- Latif SD, Azmi MSBN, Ahmed AN et al (2020b) Application of artificial neural network for forecasting nitrate concentration as a water quality parameter: a case study of Feitsui Reservoir, Taiwan. *Int J Des Nat Ecodynamics*. <https://doi.org/10.18280/ijdne.150505>
- Latif SD, Usman F, Pirot BM (2020c) Implementation of value engineering in optimizing project cost for sustainable energy infrastructure asset development. *Int J Sustain Dev Plan* 15:1045–1057. <https://doi.org/10.18280/ijdsdp.150709>
- Latif SD, Ahmed AN, Sathiamurthy E, Huang YF, el-Shafie A (2021a) Evaluation of deep learning algorithm for inflow forecasting : a case study of Durian Tunggal Reservoir, Peninsular Malaysia. *Nat Hazards*. <https://doi.org/10.1007/s11069-021-04839-x>
- Latif SD, Birima AH, Najah A et al (2021b) Development of prediction model for phosphate in reservoir water system based machine learning algorithms. *Ain Shams Eng J*. <https://doi.org/10.1016/j.asej.2021.06.009>
- Latif SD, Marhain S, Hossain S et al (2021c) Optimizing the operation release policy using charged system search algorithm : a case study of Klang Gates Dam , Malaysia. *Sustain* 13:19. <https://doi.org/10.3390/su13115900>
- Ling H, Qian C, Kang W, Liang C, Chen H (2019) Combination of Support Vector Machine and K-Fold cross validation to predict compressive strength of concrete in marine environment. *Constr Build Mater* 206:355–363. <https://doi.org/10.1016/j.conbuildmat.2019.02.071>
- Mirzahosseini M, Jiao P, Barri K et al (2019) New machine learning prediction models for compressive strength of concrete modified with glass cullet. *Eng Comput (Swansea, Wales)*. <https://doi.org/10.1108/EC-08-2018-0348>
- Naderpour H, Rafiean AH, Fakharian P (2018) Compressive strength prediction of environmentally friendly concrete using artificial neural networks. *J Build Eng* 16:213–219. <https://doi.org/10.1016/j.jobe.2018.01.007>
- Najah A, Teo FY, Chow MF, Huang YF, Latif SD, Abdullah S, Ismail M, el-Shafie A (2021) Surface water quality status and prediction during movement control operation order under COVID-19 pandemic: case studies in Malaysia. *Int J Environ Sci Technol* 18:1009–1018. <https://doi.org/10.1007/s13762-021-03139-y>
- Parsaie A, Haghiabi AH, Latif SD (2021) Predictive modelling of piezometric head and seepage discharge in earth dam using soft computational models. *Environ Sci Pollut Res*. <https://doi.org/10.1007/s11356-021-15029-4>
- Shaqadan A (2020) Prediction of concrete strength using support vector machines algorithm. *Mater Sci Forum*. <https://doi.org/10.4028/www.scientific.net/MSF.986.9>
- Vakharia V, Gujar R (2019) Prediction of compressive strength and portland cement composition using cross-validation and feature ranking techniques. *Constr Build Mater* 225:292–301. <https://doi.org/10.1016/j.conbuildmat.2019.07.224>
- Van Dao D, Ly HB, Trinh SH et al (2019) Artificial intelligence approaches for prediction of compressive strength of geopolymer concrete. *Materials (Basel)* 12. <https://doi.org/10.3390/ma12060983>

- Wang D, Ju Y, Shen H, Xu L (2019) Mechanical properties of high performance concrete reinforced with basalt fiber and polypropylene fiber. *Constr Build Mater* 197:464–473. <https://doi.org/10.1016/j.conbuildmat.2018.11.181>
- Yaseen ZM, Deo RC, Hilal A, Abd AM, Bueno LC, Salcedo-Sanz S, Nehdi ML (2018) Predicting compressive strength of lightweight foamed concrete using extreme learning machine model. *Adv Eng Softw* 115:112–125. <https://doi.org/10.1016/j.advengsoft.2017.09.004>
- Young BA, Hall A, Pilon L, Gupta P, Sant G (2019) Can the compressive strength of concrete be estimated from knowledge of the mixture proportions?: New insights from statistical analysis and machine learning methods. *Cem Concr Res* 115:379–388. <https://doi.org/10.1016/j.cemconres.2018.09.006>
- Zhong R, Wille K (2015) Material design and characterization of high performance pervious concrete. *Constr Build Mater* 98:51–60. <https://doi.org/10.1016/j.conbuildmat.2015.08.027>
- Zhong Y, Wu P (2015) Economic sustainability, environmental sustainability and constructability indicators related to concrete- and steel-projects. *J Clean Prod* 108:748–756. <https://doi.org/10.1016/j.jclepro.2015.05.095>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.