JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue

JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)

An International Scholarly Open Access, Peer-reviewed, Refereed Journal

APPLICATIONS OF MACHINE LEARNING IN CIVIL ENGINEERING

[1]Shambhavi S. Shevade,[2]Karina J. Vedmutha,[3]Shrawani D. Sawale, [4]Vaishnavi G. Saraf, [5]Dr. Deepa A. Joshi

[1]Student,[2]Student,[3]Student, [4]Student, [5]Professor
Department of Civil Engineering,
Dr. D. Y. Patil Institute of Technology Pimpri, Pune,India

Abstract: The growing demand for using soft computing techniques like Machine Learning (ML) is rising day by day and it is the most successful branch of artificial intelligence. It provides a unique way in handling complex nonlinear dataset to obtain precise solution. This paper provides comprehensive review on the growing applications of ML algorithms for Structural Engineering, Construction Management, Hydrology and Water Resource Management, Environmental Engineering, Transportation Engineering.

***IndexTerms*–Artificial Intelligence, Machine Learning (ML), Supervised Learning, Unsupervised Learning, Reinforcement Learning, Deep Learning**

Introduction:

Artificial Intelligence (AI) is developing very fast. Many industries such as engineering, medicine, finance, governments, stocks, accounting, designing, production etc. are using soft computing techniques like Machine Learning (ML), Deep Learning (DL) and natural language processing collectively known as AI.

In civil engineering, there are many practical constraints which require experience to solve those problems effectively like planning, execution, designing etc. But if a model is developed having intelligence like human brain will reduce many problems in ordinary lives.

The discipline of ML focuses on the use of specialized algorithms for extracting patterns and information from large and complex data sets.

The growing volume of information databases presents opportunities for advanced data analysis techniques from machine learning (ML) research. Thanks to the rapid increase of data availability, as well as increasing computational capacities and simplified programming methods, machine learning tools are being progressively applied in the fields of civil engineering.

Study was carried out using nine different machine learning models. The prediction performances of these machine learning models were evaluated using statistical performance metrics such as Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and R2. The results showed that out of the nine machine learning models such as ensemble learning algorithms: XGBoost, Extra Trees, AdaBoost boosted, Decision Trees, Random Forest in that order performed better in terms of prediction accuracy and reduced prediction errors compared to the traditional machine learning algorithms such as Linear Regression, SVR, MLP Neural Network, Lasso [1].

Artificial Intelligence:

John McCarthy, father of AI explained artificial intelligence as “artificial intelligence is the science and engineering of making intelligent machines, especially intelligent computer programs.” [2]

Naming “man vs. machine”, Artificial Intelligence (AI) achieved publicity in the past few years when the question was asked whether a model is capable enough for defeating a man in chess – a game that needs a high degree of intelligence and stability.

Since the beginning of this era, powerful chess programs, known as "Rebel", "Deep Blue" or "Deep Fritz" defeated world chess champion (Garry Kasparov) and international grandmasters. This event was recognized as a prominent AI success story [3].

AI covers area i.e., knowledge base, expert systems, deep learning, computer vision, machine learning, natural language processing, image processing etc.

Broadly, AI has 2 types:

1. Weak AI
2. Strong AI

Weak AI:

The concept of weak AI is that the machines act like they are intelligent. Weak AI shows various abilities like moving, talking and thinking if they are programmed in that manner [4]. Examples of weak AI are digital voice assistants like Siri and Alexa; Recommendation engines; Chatbots etc.

Strong AI:

It performs tasks like human intelligence. The concept behind the strong AI is the model will think and can work with wide range of problems. Strong AI focused to create intelligent machines which are similar to human brains. In coming decades, there will be many machines which will work on the concept of strong AI who will be more intelligent than humans.

Machine Learning:

Machine learning (ML) is a class of artificial intelligence (AI) that focuses on preparing computers which are able to predict various parameters with available datasets and algorithms. Most significantly, it provides computer systems the ability to learn and enhance themselves with their experience instead of being explicitly programmed [5].

Although machine learning in civil industry could looks like a distant idea decade away from being a reality, this technology's trend is closer. Machine learning has been acquiring popularity in the construction industry in latest years [6]. This technology is making changes in every area of the world like cost estimation and prediction, schedule management, safety management for construction site, construction project management including reduced risks and building energy demand predictions [7].

In general, ML can be classified into three main categories based on the learning process: supervised learning, unsupervised learning, and reinforcement learning.

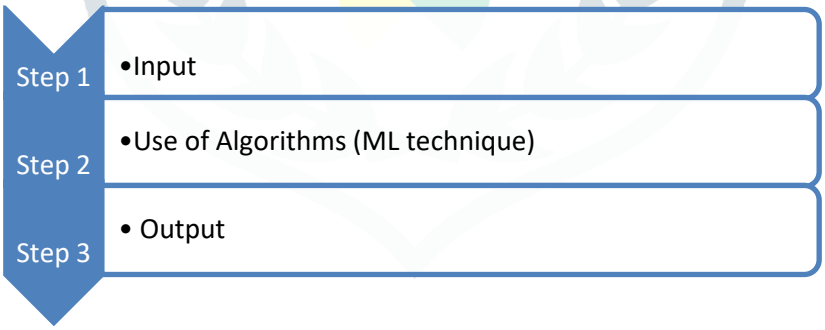

Fig 1: Basic flowchart for developing ML model

Supervised learning:

In the supervised learning algorithms, the data which is provided to the model for training is the labeled dataset, which is known as feature vector, which indicates that the datasets having observations with their known outcomes. These machines are allowed to produce relationship between input and output parameters. The widely used algorithms in supervised machine learning techniques are decision tree, support vector machine, AdaBoost, bagging, boosting, artificial neural network, gene expression programming [8].

Unsupervised learning:

Unlike supervised learning, unsupervised methods do not have much data needed for developing a model. In many cases there is not even a pre-specified dependent variable as the focus is to extract patterns in the data [9]. Examples of Unsupervised learning are K-means clustering, KNN (k-nearest neighbors) and Hierarchal clustering.

Machine learning process:

Machine learning process are the steps which are to be followed while developing any AI model. There are 7 steps which are shown in diagram below.

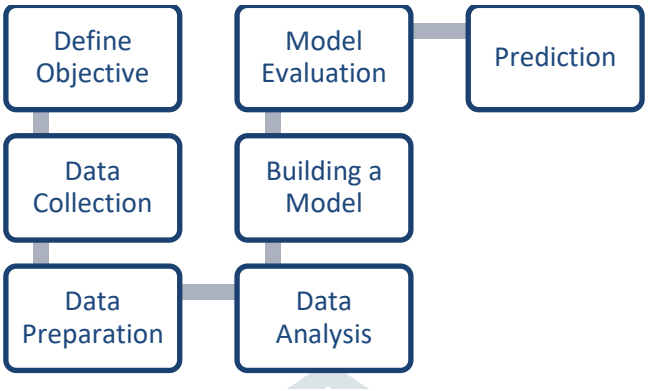

Fig 2: Flowchart showing process of ML

All the above steps are equally important while developing any model. Above steps are explained in the next section.

Step 1: **Define objective**

The goal is to find the objective i.e., finding the exact problem and what will be the output after solving this problem. The types of constraint are present while preparing this model. Also, input parameters needed for desired output is also taken into account.

Step 2: **Data collection**

This step nearly takes 60% of the time while executing the project. Input features are to be identified in this step. There are different websites from which the dataset can be collected. Provided dataset should be authentic. Input parameters are again filtered and optimum number of parameters is taken for developing efficient model.

Step 3: **Data preparation**

Dataset is formatted properly for training the model. If the dataset contains any missing value, the model will not train properly. The missing values should be filled with average value for that parameter or at least number “0” to get the better results. If any duplicate data is present, that should be removed from the dataset. Important part in this step is data scaling. If the dataset contains values which have different units, then data scaling is done. Outliers are removed from the dataset.

Step 4: **Data analysis**

Data is viewed with respect the problem solved by using AI. The quality of data is checked as per the requirements. If the data provided is noisy, then the concept of parameter tuning comes into the picture. Data should not be overcrowded and no bias should be present. All type of occurrences should be treated equally. Some data visualization tools like Tablue, Cborn, Matpotlab, Panverbi etc. can be used for data analysis.

Step 5: **Building a model**

For building an AI model, various languages can be used. Some of the languages are Python, Java, R programming, Matlab, C++ etc. Most widely Python is used for model building as it is easy and contains in-built libraries. While developing, the dataset is divided into two parts i.e., training and testing. Various types of splits are provided but generally the ratio of 4:1 is adopted.

It is important to note that, training data is always greater than testing data.

Step 6: **Model evaluation**

After developing a model, model is evaluated for various datasets which were not included for training the dataset. Various statistical methods are adopted to check the accuracy the model. Mean square root error, standard deviation, variance etc. are calculated with respect to the model performance. Some validation techniques are used in this step.

a. Percentage split method:

   This is the simplest method in validation of model. In this method, dataset is split up into two parts. Some of the common combinations are 80:20; 85:15; 70:30 etc. The advantage of this method is that it is usually adapted over remaining methods and doesn't take much time to compute. However, its evaluation can have a high tolerance [6].

b. K fold cross validation:

   The data is separated into k disjoint sets and one of it is used as testing data and remaining of them are used as training data. This process will carry on till every subset has been used once as a testing dataset. In simple words, the data set is split into k subsets, and the holdout method is repeated k times. Every time, one of the k subsets is used for testing set and the other k-1 subsets are assembled to form a training set. Every data point gets to be in a testing set exactly once, and gets to be in a training set k-1 times. The disadvantage of this method is that the training algorithm has to be run again and again from beginning k times, which means it takes k times to predict and to make an evaluation [6].

Step 7: **Predictions**

The last step is to predict the outcomes for which it is designed.

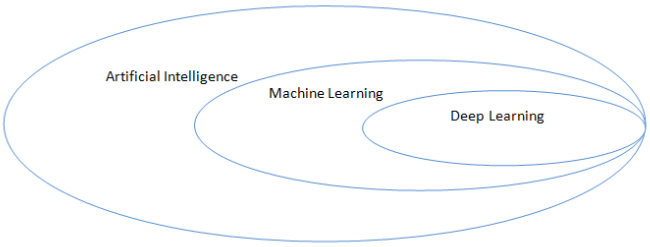

Fig 3: Relationship between AI, ML and DL

Supervised ML algorithms:

Support Vector Machine (SVM):

The support vector machine (SVM) is considered as one of the most robust and accurate methods among data mining algorithms [10]. This algorithm can be viewed as a machine which divides the dataset into parts to make it more precise to derive meaningful patterns. Thus, it forms a decision boundary in order to place the data points according to its respective class. This algorithm is more suitable for classification problems containing a linear separation between the data points [11].

Extra tree regressor (ETR):

The concept of ETR is very similar to RF regressor and the only difference is in the manner of construction of the decision trees in the forest. Each Decision Tree in the ETR is constructed from the original training sample. Then, at each test node, every single tree is provided with a random sample of k features from the feature-set from which each decision tree must select the best feature to split the data based on some mathematical criteria. This sample of features leads to the creation of multiple de-correlated decision trees. This leads to good efficiency of model. When it is compared with RF, ETR can deal with more noisy type of data. This algorithm can be used for both Classification and Regression problems. Also, it can capture nonlinear relationships [12].

Random Forest regressor (RF):

To perform this regression, Random Forest (RF) model utilizes two steps, bootstrapping and bagging, respectively. In the bootstrapping step, a set of decision trees is produced by the growth of every single individual tree that uses a random training dataset sample. In the bagging step, which process in two steps, is used to divide the decision tree nodes after achieving the ensemble, where many random subsets of training data are selected during the initial bagging process. decision-making process is accomplished by selecting the best subset and its value. [12]

XGBoost regressor:

Extreme Gradient Boosting (XGBoost) is an open-source library that provides an optimum and accurate implementation of the gradient boosting algorithm. Every time, XGBoost gives more importance to functional space while decreasing the cost of a model. This algorithm catches nonlinear relationships easily. It has carried out well on many tabular datasets with a fair amount of data. It can also deal effectively with outliers. So, while using this algorithm no need to remove outliers from dataset. This is the biggest advantage of using XGBoost regressor [13].

Decision tree regressor (DT):

Decision tree regressor is a supervised learning technique that can be adopted for both classification and regression problems, but mainly it is preferred for dealing with classification problems. It is a tree-structured algorithm, where internal nodes present the features of a dataset, branches represent the decision rules made by algorithm and each leaf node represents the output. In a Decision tree, there are two types of nodes, Decision Node and Leaf Node. Decision nodes are utilized to make any decision and can have multiple branches, whereas Leaf nodes are the outcomes of those decisions and do not carry any further branches. The decisions or the test are executed on the basis of features of the given dataset. It is a graphical representation for getting all the possible outcomes to a problem/decision based on available conditions. It is known as decision tree because, just like a tree, it starts with the root node, which expands on further branches and grow like a tree-like structure.

Linear Regression:

Linear regression is the easiest algorithm in the supervised learning algorithm. It is a method to predict dependent variable (i.e., output variable) based on the values of independent variables. This algorithm finds out a linear relation between input and output. So, it is called as Linear Regression. It can be further used when continuous quantity is to be find out. Suppose, number of covid cases is to be calculated, then that depends on various factors like city, age etc. so this is treated as independent variables and number of covid cases is the dependent variable [5].

K-nearest neighbor regressor (KNN):

In a prediction problem by KNN regression, the outcome of a data point will be decided based on the output of nearby points, and the distance between them [14]. In KNN regression, the prediction for the sample point is the average of the values for the k-nearest neighbors. It may be necessary to reduce the dimensions of the problem first [15]. One major disadvantage in

measuring distance directly from the training set is in the case where variables have different measurement scales/units or there is a combination of numerical and categorical variables. The average of these data points is finally predicted for the new point.

AdaBoost Regressor:

AdaBoost regressor is a supervised machine learning algorithm and ensemble learning technique. It is based on base learners like decision tree. In this case, every model depends on its previous model. In AdaBoost, there is only one root node and two leaf nodes called as stumps, this algorithm is used to solve both classification and regression problems during predictions we assigned weights for output variables table. Then, we predict first model based on that we have predicted second model i.e. every next model is based on its previous model. Finally, we take the combination of all models and then predicted the values [16].

Artificial Neural Network (ANN):

ANN is used to mimic like one neuron which is connected to many other neurons as in human neuron system. These neurons are known as nodes. The data is sent via input layer and weights are attached according to their bias. There are three types of layers present in the model which are input layer, hidden layer and output layer. To pass the values from input layers to hidden layers, activation function is called off. Activation functions are necessary to activate the neurons same as human neurons. There are many types of activation function such as sigmoid, ReLU, Tanh, softmax etc. There can be "n" number of hidden layers [17-18].

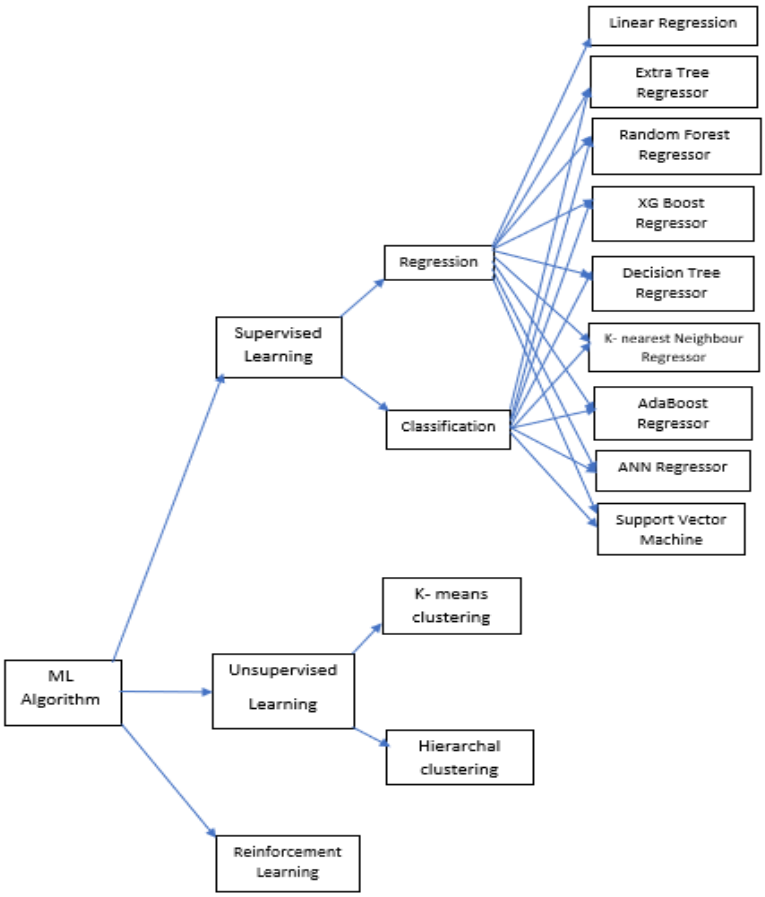

Fig 4: Classification of algorithms

ML Applications for Prediction of Civil Engineering Fields:

There are many methods/ algorithms by which development of model takes place. For solving civil engineering problems, widely supervised learning algorithms can be used. This paper focuses on five sectors of civil engineering such as:

1. Structural engineering
2. Construction management
3. Hydrology and water resource management
4. Environmental engineering
5. Transportation engineering

All these sectors are using ML techniques to strengthen the process and effective working. To solve a particular problem, various languages like C; C++; Matlab; Java; Python etc. can be used. Working in python can be easy as it have various in-built libraries which can reduce coding part. For choosing efficient algorithm to develop a model, "AutoML" library can be used.

Structural engineering:

The equation was built by executing linear regression on a dataset of 164 physical experiments by utilizing the ratio of wall neutral axis depth-to-compression zone width, the ratio of wall length-to-compression zone width, wall shear stress ratio, and the configuration of the boundary zone reinforcement as the predictor (or independent) variables [20]. Jiang et al. introduced a fuzzy stochastic neural network model for nonparametric identification of civil structures using the nonlinear autoregressive moving average with exogenous inputs model through the combination of two computational intelligence techniques, i.e., fuzzy logic and neural networks [19]. Ibrahim et al. [21] compared the classification performance of three machine learning algorithms, SVM, K-Nearest Neighbor (KNN), and Convolutional Neural Network (CNN), for evaluating the health condition of two simulated four- and eight story building structures subjected to earthquakes. They conclude that CNN performed well over SVM and KNN in terms of accuracy for damage detection. But, for developing CNN skill requirement in the field of data science is also high.

Huang and Burton (2019) studied six machine learning algorithms- Logistic Regression, Decision Tree, Random Forests, Adaptive Boosting, Support Vector Machine and Multilayer Perceptron- for the classification of 114 masonry-infilled RCC frame specimens in to four distinct failure modes, utilizing nine structural parameters as inputs, of which masonry is also a variable with binary values (0 or 1). Most of the models were able to achieve more than 80% prediction accuracy, based on recall score, with the highest value of 85.7% achieved with the Adaptive Boosting and SVM algorithms [22]. Feng et al. (2020) [23] employed ensemble learning algorithms such as Adaptive Boosting (AdaBoost), for classification of failure modes of RC columns. Seismic performance evaluated by lateral cyclic loading utilizing 254 available examples achieved an R2 value of 0.96 for failure mode classification and 0.98 for capacity prediction.

Construction management:

The number of accidents on a construction site is always on high rate. Thus, construction industry requires various techniques to manage the construction activities.

To differentiate construction documents based on various project factors, Soibelman and Caldas applied the support vector machine (SVM) algorithm. It was used to upgrade the quality of the existing construction information management system [24].

Chen used the KNN classification algorithm to build a model of awareness and information exchange [25].

The cost of construction projects and a schedule using an artificial neural network (ANN) and support a vector machine (SVM) was projected by Fang, et al. [26].

Using neural networks and linear regression, Elfahham [27] estimated a construction cost index for concrete buildings based on historical reports of the main construction costs. Elfahham [27] analysis key contribution was to provide stakeholders with a credible method for predicting prices for future project developments.

In Nguyen and Nguyen [28] research, in-depth talks with industry experts and obtained correlation relationships to find key socioeconomic variables that influence the construction price index in Vietnam.

Using Python language and scikit-learn library, the construction price index prediction model based on K-Nearest Neighbors was developed using raw data of seven socio-economic variables [14].

The study of [29] give a recurrent neural network model to make medium-to long-term forecast (i.e., the time horizon of the week, of one-hour electricity consumption profiles in commercial and residential buildings).

The study of [31] investigated how the deep-recurring neural network model predicts heating demand for campus buildings at the University of Utah over a period of weeks. The researchers then generated an optimization methodology that could provide conclusive advice on the size of a stratified thermal storage tank without requiring high-performance computational resources.

The study of zin et al. given the development of a ML model for delay risk assessment in tall building projects. In this analysis, the techniques such as Artificial Neural Networks (ANN), K-Nearest Neighbors (KNN), Support Vector Machines (SVM), and Ensemble methods were taken to develop a dataset suitable for ML application [30].

Hydrology and water resource management:

Liao and Sun [32] implemented a model to forecast the WQ of China's Chao Lake by pairing the ANN and decision tree algorithm.
Solanki et al. [33] examined and predicted the chemical eigenvalues of water, especially dissolved oxygen and pH using the deep learning network model.
Yan et al. [34] suggested a genetic algorithm (GA) and particle swarm optimization (PSO) algorithm to enhance the back propagation (BP) neural network to predict the oxygen demanded in a lake.
Shafi et al. [35] suggested four machine learning algorithms, namely, Deep Neural Networks, k-Nearest Neighbors (kNN), Support Vector Machines (SVM), Neural Networks (NN) for the prediction of water quality.
Abyaneh [36] has applied the machine learning approaches like ANN and regression to predict the chemical oxygen demand (COD).
Nikoo&Mahjouri (2013) invented a Probabilistic Support Vector Machines (SVMs) model associated with GIS technique for planning the classification and distribution of surface and groundwater in Iran [37].

Environmental engineering:

GA was utilized in a water quality conflict resolution model for a reservoir management system. The system model incorporated both physical parameters, such as pollutants, water demand, and water availability, in addition to human factors which influence decision making [9].
Tree based supervised models like random forest algorithms can help in making anticipations (predictions) in detecting the facilities are likely to fail an inspection and utilise the anticipated risk scores to recommend deputy surrogate inspections [38]
Decision tree learning , neural network learning are widely utilized in building reasonably precise and efficient learning models for handling scenarios like downscaling, and rainfall management [39].
AI holds huge benefits for soil science in the form of soil resistance, soil nutrition, soil moisture, soil health and many. When it comes to soil moisture, precise prediction helps in flood and drought surveillance and forecasting. Estimation and classification of soil moisture relies on assorted factors, like vegetation, soil surface, vegetation and roughness [40].
The combination of Artificial Neural Networks (ANN) and Support Vector Machines (SVM) can be used to predict soil penetration resistance based on the composition of clay and sand, soil and particle density, sand and silt and soil volumetric moisture The ANN model is further connected to a multilayer perceptron (MLP) and suitable back propagation algorithms can be applied to predict resistance penetration [41].
AI based agriculture controls the surplus of water, herbicides and pesticides, preserves soil fertility and helps the farmers in improving the quality and increased productivity of crops [42].

Transportation engineering:

Implementing the traffic sensors on the road that detects accidents automatically and predicts future traffic conditions precisely. ANNs can be used for road planning, public transport, Traffic Incident Detection; and predicting traffic conditions [43].

[44] focused on finding a good transportation system management and safety plan to balance the transport demand with a proper distribution of highways, railways, and airways.

Previous researches in the 90s focused on NNs for designing, road planning and modelling. For example, reference [45] developed the spatial relationship between transportation and land-use planning by using a parallel NN system.

Result and discussion:

In this paper, 10 algorithms are studied with its applications in civil industry. ML is growing field as it does not required much coding. Algorithms which are used for developing any model will take the dataset and make rules accordingly. So the developer does not have to take care of every condition to do the programming part. Also, these algorithms are easily adaptable and can handle complex multi-dimensional problems. Every algorithm has its own advantages and disadvantages. Using any type of algorithm depends on problem statement and provided features to the model. The table below represents used algorithms in the past for solving problem in particular industry which are listed in section 3.

Table 1: List of widely used algorithms in civil engineering

	Structural engineering	Construction management	Hydrology and water resource management	Environmental engineering	Transportation engineering
Extra tree regressor					
Random forest					
Support vector machines	Y	Y	Y	Y	
XGBoost regressor	Y				
Decision tree regressor	Y	Y			
Linear regression		Y			
Knearest neighbour regressor	Y	Y	Y		
AdaBoost	Y				
Artificial Neural Network	Y	Y	Y	Y	Y

Conclusion:

According to the research, ANN is widely used in most of the sectors. To get the optimum solution for a problem, different algorithms like SVM, DT, KNN can be used. ML is gaining popularity because of its advantages over rest of the techniques. Thus, to solve engineering problems ML is effective tool.

References:

1. Fidelis C. Obodoeze , Chris A. Nwabueze, Silas A. Akaneme Comparative Evaluation of Machine Learning Regression Algorithms for PM2.5 Monitoring.

2. Raffaele Cioffi, Marta Travaglioni, Giuseppina Piscitelli, Antonella Petrillo and Fabio De Felice Review Artificial Intelligence and Machine Learning Applications in Smart Production: Progress, Trends, and Directions
3. K. Smarsly, D. Hartmann Artificial intelligence in structural health monitoring.
4. Mahind Rupali, Patil Amit A Review Paper on General Concepts of "Artificial Intelligence and Machine Learning".
5. Huu-Tai Thai Machine learning for structural engineering: A state-of-the-art review
6. Jameela Ali Akrimi, Abdul Rahim Ahmad, Loay E. George, Sherna Aziz Review of Artificial Intelligence
7. Nguyen Van Tam, Nguyen Quoc Toan Research Trends on Machine Learning in Construction Management: A Scientometric Analysis
8. Hongwei Song, Ayaz Ahmad, Furqan Farooq, Krzysztof Adam Ostrowski, Mariusz Ma´slak, Slawomir Czarnecki, Fahid Aslam Predicting the compressive strength of concrete with fly ash admixture using machine learning algorithms
9. Catherine Wilcox, Wei Lee Woon, Zeyar Aung Applications of machine learning in environmental engineering
10. HeejaeAhna ,DongminLeeb , SeongsooLeec , Taehoon Kimd , Hunhee Choe , and Kyung-In Kangf Application of Machine Learning Technology for Construction Site
11. Tharsanee Raman Maganathan, Soundariya Ramasamy Senthilkumar, Vishnupriya Balakrishnan Machine Learning and Data Analytics for Environmental Science: A Review, Prospects and Challenges
12. Mohammed Majeed Hameed, Mohamed Khalid AlOmar, Faidhalrahman Khaleel and Nadhir Al-Ansari An Extra Tree Regression Model for Discharge Coefficient Prediction: Novel, Practical Applications in the Hydraulic Sector and Future Research Directions
13. Henry V. Burton, S.E., Ph.D., and Michael Mieler, Ph.D. Machine Learning Applications Hope, Hype, or Hindrance for Structural Engineering
14. Phong Thanh Nguyen Application Machine Learning in Construction Management
15. H. Lange and S. Sippel Machine Learning Applications in Hydrology
16. De-Cheng Fenga, Zhen-Tao Liub, Xiao-Dan Wang Yin Chen, Jia-Qi Chang, Dong-Fang Wei, Zhong-Ming Jiang Machine learning-based compressive strength prediction for concrete: an adaptive boosting approach
17. Akshata Patil, Lata Patted, Mahesh Tenagi, Vaishnavi Jahagirdar, Madhuri Patil, Rahul Gautam Artificial Intelligence as a Tool in Civil Engineering – A Review
18. Pengzhen Lu, Shengyong Chen and Yujun Zheng Review Article Artificial Intelligence in Civil Engineering
19. J.P. Amezquita-Sancheza, M. Valtierra-Rodrigueza and H. Adelib Machine learning in structural engineering
20. Abdullah, S., and Wallace, J. (2019). Drift capacity of reinforced concrete structural walls with special boundary elements, ACI Structural Journal 116(1), 183-194
21. Ibrahim, A., Eltawil, A., Na, Y., and El-Tawil, S. \A machine learning approach for structural health monitoring using noisy data sets", IEEE Transactions on Automation Science and Engineering, 17(2), pp. 900{908 (2019).
22. Honglan Huang, and Henry V Burton 2019 Classification of in-plane failure modes for reinforced concrete frames with infills using machine learning Journal of Building Engineering
23. Feng Jia, YaguoLei, JingLin, XinZhou, and NaLu 2015 Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data Mechanical systems and signal processing
24. Caldas, C. H., &Soibelman, L. (2003). Automating hierarchical document classification for construction management information systems. Automation in Construction, 12(4), 395-406.
25. Chen, J. H. (2008). KNN based knowledge-sharing model for severe change order disputes in construction. Automation in Construction, 17(6), 773- 779.
26. Xie, S., & Fang, J. (2018). Prediction of construction cost index based on multi variable grey neural network model. International Journal of Information Systems and Change Management, 10(3), 209-226.
27. Elfahham, Y. (2019). Estimation and prediction of construction cost index using neural networks, time series, and regression. Alexandria Engineering Journal, 58(2), 499-506.
28. Nguyen, P. T., & Nguyen, Q. L. H. T. T. (2020). Critical factors affecting construction price index: An integrated fuzzy logic and analytical hierarchy process. The Journal of Asian Finance, Economics, and Business, 7(8), 197-204.

29. Rahman, A., V. Srikumar, and A.D. Smith, Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks. Applied energy, 2018. 212: p. 372-385.
30. Sanni-Anibire, M.O., R.M. Zin, and S.O. Olatunji, Machine learning model for delay risk assessment in tall building projects. International Journal of Construction Management, 2020.
31. Rahman, A. and A.D. Smith, Predicting heating demand and sizing a stratified thermal storage tank using deep learning algorithms. Applied Energy, 2018
32. H. Liao and W. Sun, "Forecasting and evaluating water quality of Chao Lake based on an improved decision tree method," Procedia Environmental Sciences, vol. 2, pp. 970– 979, 2010.
33. Solanki, H. Agrawal, and K. Khare, "Predictive analysis of water quality parameters using deep learning," International Journal of Computers and Applications, vol. 125, no. 9, pp. 29–34, 2015.
34. J. Yan, Z. Xu, Y. Yu, H. Xu, and K. Gao, "Application of a hybrid optimized BP network model to estimate water quality parameters of Beihai Lake in Beijing," Applied Sciences, vol. 9, no. 9, p. 1863, 2019.
35. U. Shafi, R. Mumtaz, H. Anwar, A. M. Qamar, and H. Khurshid, "Surface water pollution detection using internet of things," in 2018 15th International Conference on Smart Cities: Improving Quality of Life Using ICT & IoT (HONET-ICT), pp. 92–96, Islamabad, Pakistan, October 2018.
36. H. Z. Abyaneh, "Evaluation of multivariate linear regression and artificial neural networks in prediction of water quality parameters," Journal of Environmental Health Science and Engineering, vol. 12, no. 1, p. 40, 2014.
37. Nikoo, M. R. &Mahjouri, N. 2013 Water quality zoning using probabilistic support vector machines and self-organizing maps. Water Resour. Manage.
38. Hino, M., Benami, E. & Brooks, N.2018 Nat Sustain 1 pp.583–588
39. Manish Kumar Goyal, Chandra S. P. Ojha, and Donald H. Burn 2017 Sustainable water resources management pp.165-178.
40. Qiangqiang Yuan, Huanfeng Shen, Tongwen Li, Zhiwei Li, Shuwen Li, Yun Jiang, Hongzhang Xu, Weiwei Tan, Qianqian Yang, Jiwen Wang Jianhao Gao, Liangpei Zhang 2020 Remote Sensing of Environment 241.
41. Pereira, Tonismar dos S. et al 2018 Eng. Agríc. [online].38, pp.142-148.
42. Tanha Talaviya, Dhara Shah, Nivedita Patel, Hiteshri Yagnik, Manan Shah 2020 Artificial Intelligence in Agriculture 4, pp. 58-73.
43. RusulAbduljabbar, Hussein Dia, Sohani Liyanage and Saeed AsadiBagloee Applications of Artificial Intelligence in Transport: An Overview
44. Dogan, E.; Akgüngör, A.P. Forecasting highway casualties under the effect of railway development policy in ˇ Turkey using artificial neural networks. Neural Comput. Appl. 2013, 22, 869–877.
45. Rodrigue, J.P. Parallel modelling and neural networks: An overview for transportation/land use systems. Transp. Res. Part C Emerg. Technol. 1997, 5, 259–271