

Comparison of the Performance of K-Nearest Neighbours and Generalized Neural Network in Construction Crew Productivity Prediction

Okyanus ORAL¹, Emel LAPTALI ORAL^{*2}, Mehmet Sait ANDAÇ³

¹Middle East Technical University, Electrical-Electronics Engineering Department, BSc Student, Ankara

²Çukurova University, Civil Engineering Department, Adana

³Alparslan Mah. Uzay Cad. 40/17 Melikgazi, Kayseri

Geliş tarihi: 01.02.2021 Kabul tarihi: 31.03.2021

Abstract

Construction crew productivity prediction is one of the most important issues that affect the realistic prediction of construction duration and cost. Use of different search algorithms like Feed Forward Neural Network, Ant Colony, Artificial Bee Colony, Particle Swarm Optimization, Radial Based Neural Networks and Self Organizing Maps for crew productivity prediction problem have been discussed in previous studies. However, the significant effect of the coherence between the nature of the data and the characteristics of the method used in prediction performance has generally been neglected. The aim of the current research thus has been to analyse the prediction performance of two contemporary learning algorithms; K- Nearest Neighbour (K-NN) and Generalized Neural Network (GRNN) when applied to three different crew (formwork, tiling and masonry) productivity related data sets with different distribution characteristics. Performance of both methods varied with the changing coefficient of variation values. K-NN outperformed GRNN for all data sets and both of the methods had their worst performance on the dataset with the highest variance.

Keywords: Construction crew, Productivity, Artificial neural networks, K-nearest neighbour, Generalized neural network

İnşaat Ekibi Üretkenlik Tahmininde K-En Yakın Komşu ve Genelleştirilmiş Sinir Ağının Performansının Karşılaştırılması

Öz

İnşaat işlerinde ekip verimliliğinin tahmini, inşaat süresi ve maliyetinin gerçekçi tahminini etkileyen önemli faktörlerden biridir. Ekip verimliliği tahmini için İleri Besleme Sinir Ağı, Karınca Kolonisi, Yapay Arı Kolonisi, Parçacık Sürü Optimizasyonu, Radyal Tabanlı Sinir Ağları ve Kendi Kendini Düzenleyen Haritalar gibi farklı arama algoritmalarının kullanımı önceki çalışmalarda tartışılmıştır. Ancak, bu çalışmalarda tahmin performansında kullanılan yöntemin özellikleri ile verinin niteliği arasındaki tutarlılığın etkisi genellikle ihmal edilmiştir. Dolayısıyla mevcut araştırmanın amacı, iki çağdaş öğrenme algoritması olan K- En Yakın Komşu (K-NN) ve Genelleştirilmiş Sinir Ağı (GRNN) kullanılarak farklı dağılım özelliklerine sahip üç farklı ekibe (kalıp, döşeme ve duvar) ait verimlilikle

*Sorumlu yazar (Corresponding author): Emel LAPTALI ORAL, emellaptalioral@yahoo.com

ilgili veri seti için tahmin performansını analiz etmektir. Her iki yöntemin performansı da, değerlerin değişen katsayıları için farklılık göstermiştir. K-NN, tüm veri setleri için GRNN'den daha iyi performans göstermiş olup, her iki yöntem de en yüksek varyansa sahip veri kümesinde en kötü performansa sahiptir.

Anahtar Kelimeler: İnşaat ekibi, Verimlilik, Yapay sinir ağları, K-en yakın komşu, Genelleştirilmiş sinir ağı

1. INTRODUCTION

Precise estimation of duration and in turn cost of a project is one of the vital activities of construction project management as precise estimates together with a proper execution result in successful completion of the projects. Construction management related research has focused on two important issues related to the precise estimation of duration of construction activities. These are firstly determination and secondly prediction of the crew (labour) productivity (man-hour) values.

While man-hour/ crew productivity values can be determined by collecting related data through methods like time-study or work-study, and are used for prediction with a certain amount of confidence in industries where the production is based on repetitive tasks under steady working conditions, prediction becomes a problematic issue for construction works where project-based production leads to varying conditions related with labour, management, working space and the environment from one project to the other. Thus, man hour/crew productivity values determined during one project may be insignificant for the other. Research has focused on predicting crew productivity values for these varying conditions by using methods based on initially statistics and recently machine learning. While continuous developments in neural networks resulted in the application of many new methods to the problem, either focusing on the performance of a single prediction method or comparison of two or more methods by only focusing on one particular activity like formwork, tiling, and so on have not created a firm ground for future research. The significant effect of the coherence between the nature of the data and the characteristics of the method used in prediction performance has generally been neglected. The aim of the current

research thus has been to analyse the prediction performance of two different learning algorithms when the distribution characteristics of data are changed. In order to achieve the aim of the research, productivity data related to three different types of crews; i.e. ceramic tiling, brick wall building, and timber formwork, with different distribution characteristics, i.e. standard variation, coefficient of variation, coefficients of skewness, and kurtosis has been used.

2. LITERATURE REVIEW

Artificial Neural Networks (ANNs), unlike conventional programmes, learn from observational data and figure out their own solutions to the problems. They are a series of algorithms that recognize the underlying relationships in a set of data through a process that mimics the way the human brain operates. Since their first introduction by [1], ANNs have evolved through a broad family of learning algorithms that have advanced the state of the art across multiple domains. Feed Forward Neural Network (FFNN) was the first and has been the most commonly used ANN and [2] were the first to use FFNN for construction crew productivity prediction problem. The productivity of some formwork crews were predicted by training the FFNN with the data related to the place of work, skills of the superintendents, quantity of the formwork, thickness of the walls, degree of repetitions, number of reuses of the formworks, season, site conditions, material handling problems and project gross building area. Data was collected from a previously completed building project. Results were compared with the estimators' predictions and FFNN was reported to be superior in predicting formwork crew productivity rates than the estimators. [3] followed [2] and used FFNN in predicting the performance of not only formwork crews but also concrete

pouring and concrete finishing crews. Applying two variations of FFNN with different number of hidden layers let the researchers to undertake sensitivity analysis related to the effect of variables like crew size, weather conditions (temperature, precipitation, humidity) and amount of work done on prediction performance of FFNNs. Data was collected from eight building projects and results were compared with regression analysis results. [4] then applied FFNN to the prediction of pipe installation crew productivity and like [2] compared the results with the estimators' results. After these pioneering studies, FFNN has not lost its popularity over time. While researchers like [5] and [6] still focused on the performance evaluation

of FFNN by applying it to formwork and concrete crews respectively, some researchers introduced more contemporary methods like Generalized Neural Network (GRNN) [7,8], Self-Organizing Maps (SOM) [8,9], Radial Basis Neural Network (RBNN) [10], Artificial Bee Colony (ABC) [11, 12] for construction crew productivity problem and mostly used FFNN as a comparison benchmark (see Table 1). As summarised in Table 1, while input variables and the type of crews changed depending on the availability of real-life data during each research, method selection depended mainly on the academic trends of the time of research.

Table 1. Literature findings related to ANN applications in construction crew prediction problem

Reference	Nature of Data Crew Type/ Input Variables	Method Proposed/Alter native Method	Superior method/ Results
1) Postas and AbouRizk (1997)	Formwork crews from one particular building project / place of work, skills of the superintendent, quantity of formwork, thickness of the wall, degree of repetition, number of reuses, season, site conditions, material handling problems, project gross building area.	FFNN /Estimators' predictions	FFNN/ Prediction values lying within the interval of $\pm 5\%$ of the actual values.
2) Sönmez and Rowings (1998)	Concrete pouring, formwork, concrete finishing crews from eight building projects/ crew size, temperature, precipitation, humidity, amount of work done.	FFNN/ Multiple Regression Analysis	FFNN / Best MAPE values of 71.7% and 68.1% for concrete pouring and formwork crews, respectively.
3) AbouRizk et al. (2001)	Pipe installation crews from twenty-seven projects/ location, state/province, administrative requirements, year of construction, quality requirements, degree of difficulty.	FFNN/Estimators' predictions	FFNN/ Predicted values within the interval of $\pm 15\%$ of the actual values for 84% of the cases.
4) Dissanayake et al. (2005)	Hydro testing of pipe fabrication crews / no. of pipe modules in progress, equipment availability, no. of cranes, manpower availability, ratio between no. of pipefitters and no. of modules, temperature, precipitation, rework, quality control/testing ratio.	A hybrid neural network combining the GRNN, Fuzzy Logic (FL) and Genetic Algorithms (GA)/ Real activity data	MAE equals to 0,018, MSE equals to 0,001.

Comparison of the Performance of K-Nearest Neighbours and Generalized Neural Network in Construction Crew Productivity Prediction

5) Ezeldin and Sharara (2006)	Formwork, steel fixing, concrete pouring crews from six projects.	ANN/Real life data	Best MAPE values of 44%, 38%, and 42% for formwork assembly, steel fixing, and concrete pouring, respectively.
6) Oral and Oral (2010)	Concrete pouring, reinforcement, formwork /crew size, experience of the crew on the particular site, age of the crew members, payment system, weekly working hours, daily working hours, accommodation distance of the crew members from the site.	SOM/ Multiple Regression Analysis	SOM/ Optimum values when crew size, experience of the crew on the particular site, age of the crew members, payment system is used as an input. Best MAPE values of 25.68%, 38.04%, 25.05% for concrete pouring, reinforcement, formwork crews, respectively.
7) Oral et. al (2012)	Plastering crews/ crew size, experience of the crew on the particular site, age of the crew members.	SOM /GRNN, FFNN	SOM/ Best MAPE values of 41,27%, 45,87% and 45,97% for SOM, GRNN and FFNN, respectively.
8) Gerek et al. (2015)	Masonry crews /crew size, experience of the crew on the particular site, age of the crew members, payment system, weekly working hours, daily working hours, accommodation distance of the crew members from the site, experience of the crew members in the same team, accommodation type, materials used.	RBNN/ FFNN	RBNN / Best MAPE values of 10,152% and 14,942% for RBNN and FFNN, respectively.
9) Dikmen and Sönmez (2015)	Formwork crews data from twenty-two different concrete building projects)/ building size, climatic conditions, work culture, building height, working height.	FFNN/Real life data	FFNN/ min.11% , max. 22% variation from the actual field data.
10) Heravi and Eslamdoost (2015)	Concrete foundation works in two real power plant construction projects / labour competence, poor decision making, motivation of labour, suitable site layout, and proper planning w	FFNN trained with back propagation learning algorithm using Bayesian Regularization and Early stopping methods.	Training based on Bayesian Regularization gives better results in small sized data sets.
11) Oral et al. (2016)	Ceramic tiling crews/ Crew sizes, age distribution of crew members, payment methods.	SOM/ABC	SOM / Best MAPE values of 22,54%, 39,96% for SOM and ABC, respectively.
12) Andaç, Oral (2019)	Plastering crews / crew size, experience of the crew on the particular site, age of the crew members.	ABC/FFNN	ABC / Best MSE values of 0,0523 and 0,0583 for ABC and FFNN, respectively.

3. MATERIAL AND METHODOLOGY

3.1. Material

As the aim of the research was to analyse and discuss the relationship between the distribution characteristics of data (like coefficient of variation, skewness and kurtosis) and performance of the methods, three different data sets were chosen from a group of data sets composed through time studies undertaken by 288 different construction crews (121 brick wall building, 100 formwork, 67 ceramic tiling crews) working around 200 construction sites in various regions of Turkey. Standard time study sheets designed by considering the details of each work and crew information sheets were used in order to collect data consistently. Age, accommodation type (on site or not), method of payment(daily, monthly, lump sum), experience duration on site and size of the crews were selected as the input (independent) variables depending on the previous experience of the researchers discussed in various studies [9–11]. Output (dependent variable) was—the productivity value (hr/m²) of the crews recorded during time study. Distribution characteristics of the productivity values for each crew type are summarised in Table 2.

3.2. Methods Used K- Nearest Neighbour (K-NN) and Generalized Regression Neural Network (GRNN)

While K-NN and GRNN are reported and used to be two different prediction methods, they are both based on Radial Basis Function Kernel (RBF Kernel) network and both do not require an iterative training procedure as back propagation networks. They approximate any arbitrary function between input and output data set, drawing the function estimate directly from the training data [13] Meanwhile the main difference between these two methods is that; while GRNN uses all of the

training data sets for prediction of the output of a query, K-NN uses k number of data sets nearest to the query data set for prediction [14]. Like GRNN, K-NN also finds the distances between a query and all the examples in the training data set but unlike GRNN it only selects the specified number (k) of examples closest to the query and determines the similarities of only K- nearest neighbour to the query. In GRNN this similarity is calculated between the whole training data set and the sample of query. Finally, both methods predict the output of the query by considering the calculated similarity values. Following steps for further explanation of the application of the networks to construction crew productivity prediction problem.

Step 1) Preparing the Data Sets for Training and Cross Validation:

Original data sets of the three crew types (i.e. brick wall, formwork and ceramic tiling) were initially divided into 5 randomly formed sets (each containing 20% of the original data) and each neural network was trained by feeding the network with 4 of the 5 sets (i.e. 80% of the data) at a time leaving other 20% for validation. Thus, the following steps were undertaken five times and performance of the neural networks were then evaluated accordingly, as discussed below.

Step 2) Training the Neural Networks:

As stated above, each neural network was fed by 80% of the total samples which comprised 97, 80 and 54 sample data sets for brick wall, formwork and ceramic tiling crews, respectively. Input variables were age, accommodation type, method of payment and experience of the crew members on the particular site, which in turn determined the number of input nodes, i.e. neurons (n) to be 4.

Table 2. Distribution characteristics of data

Formwork	
Man Hour (hr/m ²)	0,87
Standart Deviation (σ)	0,85
Coefficient of Variation (V)	0,97
Coefficient of Skewness	2,82
Coefficient of Kurtosis	9,57
Ceramic Tiling	
Man Hour (hr/m ²)	0,68
Standart Deviation (σ)	0,49
Coefficient of Variation (Değişim katsayısı) (V)	0,72
Coefficient of Skewness	2,05
Coefficient of Kurtosis	4,3
Brick Wall	
Man Hour (hr/m ²)	0,69
Standart Deviation (σ)	0,35
Coefficient of Variation (V)	0,5
Coefficient of Skewness	1,39
Coefficient of Kurtosis	2,21

Step 3) Deciding the k Value:

As pointed out above, GRNN covers all the training data set for prediction, in fact resulting in the value of k to be equal to the number of data samples in the training set. It is stressed in the literature [18] that the value of k may have a strong effect on the performance of K-NN, with small values resulting in smaller biases and higher variances and large values and vice versa, which also stresses the potential differences between the prediction performances of K-NN and GRNN when applied to the same data set. In order to analyse the effect of the varying values of k on prediction accuracy, k values between 1 to 30 were tested during the application of K-NN.

Step 4) Calculating the Distance and the Similarity Between the Query and the Training Data Set:

One needs to calculate the distance between the query and the training data set in order to be able to determine their similarities. While there are different distance measures like Euclidean Distance, Manhattan Distance and Angular Similarities [15–17]. Euclidean Distance, i.e. the most popular similarity measure, was used both

during the application of GRNN and K-NN in the current study.

$$ED(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

where;

ED (x,y): Euclidean Distance between the sample of query (y) and a particular sample (x) of the training set

n: number of independent variables in one sample (input nodes) (n=4 for the current research)

x_i : value of the i^{th} independent variable in the particular sample of the training set

Y_i : value of the i^{th} independent variable in the sample of query

While GRNN uses all of the calculated distances for similarity determination, K-NN requires selection of k nearest distances, i.e. neighbours, to the query to be used. The similarity of each sample to the query is calculated using RBF Kernel (see Equation 2), which outputs a similarity value between 0 (complete opposite) and 1 (exactly the same).

$$S(x, y) = \exp\left[-\frac{ED(x, y)^2}{2\sigma^2}\right] \quad (2)$$

The hyper-parameters of RBF Kernel are:

$S(x, y)$: Similarity value between a particular sample (x) and the sample of query (y)

$ED(x, y)$: Euclidean Distance between a particular sample (x) and the sample of query (y)

σ : the parameter for adjusting the gradient of associated similarity with respect to distance.

Like k value, σ is also selected by the researcher and its value has a strong effect on the similarity of a particular sample in the training set to the query. While a ‘too large’ σ value would result in every sample in the training sample to be equally similar to the sample of the query, a ‘too small’ σ value would result in ‘no similarity’ between any of the samples in the training data set and the sample of query. In order to analyse the effect of σ value on prediction performance of the two neural networks, σ values between 0.01 to 30.00 were tested consecutively.

Step 5) Predicting the Value of the Output of the Sample of Query:

The output of the sample of query, i.e. crew productivity is calculated by using Equation 3 which includes similarity values as weighted averages.

$$P(y) = \frac{\sum_{i=1}^k S(x, y) P(x)}{\sum_{i=1}^k S(x, y)} \quad (3)$$

$P(y)$: Productivity (output) value of the sample of query (y)

$S(x, y)$: Similarity value of a particular sample (x) to the query sample (y)

$P(x)$: Productivity value of a particular sample (x)

k : selected number of the nearest neighbours (for GRNN $k=n$)

Performance Measurements: In this study, results of the analyses are measured by using 3 different statistical error measurement techniques.

Mean Absolute Percentage Error (MAPE):

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{A_i - P_i}{A_i} \right| \quad (4)$$

Mean Square Error (MSE):

$$MSE = \frac{1}{n} \sum_{i=1}^n (A_i - P_i)^2 \quad (5)$$

Mean Absolute Error (MAE):

$$MAE = \frac{1}{n} \sum_{i=1}^n |A_i - P_i| \quad (6)$$

where;

n : Number of data sets used for estimation.

A_i : Actual value of the i^{th} element of the data set.

P_i : Predicted value of the i^{th} element of the data set.

4. RESULTS AND DISCUSSION

4.1. Prediction Performance of the Models

Discussion under Material and Methodology Section on the similarities and dissimilarities of K-NN and GRNN methods led the choice of data sets as stated in Table 2. As shown in the table, distribution characteristics of formwork, ceramic tiling and masonry crews varied. Productivity values between formwork crews had the largest variation. Depending on the nature of the data and the characteristics of the two methods, following were expected to be fulfilled after the methods were applied to these three different data sets. Due to the relatively high variation of productivity related data of formwork crews;

- 1) Both GRNN and K-NN are expected to have the worst performance for formwork crew productivity prediction

- 2) K-NN is expected to give better results than GRNN, especially for formwork crew productivity prediction
- 3) In K-NN application k value with the best (minimum) MAPE value is expected to be the largest for formwork crews, ceramic tiling crews and masonry crews following it, respectively.
- 4) σ values with best (minimum) MAPE values for both GRNN and K-NN applications are expected to be sorted for the masonry, ceramic tiling and formwork crews in ascending order.
- Table 3 shows the best MAPE, MAE and MSE values achieved when GRNN and K-NN were
- 1) Both GRNN and K-NN had the worst performance for formwork crew productivity prediction, as expected.
- 2) K-NN gave better results than GRNN for all of the crew types, as expected.
- 3) Optimum k value with minimum MAPE was the largest for formwork crew data set, and ceramic tiling and masonry crews followed it, as expected.
- 4) Optimum σ values in the GRNN application were as expected, but not as expected for the K-NN application.

Table 3. MAPE, MSE and MAE Values for GRNN and K-NN Applications

<i>K-NN</i>									
	<i>Formwork</i>			<i>Tiling</i>			<i>Masonry</i>		
	k	σ	err	k	σ	err	k	σ	err
MAPE	23	10	0,70032	17	0,15	0,134518	5	10	0,420937
MAE	28	10	0,347715	12	0,38	0,082592	5	10	0,255795
MSE	24	2,66	0,179473	12	0,48	0,063814	5	10	0,113705
<i>GRNN</i>									
	<i>Formwork</i>			<i>Tiling</i>			<i>Masonry</i>		
	k	σ	err	k	σ	err	k	σ	err
MAPE	23	10	0,470036	13	10	0,175702	5	0,59	0,415452
MAE	28	10	0,347726	17	10	0,087171	5	0,69	0,251965
MSE	24	10	0,179497	12	10	0,066764	30	0,53	0,11105

With σ values between 0,01 to 30,00, and k between 1 to 30 (for K-NN) tested consecutively, results show that K-NN gave better results than GRNN for all crew types. Best MAPE values in K-NN applications were for the k values equal to 23, 17 and 10 for formwork, floor tiling and masonry crews respectively which was in good agreement with the fact that; optimum k value varies depending on the data set and, the more homogenous the data, the smaller the value of optimum k is. Optimum MAE and MSE values also provided a similar comparison. When optimum σ values are analysed in the GRNN application, the results fit with the fact that the more homogenous the data, the larger the value of

optimum σ giving minimum MAPE is (ie. σ values 18, 9,95, 9,39 for brick wall, ceramic tiling and formwork crews, respectively). However, such a relationship is not observed for the K-NN application.

When the results are compared with the results of previous studies that were based on the similar data sets used during the current research, it is observed that;

- (1) for formwork crew productivity prediction, SOM [9] with 25,05% MAPE values had better results than both K-NN and GRNN with 47,00% and 142,36% MAPE values,

- (2) for masonry crew productivity prediction, RBNN and FFNN [10] with 10,15% and 14,92% MAPE values had better results than both K-NN and GRNN with 42,09% and 99,50 % MAPE values, respectively,
- (3) for ceramic tiling crew productivity prediction, SOM and ABC [11] with 22,54 % and 39,96% MAPE values had worse results than KKN with MAPE 13,45%, but better results than GRNN with MAPE 106,80%.

5. CONCLUSION

Previous research that focused on application of artificial neural networks to construction crew productivity problem inevitably followed academic trends of their time in computer science and usually ignored or had to ignore the significant effect of the coherence between the nature of the data and the characteristics of the method used in prediction performance. Recent developments in machine learning methods now allow any researcher to be able to select an appropriate method from different alternatives. K-NN and GRNN are two of these methods in which the relationship between the similarity related parameters like k , σ , $S(x,y)$, $ED(x,y)$ and nature of the data distribution are to some extent predictable and have significant impact on the model performance. Thus, while these two methods reduce the risk of overfitting, they have the complication of having to choose the best values of the similarity related parameters. In this research three different data sets with different distribution properties were used to analyse the performance of these methods. Performance of both methods varied with the changing coefficient of variation values and K-NN outperformed GRNN for all data sets as expected and both of the methods performed worst on the dataset with the highest variance, i.e. the formwork crew data set. While the relationship between the optimum values of parameter k with the nature of the data was as expected, optimum values of σ was as expected for GRNN only.

6. ACKNOWLEDGEMENTS

This paper is based on some of the findings of the research project 106M055 which is supported by TÜBİTAK (The Scientific and Technical Research Council of Turkey).

7. REFERENCES

1. McCulloch, W.S., Pitts, W., 1943 Learning Based Industrial Bin-picking Trained with Approximate Physics Simulator. Bull Mat Biophys. 5, 115–133. https://doi.org/10.1007/978-3-030-01370-7_61.
2. Portas, J., Abou Rizk, S., 1997. Neural Network Model for Estimating Construction Productivity. J Constr Eng Manag. 123, 399–410. [https://doi.org/10.1061/\(ASCE\)0733-9364\(1997\)123:4\(399\)](https://doi.org/10.1061/(ASCE)0733-9364(1997)123:4(399))
3. Sonmez, R., Rowings, J.E., 1998. Construction Labor Productivity Modeling with Neural Network. J Constr Eng Manag, 498–504.
4. AbouRizk, S., Knowles, P., Hermann, U.R., 2001. P Roduction R Eport/a Creage R Eport P Roduction R Eport/a Creage R Eport. J Constr Eng Manag. 127, 502–511.
5. Dikmen, U.S., Sonmez, M., 2011. An Artificial Neural Networks Model for the Estimation of Formwork Labour. J Civ Eng Manag, 17, 340-347. <https://doi.org/10.3846/13923730.2011.594154>.
6. Heravi, G., Eslamdoost, E., 2015. Building Information Modeling Education for Construction Engineering and Management. II: Procedures and Implementation Case Study. J Constr Eng Manag. 1–13. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862](https://doi.org/10.1061/(ASCE)CO.1943-7862).
7. Dissanayake, M., Fayek, A.R., Russell, A.D., Pedrycz, W., 2005. A Hybrid Neural Network for Predicting Construction Labour Productivity. Proc 2005 ASCE Int Conf Comput Civ Eng. 819–830. [https://doi.org/10.1061/40794\(179\)78](https://doi.org/10.1061/40794(179)78).
8. Oral, M., Oral, E.L., Aydin, A., 2012. Supervised vs. Unsupervised Learning for Construction Crew Productivity Prediction. Autom Constr. 22, 271–276. <https://doi.org/10.1016/j.autcon.2011.09.002>

9. Oral, E.L., Oral, M., 2010. Predicting Construction Crew Productivity by Using Self Organizing Maps. *Autom Constr*, 19, 791–797. <https://doi.org/10.1016/j.autcon.2010.05.001>.
10. Gerek, I.H., Erdis, E., Mistikoglu, G., Usmen, M., 2015. Modelling Masonry Crew Productivity Using Two Artificial Neural Network Techniques. *J Civ Eng Manag*, 21, 231–238. <https://doi.org/10.3846/13923730.2013.802741>.
11. Oral, E.L., Oral, M., Andaç, M., 2016. Construction Crew Productivity Prediction: Application of Two Novel Methods. *Int J Civ Eng*, 14, 181–186. <https://doi.org/10.1007/s40999-016-0009-2>.
12. Andac, M.S., Oral, E., 2019. Crew Productivity Prediction by Using Artificial Bee Colony and Levenberg-Marquardt Algorithms-Comparison of Performances. *Int Civ Eng Archit Conf*, 29-44.
13. Firat, M., Gungor, M., 2009. Generalized Regression Neural Networks and Feed Forward Neural Networks for Prediction of Scour Depth Around Bridge Piers. *Adv Eng Softw*, 40, 731-737. <https://doi.org/10.1016/j.advengsoft.2008.12.001>.
14. Harikumar, S., Aravindakshan Savithri, A., Kaimal, R., 2019. A Depth-based Nearest Neighbor Algorithm for High-dimensional Data Classification. *Turkish J Electr Eng Comput Sci*, 27, 4082–4101. <https://doi.org/10.3906/ELK-1807-163>.
15. Prasath, S., Abu Alfeilat, H., Lasassmeh, O., Hassanat, A.B.A., Tarawneh, A.S., 2017. Distance and Similarity Measures Effect on the Performance of K-Nearest Neighbor Classifier -A Review. 1–39. <https://doi.org/10.1089/big.2018.0175>.
16. Mulak, P., Talhar, N., 2015. Analysis of Distance Measures Using K-Nearest Neighbor Algorithm on KDD Dataset. *Int J Sci Res*, 4, 2319–7064.
17. Yu, X.G., Yu, X.P., 2008. A New K-nearest Neighbor Searching Algorithm Based on Angular Similarity. *Proc 7th Int Conf Mach Learn Cybern ICMLC 3*, 1779–1784. <https://doi.org/10.1109/ICMLC.2008.4620693>.
18. A Complete Guide to K-Nearest-Neighbors with Applications in Python and R, <https://kevinzakka.github.io/2016/07/13/k-nearest-neighbor/#more-on-k>, Jul 13, 2016.