Application of Machine Learning in Predicting Key Performance Indicators for Construction Projects

Seyedeh Sara Fanaei[1], Osama Moselhi[2], Sabah T. Alkass[2], Zahra Zangenehmadar[3]

[1]Ph.D. Candidate, Dept. of Building, Civil and Environmental Engineering, Concordia Univ., Montreal, QC, Canada H3G 1M8
[2]Professor, Dept. of Building, Civil and Environmental Engineering, Concordia Univ., Montreal, QC, Canada H3G 1M8.
[3]Post-Doc Fellow, School of Information Technology and Engineering, University of Ottawa, Ottawa, ON, Canada K1N 6N5

---***---

Abstract - *Different projects aspects are evaluated by Key Performance Indicators (KPIs) which are utilized for monitoring and controlling construction projects. While there is considerable work on applying different techniques to forecast construction project performance, few studies focus on predicting KPIs qualitatively during construction projects. This research applies and compares different machine learning techniques to predict whole project KPIs qualitatively in critical project stages. Artificial neural network (ANN), as well as the neuro-fuzzy technique using fuzzy C-means (FCM) and subtractive clustering, were used to predict project KPIs. Models were developed to map the KPIs of three critical project stages to the whole project KPIs. Validation was done using the data of real projects to confirm models' effectiveness and to compare the results of the employed machine learning techniques. Mean absolute percentage error (MAPE) models developed using neuro-fuzzy technique with subtractive clustering for six KPIs are 1.1% to 8.3% lower compared to ANN technique. Thus, Results show that neuro-fuzzy technique with subtractive clustering have better prediction ability in this study.*

Key Words: Key Performance Indicators (KPIs), Neuro-Fuzzy, Artificial Neural Network, Performance Forecasting, Construction Project

1. INTRODUCTION

Performance is defined as the amount of efficiency and effectiveness in all project objectives [1]. Key Performance Indicators (KPIs) are a benchmark for assessing project performance. In the construction industry, project performance prediction is vital for monitoring and controlling construction projects. In Rethinking Construction, published as a report in 1998, Egan [2] launched the KPI method for measuring performance. However, additional KPIs needed to be recognized to understand the true status of a project [3]. Defense Construction Canada (DCC) offers guidance on the development, measurement, and reporting of KPIs for evaluating the success of Canadian projects [4].

Freeman and Beale [5] measured construction performance using seven criteria of technical performance, the efficiency of project execution, managerial and organizational expectations, personal growth, project termination, technical innovativeness, and manufacturability and business performance. The limitation is that they need complicated information that may not be calculated until a couple of years after project completion, and the ways that criteria's are defined is not inclusive enough to cover all aspects of the project. Chua, et al. [6] measured project achievement using three objectives: cost, schedule, and quality. Sixty-seven critical success factors (CSFs) that influence the performance of the three objectives and overall project success are defined by a survey using expert's opinion. This model does not give a special way of measuring the three objectives and also does not consider other criteria that affect the success of a project such as safety and client satisfaction. Both the client and contractor's point of view were considered by Tucker, et al. [7] in a model for quantifying construction phase success. However, it does not consider criteria such as client satisfaction and is not applicable to the project and can only be used when the project is finished, which is too late to carry out corrective actions.

Li, et al. [8] used the forecasting method for predicting potential cost overrun and schedule delay in construction projects using a set of performance indicators by employing a fuzzy inference process. A model for monitoring performance and identifying the causes of performance failures for construction projects was proposed by Dissanayake and Fayek [9] using fuzzy, ANN, and Genetic Algorithm (GA) methods. The proposed approach is daily-based and for individual activities and does not consider project level. Cha and Kim [10] developed a mathematical model for quantitative performance measurement of residential building projects using 18 KPIs. Haponava and Al-Jibouri [11] designed a system for measuring process performance in three main project stages. Their proposed measurement system considers a number of process-based KPIs. In this model, the KPIs are not comprehensive, and some criteria are not mentioned. Moreover, the performance of the project is not measured during the project. A quantitative approach to measuring project performance from the contractor's point of view was developed by Nassar and AbouRizk [1] during the construction phase of projects. Case studies are needed to validate the proposed technique. Ngacho and Das [12] developed a framework based on six KPIs of time, cost, and quality, safety, minimum site disputes and environmental impact for construction project performance assessment. Oral, et al. [13] applied two methods of self-organizing maps (SOM) and artificial bee colony (ABC) to predict construction crew performance and compared their results using MAPE, MAE and MSE values. Reenu, et al. [14] forecasted project performance using four performance metrics: prediction of cost performance, schedule performance, quality performance, and satisfaction level by developing an ANN technique.

Leon, et al. [15] developed a system dynamics (SD) model for forecasting project performance using eight construction indicators of cost, schedule, quality, profitability, safety, environment, team satisfaction, and client satisfaction. The model emphases on the construction phase of projects for using by contractors. Shaikh and Darade [16] focused on quality of activities by considering KPIs in planning stage. This study tried to find KPIs of activities and prepared a Project Quality Plan (PQP) for activities and their importance. However, this research did not predict performance and only focus on quality indicators without considering other KPIs. Nilashi, et al. [17] applied machine learning techniques to develop a hybrid intelligent system for predicting heating and cooling loads of residential buildings. Adaptive Neuro-Fuzzy Inference System was compared with other techniques for predicting buildings' energy performance. The results indicate better prediction accuracy when using neuro-fuzzy. However, neuro-fuzzy has not been previously utilized for predicting construction performance indicators

Based on the literature, some limitations were identified. First, most studies focus on quantitative project performance forecasting and less attention has been directed toward qualitative methods, although some construction KPIs are qualitative in nature and cannot be measured quantitatively. Therefore, this research aims to measure and predict the performance of construction projects qualitatively. Second, most of the previous research measured performance only after the project was completed and not during the construction phase. The benefit of measuring performance during the project is that stakeholders can suggest corrective actions and predict remaining project performance.

Therefore, the main objective of this research is to define a framework for measuring and forecasting construction project performance qualitatively. To reach this objective, a framework for qualitatively measuring and forecasting performance was defined during construction projects. The performance evaluation process was then formalized by defining a set of key performance indicators. Ultimately, different soft computing techniques were applied to forecast the performance for rest of the project.

2. RESEARCH METHODOLOGY

The first step in this research methodology was to define the project objectives and goals. Objectives or goals provide a sense of direction to the project management team. Defining the objectives means that the attention of the team directed toward priorities in order to better monitor progress during the construction phase [18]. Because of the different benefits for the different stakeholders, from whose point of view project success is defined should be specified.

Then, the main project KPIs were identified through both the literature review and expert opinion. A list of the KPIs used in the literature at the project level was identified based on their frequency found in 34 references [19]. To develop the model, the six KPIs used were referenced by approximately 50 percent or more of the studies; these include cost, time, quality, safety, client satisfaction, and project team satisfaction. The expert questionnaire results confirmed the importance of these six KPIs, which got 40 to 90 percent of the score, with the seventh KPI receiving a score of 27, and the remaining KPIs receiving a score of 20 to 27 percent. Given the tangibly larger score ratio between KPI number six and seven in the questionnaire, it was concluded that the experts found the first six KPIs more important than the others. Based on the literature review and the questionnaire, this study decided to select the first six KPIs since these KPIs also comply with previous studies [1, 4, 9]. These six KPIs were used to design the questionnaire for data collection. All indicators were measured considering the profits and damages to the owner qualitatively.

All KPI data was collected using a questionnaire. A qualitative method using a 1 to 7 scale based on a method proposed by Dissanayake and Fayek [9] was used for measuring the KPIs. The project was assumed to have three stages: initial stage (0 to 30 % physical progress), middle stage (30% to 70% physical progress) and finishing stage (70% to 100% physical progress). The questionnaire was designed using the above scale to collect the KPIs for three critical project stages. The prediction models for forecasting KPIs were developed in this step.

A framework was proposed for measuring and forecasting a set of qualitative KPIs using different machine learning techniques. This research applied the neuro-fuzzy as well as the neural network technique to forecast KPIs of building construction projects. The neuro-fuzzy technique was applied with both FCM and subtractive clustering.

2.1 Neuro-fuzzy technique

A novel framework for qualitatively measuring and predicting six important construction project KPIs using neuro-fuzzy technique was proposed. To map the KPIs of three critical project stages to the whole project KPIs, neuro-fuzzy models were developed. In the proposed framework, the neuro-fuzzy technique was applied to forecast the whole project KPIs automatically from data. The neuro-fuzzy technique is a combination of ANN and fuzzy logic, which is used in resolving different research problems in construction management. Furthermore, an advantage of the neuro-fuzzy technique is that it does not assume a pre-defined mathematical expression. In addition, it also captures each variable's effect on the output without requiring a priori knowledge [20].

The neuro-fuzzy technique provides only one output [20]. Therefore, in this research, multiple neuro-fuzzy models were developed to predict each whole project performance indicator for three critical project stages. Six different neuro-fuzzy models were developed for the initial stage. Each model in this stage had six inputs and one output. For each of the middle and finishing stages, six different neuro-fuzzy models were developed as well, resulting in a total of 18 neuro-fuzzy models. Development of neuro-fuzzy models consisted of two main steps: the development of an initial Fuzzy Inference System (FIS) and the optimization of the initial FIS model using the ANN technique. Subtractive clustering and FCM methods were utilized to automatically generate initial FIS models.

The collected data for the KPIs was used to develop the prediction models. The data was divided into two groups: train and test. The training data included the experimental data utilized for tuning the FIS model parameters during the training stage, whereas, the test dataset included new data that had not yet been introduced to the FIS model. A test dataset was utilized to make sure the model is not overfitted. Overfitting means a model is able to predict the training dataset but can not predict future data precisely. Considering statistical consistency between the train and test data can improve the performance of the prediction model. In this research, data division processes considered the statistical parameters between the test and train data [21]. Therefore, the test and train dataset statistical parameters were considered to be as close to each other as possible to represent the same statistical population in order to achieve an optimal model. Mean, standard deviation, minimum, and maximum are statistical parameters that are compared using trial and error to achieve this objective [21]. In this research, the data was divided based on 70% for train and 30% for the test dataset.

Subtractive clustering

Subtractive clustering was utilized to develop the initial FIS model. This is because subtractive clustering has been proposed as a reliable and precise method for developing prediction models. Subtractive clustering can be used to extract cluster centers that represent the FIS model [22]. Each cluster center is used to represent fuzzy rules based on the input/output relationship characteristics as follows: "IF input is near a cluster center THEN output is near the output value of the cluster center" [23].

The number of fuzzy rules in subtractive clustering is impacted by the radius chosen for developing clusters. A bigger radius results in a smaller number of fuzzy rules, while a smaller radius results in a higher number of fuzzy rules but increases the chance of overfitting. Therefore, to achieve optimum precision without overfitting the training dataset, the cluster radius should be optimized [23].

The cluster radius of each FIS model was optimized by changing the cluster radius from 0 to 1, which is the acceptable range in subtractive clustering (Nasrollahzadeh and Basiri, 2014). A FIS model was generated, resulting in various FIS models by changing the cluster radius. For each cluster radius, the developed FIS models' errors were measured in two groups: train and test datasets. Several error measures including Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), and Coefficient of Variation (COV) were considered between the train and test datasets.

A model is optimum when the errors calculated for the test dataset are at their lowest but also as close as possible to the training dataset. This approach ensures the generalization capability of the model and prevents the problem of overfitting (Nasrollahzadeh and Basiri, 2014). When two FIS models perform similarly regarding error measures, the model with the less number of rules (smaller cluster radius) is preferred. This approach of changing radius values was used to select the best initial FIS models. The neuro-fuzzy technique was then utilized to optimize the initial FIS models' parameters to reduce model error as much as possible. The neuro-fuzzy inference system develops a FIS whose membership function parameters were tuned using a backpropagation algorithm in combination with the least squares method. This tuning allowed the FIS model to learn from the data it is modeling [20].

Fuzzy C-means clustering

To develop the initial FIS model, a FCM clustering approach was applied as a comparison with the subtractive clustering approach used in the previous section. FCM is a very common and popular approach for fuzzy clustering. It provides a methodology for grouping data points to populate a multidimensional space into a specific number of different clusters. FCM assigns a membership degree for each cluster and iteratively updates the cluster centers and the membership degrees to minimize objective function. The objective function is the distance from any given data point to a cluster center weighted by that data point's membership degree [24].

In FCM, instead of changing the cluster radius as described in subtractive clustering, the number of rules (clusters) is changed to find the optimum number of clusters. Thus, in this research, the number of rules was changed between 1 to 50 to find the optimum number of cluster centers. This is equal to the number of fuzzy rules in the initial FIS model. The developed FIS models' errors were measured for each number of clusters in two separate groups: train and test datasets. Several error measures including MAPE, RMSE, and COV were calculated between the train and test dataset model results. The optimum model was chosen when the errors calculated for the test dataset were at their lowest but also as close as possible to the training dataset.

After selecting the best initial FIS models, the neuro-fuzzy technique was used to improve the initial FIS models' parameters to reduce model error as much as possible. The neuro-fuzzy inference system tunes the membership function parameters similar to the approach applied for subtractive clustering. After developing the models for predicting KPIs using both subtractive clustering and FCM, the results of the models were compared using validation data to decide which method performed better. For this purpose, different error measures were compared based on the output of the developed models and the actual validation data value.

2.2 Artificial Neural Network technique

ANN was applied to estimate construction project KPIs. The main advantage of ANN is its learning ability. ANN models can find relationships between inputs and outputs using training data. The trained model can then be used to predict the outputs of new inputs.

In this research, the ANN model was used to predict the whole project KPIs. The input of the prediction models was 18 KPIs, six KPIs for each of the three stages. The outputs were six whole project KPIs. ANN models were developed, trained, and tested in MATLAB 2016a. Models were developed using three training algorithms available for neural networks: Levenberg–Marquardt (LM), Bayesian Regularization (BR), and Scaled Conjugate Gradient (SCG). Furthermore, different models were developed using different neuron numbers. The neuron numbers were set from five to 50 neurons with a spacing of five which led to ten models for each training algorithm. Considering the different training algorithm and neurons, 30 different models were developed, and the results compared.

Model performance was assessed based on Coefficient of Determination (R2), Mean Absolute Error (MAE), relative absolute error (RAE), root relative square error (RRSE), and mean absolute percentage error (MAPE) and each error index amount. R2 is a statistical analysis method that ranges between [0, 1] and evaluates the total differences percentage between the target value (ti) and predicted values (oi), as shown in Equation 1. Higher values of R2 indicate a better performing model. MAE is an absolute measure that ranges from 0 to + ∞ and is calculated from Equation 2.

$$R^2 = 1 - \frac{\sum_i (t_i - o_i)^2}{\sum_i \left(t_i - \frac{1}{n}\sum_i t_i\right)^2} \qquad (1)$$

$$MAE = \frac{1}{n}\sum_i |t_i - o_i| \qquad (2)$$

RAE and RRSE were also used to assess forecasting model performance in the same way as shown in Equation 3 and 4 [25]. Lower RAE and RRSE values indicate better forecasting model performance.

$$RAE = \frac{\sum_i |t_i - o_i|}{\sum_i \left|t_i - \frac{1}{n}\sum_i t_i\right|} \qquad (3)$$

$$RRSE = \sqrt{\frac{\sum_i (t_i - o_i)^2}{\sum_i \left(t_i - \frac{1}{n}\sum_i t_i\right)^2}} \qquad (4)$$

MAPE is usually used for evaluating the model accuracy, calculated based on Equation 5. Lower MAPE values indicate a higher model accuracy. Four ranges can be used to divide this index: high accuracy forecast (MAPE < 10%), sound forecast (10% < MAPE < 20%), feasible forecast (20% < MAPE < 50%), and error forecast (MAPE > 50%) [26].

$$MAPE = \frac{100}{n}\sum_i \frac{|t_i - o_i|}{t_i} \qquad (5)$$

The error values mentioned above were calculated for the different BR, LM and SCG algorithms with 5 to 50 neurons. The model with the lowest error values was chosen as the final model for predicting whole project KPIs.

3. DATA COLLECTION

The data in this study was collected from experts. Two sets of questionnaires were designed. The first questionnaire was used to collect data from different projects to run the model and the second questionnaire was used to justify the choice of six KPIs used in this research. Furthermore, another 16 questionnaires were used to validate the models.

To collect data, a questionnaire was designed and distributed in 2016 to owners of building construction projects in Tehran, Iran. For the first questionnaire, two-hundred questionnaires were distributed to experts of which 119 with consistent results were selected for analysis. The questionnaire was designed using an online data collection platform, Qualtrics. In this study, the questionnaires were designed based on KPIs that get from the literature. The first part of the questionnaire get some general information from the respondents. The next part included information about the specific project the questionnaire is designed for. The experts were then asked to rank the six KPIs in three stages of the project and also rank whole project performance. At the end, questions about any possible problems that the project might encounter were included. The data was collected using a 1 to 7 scale, where 1=Very Low, 2=Low, 3=Medium-Low, 4=Medium, 5=Medium-High, 6=High, and 7=Very High [9]. Using quantitative scales allowed machine learning methods to be applied to the qualitative data.

4. MODEL DEVELOPMENT

The model was developed using the neuro-fuzzy technique with both FCM and subtractive clustering. Afterward, the ANN technique was applied for developing the prediction models.

4.1 Neuro-Fuzzy Technique

As described in the methodology section, this study proposes a model that can forecast the project performance using KPIs by a neuro-fuzzy technique in MATLAB using subtractive clustering and FCM clustering. Data were collected by questionnaires in three stages of construction projects, for six KPIs, and a model was developed for each. A large database containing the results of the building projects assembled from an extensive survey of questionnaires was used to develop the neuro-fuzzy models. For each stage, six models were trained. Therefore, 18 FIS models were trained to get the prediction model. In the middle stage, 12 inputs were used to train the model with one output and in the finishing stage, 18 inputs were applied with one output. This model is used to predict the six KPIs in the whole project, which means the average of the whole project KPIs.

In developing the neuro-fuzzy model using subtractive clustering, the codes were written in to generate the models. By writing the codes, the cluster radius amount was optimized, whereas, with the MATLAB toolbox, the radius needs to be provided by the user. Secondly, by writing the codes, data division did not rely on the user to randomly divide data into train and test.

The number of rules was calculated by subtractive clustering in Sugeno-type fuzzy inference system and tuning by the neuro-fuzzy technique. Fig-1 represents an example of the developed models. Here, the six KPIs of the initial stage are the inputs and the output is the cost of the whole project.

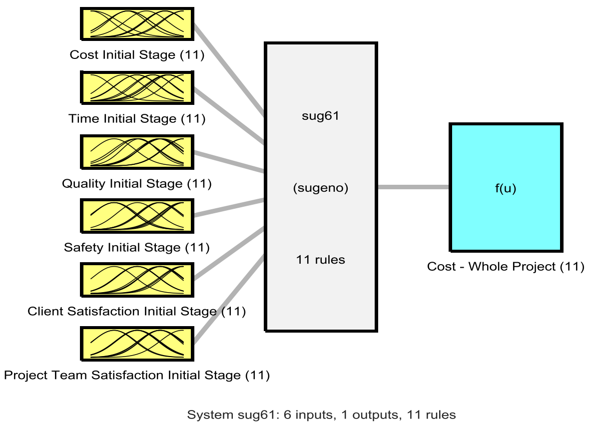

Fig-1: Neuro-fuzzy model using subtractive clustering in the initial stage

In developing the neuro-fuzzy model using FCM, the codes were written in MATLAB software to generate the models. Eighteen neuro-fuzzy models were developed. The number of membership functions is greater in the FCM approach as compared to the subtractive clustering approach, resulting in a more complicated fuzzy rule-based system. In Fig-2 six KPIs of the initial stage are the inputs, and the output is the cost of the whole project.

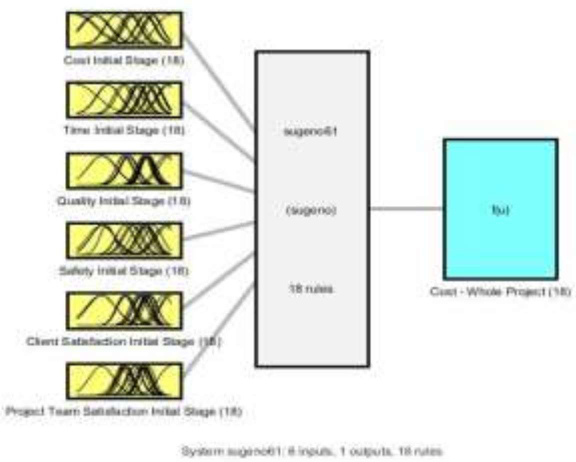

Fig-2: Neuro-fuzzy model using FCM in the initial stage

4.2 Artificial Neural Network

As described in the research methodology, ANN was also applied to predict the whole project KPIs. Different number of neurons and training algorithms were used to find the best model for predicting the KPIs. The dataset was divided into 70 and 30% groups, which were used for training and testing respectively. Three different algorithms were used for the training models, namely: LM, BR, and SCG. For each ANN model, ten different number of neurons (5, 10, 15, ..., 50) with one hidden layer was tested. The performance of the models was evaluated based on different error values: R2, MAE, RAE, RRSE, and MAPE.

Fig-3 to Fig-7 illustrates the MAE, RAE, RRSE, MAPE, and R2 comparison between each of the ten models and three different algorithms; LM, BR, and SCG. As can be seen from the figures, when using the BR algorithm, among all the errors, the model with the 35 neurons has the lowest MAE, RAE, RRSE, and MAPE error value. This model also has the highest R2 value. In addition, the model has a high accuracy forecast model MAPE at less than 10% and the RAE and RRSE values are also very small. Moreover, the model's MAE is 0.059, which is reasonably good. Therefore, the BR model with 35 neurons shows the best performance and was chosen as the final model for predicting project KPIs.

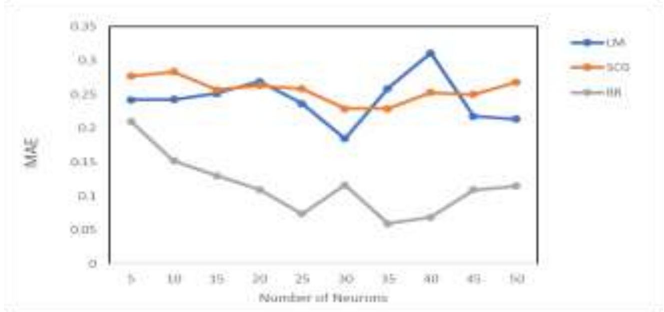

Fig-3: MAE values of ANN models

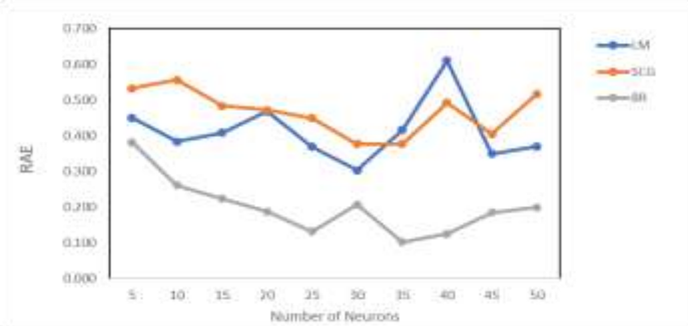

Fig-4: RAE values of ANN models

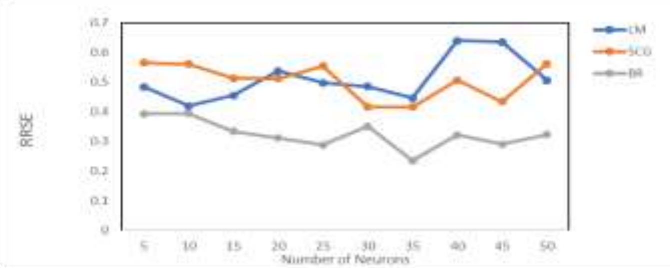

Fig-5: RRSE values of ANN models

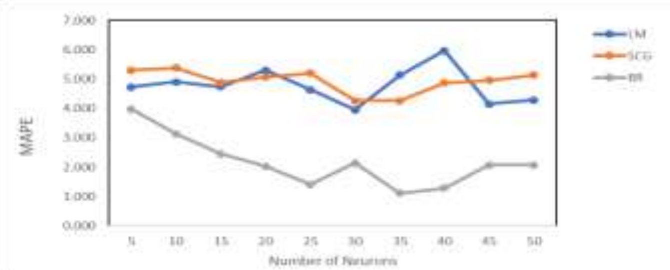

Fig-6: MAPE values of ANN models

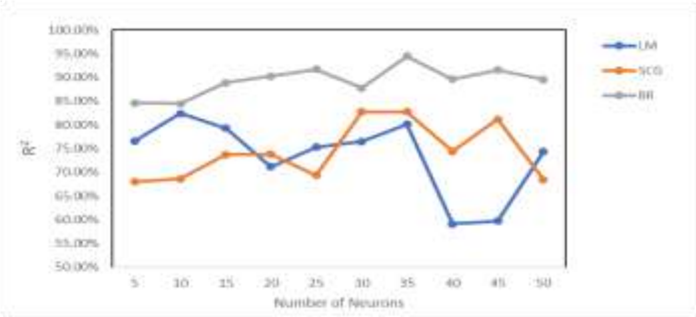

Fig-7: R2 values of ANN models

Therefore, the BR model with the 35 neurons was applied to the collected dataset for predicting the KPIs of building construction projects.

5. VALIDATION

Sixteen questionnaires were used to get data from sixteen real projects to validate the model. First, the developed models' finishing stage KPIs for all employed techniques; ANN and neuro-fuzzy technique with subtractive clustering and FCM were compared. Then, the developed models' finishing stage outputs were compared with the actual data. Table 1 illustrates the error comparison between the three different techniques.

Table -1: Validation data error value comparison for different applied techniques

KPIs	Method	MAPE (%):	RAE	MAE	RRSE
KPI1	ANN	38.55%	0.85	1.34	0.85
	subtractive clustering	37.13%	0.94	1.48	0.91
	FCM	65.72%	1.61	2.53	1.6
KPI2	ANN	28.16%	0.99	1.13	1.09
	subtractive clustering	24.55%	1.15	1.31	1.19
	FCM	29.30%	1.17	1.34	1.25
KPI3	ANN	22.09%	0.99	1.03	0.93
	subtractive clustering	18.09%	0.9	0.94	0.84
	FCM	68.18%	3.41	3.57	3.64
KPI4	ANN	50.95%	1.34	1.42	1.19
	subtractive clustering	42.60%	1.11	1.18	1.1
	FCM	55.06%	1.36	1.44	1.23
KPI5	ANN	24.08%	1.29	1.31	1.46
	subtractive clustering	22.95%	1.12	1.14	1.04
	FCM	25.88%	1.17	1.18	1.11
KPI6	ANN	26.00%	1.03	1.11	0.98
	subtractive clustering	22.00%	0.87	0.94	0.95
	FCM	27.73%	1.24	1.34	1.34

For error measures RAE, MAE, and RRSE, the ANN technique error value is slightly lower for predicting KPI1 and KPI2. However, these error values are lower for predicting KPI 3, KPI 4, KPI 5, and KPI6 when using the subtractive clustering technique. In all cases, the FCM technique produced poor predictive accuracy.

As Table 1 illustrates, the MAPE% error is lower when using the subtractive clustering technique in all six models. The MAPE% error ranges between 18 and 43 percent. According to Jia et al. (2015), MAPE values lower than 50% indicate a sound forecast. As a result, the neuro-fuzzy technique using subtractive clustering is the only approach that provides a sound forecast for predicting all six KPIs [26].

6. DISCUSSION

The neuro-fuzzy technique using subtractive clustering produces better and more acceptable results in predicting KPIs as compared to other techniques. The validation results indicate that the error value for Neuro fuzzy technique with subtractive clustering is the lowest compared with ANN and Neuro fuzzy technique with FCM. MAPE of models developed using neuro-fuzzy technique with subtractive clustering for six KPIs are 1.1% to 8.3% lower compared to ANN technique. Also, MAPE of neuro-fuzzy technique with subtractive clustering technique is 2.9% to 50% lower compared to FCM. The error values also indicate that ANN predictive accuracy is higher compared to FCM technique. These results comply with the results of other research showing that neuro-fuzzy models with subtractive clustering have higher predictive accuracy as compared with other tools such as ANN [17, 27].

7. CONCLUSIONS

Key Performance Indicators (KPIs) evaluate different projects aspects and are used to determine the health status of projects. Therefore, by evaluating and predicting KPIs, monitoring and controlling project progress can be facilitated. A comprehensive literature review of the existing research indicates that limited work has been done on forecasting project performance using KPIs at the project level. Furthermore, there is little focus on performance measurement and forecasting during the project.

Based on the above-mentioned limitations, this research was motivated to develop a framework for measuring and forecasting project KPIs. Because of different stakeholder investments and benefits, the first thing that needed to be defined was from whose viewpoint the performance was measured. Then, a list of KPIs used in literature at the project level was prepared and their frequencies indicated. Six KPIs were chosen due to the frequency of their use in the literature; there are cost, time, quality, safety, client satisfaction, and project team satisfaction.

KPIs of three critical project stages (initial stage, middle stage, and finishing stage) were used to predict whole project KPIs using two main techniques: Artificial Neural Networks (ANN) and neuro-fuzzy. In the ANN, the best model was selected by changing the number of neurons in the hidden layer. Neuro-fuzzy models were developed in two steps; first, the initial FIS models were developed using both subtractive clustering and FCM. In subtractive clustering, cluster radius was optimized to achieve optimum precision without overfitting. Second, the optimization of the initial FIS model was performed using ANN. In the neuro-fuzzy technique, 18 different models were developed, six models for each of the three critical project stages.

All KPIs were measured qualitatively by designing a questionnaire. This research measures project performance from the owner's point of view. Models were developed using the above-proposed framework to forecast project performance using MATLAB software. ANN and neuro-fuzzy techniques using both FCM and subtractive clustering were applied to develop models for predicting whole project KPIs. Results indicate that the neuro-fuzzy technique using subtractive clustering performs better and has lower error values as compared with the other two methods. Thus, for predicting construction project KPIs, the neuro-fuzzy technique using subtractive clustering is recommended. The proposed framework is designed to be flexible and can be applied to other countries and other types of projects.

REFERENCES

[1] N. Nassar and S. AbouRizk, "Practical Application for Integrated Performance Measurement of Construction Projects," Journal of Management in Engineering, p. 04014027, 2014.

[2] J. Egan, "The Egan report-rethinking construction," report of the construction industry task force to the deputy prime minister. London, 1998.

[3] H. R. Kerzner, Project management metrics, KPIs, and dashboards: a guide to measuring and monitoring project performance: John Wiley & Sons, 2011.

[4] DCC, "Corporate Plan Summary 2016–2017 to 2020–2021," Defense Construction Canada, Canada2016.

[5] M. Freeman and P. Beale, "Measuring project success," 1992.

[6] D. K. H. Chua, Y.-C. Kog, and P. K. Loh, "Critical success factors for different project objectives," Journal of construction engineering and management, vol. 125, pp. 142-150, 1999.

[7] R. L. Tucker, D. R. Shields, and S. R. Thomas, "Measurement of Construction Phase Success of Projects," in Construction Research Congress, ed, 2003, pp. 1-8.

[8] J. Li, O. Moselhi, and S. Alkass, "Forecasting project status by using fuzzy logic," Journal of construction engineering and management, vol. 132, pp. 1193-1202, 2006.

[9] M. Dissanayake and A. R. Fayek, "Soft computing approach to construction performance prediction and diagnosis," Canadian Journal of Civil Engineering, vol. 35, pp. 764-776, 2008.

[10] H. S. Cha and C. K. Kim, "Quantitative approach for project performance measurement on building construction in South Korea," KSCE Journal of Civil Engineering, vol. 15, pp. 1319-1328, 2011.

[11] T. Haponava and S. Al-Jibouri, "Proposed System for Measuring Project Performance Using Process-Based Key Performance Indicators," Journal of Management in Engineering, vol. 28, pp. 140-149, 2012.

[12] C. Ngacho and D. Das, "A performance evaluation framework of construction projects: insights from literature," International Journal of Project Organisation and Management, vol. 7, pp. 151-173, 2015.

[13] E. L. Oral, M. Oral, and M. Andaç, "Construction crew productivity prediction: application of two novel methods," International Journal of Civil Engineering, vol. 14, pp. 181-186, 2016.

[14] M. Reenu, R. Kumar, and S. Babu, "Construction Project Performance Model Using Artificial Neural Network," International Journal of Recent Trends in Engineering & Research (IJRTER), 2017.

[15] H. Leon, H. Osman, M. Georgy, and M. Elsaid, "System Dynamics Approach for Forecasting Performance of Construction Projects," Journal of Management in Engineering, vol. 34, p. 04017049, 2017.

[16] S. S. Shaikh and M. M. Darade, "Key Performance Indicator for Measuring and Improving Quality of Construction Projects," International Research Journal of Engineering and Technology (IRJET), 2017.

[17] M. Nilashi, M. Dalvi-Esfahani, O. Ibrahim, K. Bagherifard, A. Mardani, and N. Zakuan, "A soft computing method for the prediction of energy performance of residential buildings," Measurement, vol. 109, pp. 268-280, 2017.

[18] N. K. Nassar, "An integrated framework for evaluation of performance of construction projects " PMI Global Congress 2009--North America. Proceedings, RSC04.PDF., 2009 2009.

[19] S. S. Fanaei, O. Moselhi, and S. Alkass, "Dynamic performance forecasting model and measurement system in construction project " presented at the CSCE Annual Conference, 2016.

[20] J. R. Jang, MATLAB: Fuzzy logic toolbox user's guide: Version 1: Math Works, 1997.

[21] M. A. Shahin, H. R. Maier, and M. B. Jaksa, "Data division for developing neural networks applied to geotechnical engineering," Journal of Computing in Civil Engineering, vol. 18, pp. 105-114, 2004.

[22] S. L. Chiu, "Fuzzy model identification based on cluster estimation," Journal of intelligent and Fuzzy systems, vol. 2, pp. 267-278, 1994.

[23] [23] K. Nasrollahzadeh and M. M. Basiri, "Prediction of shear strength of FRP reinforced concrete beams using fuzzy inference system," Expert Systems with Applications, vol. 41, pp. 1006-1020, 2014.

[24] J. C. Bezdek, "Objective Function Clustering," in Pattern recognition with fuzzy objective function algorithms, ed: Springer, 1981, pp. 43-93.

[25] S. Makridakis and M. Hibon, Evaluating accuracy(or error) measures: INSEAD, 1995.

[26] C. Jia, L. Wei, H. Wang, and J. Yang, "A hybrid model based on wavelet decomposition-reconstruction in track irregularity state forecasting," Mathematical Problems in Engineering, 2015.

[27] S. K. Paul, A. Azeem, and A. K. Ghosh, "Application of adaptive neuro-fuzzy inference system and artificial neural network in inventory level forecasting," International Journal of Business Information Systems, vol. 18, pp. 268-284, 2015.