# High Prediction Accuracy and Low Error for ERP User Satisfaction by Hybrid of ANFIS and KNN Classification



Pinky Kumawat
Politecnico di Torino

# High Prediction Accuracy and Low Error for ERP User Satisfaction by Hybrid of ANFIS and KNN Classification

Pinky Kumawat
Department of Computer Engineering
Rajasthan College of Engineering for Women,
Jaipur, RTU, Kota.
kumawat.pinky3@gmail.com

**Abstract—The incoming era is becoming more friendly and dependent on Information Technology. Enterprise Resource Planning (ERP) Systems are one of the most widely used latest examples of Information Systems (IS) technology. They provide a single window system to the organizations by integrating the whole functions of them. Today, all enterprises are rapidly adopted ERP systems. But, their adoption and implementation is not being without any problem. The implementation process of ERP is also a very challenging, time consuming and costly task. Therefore, instead of many efforts if the implementation process is failed. Then it will be a big failure for the organization. Hence, to overcome this failure and increase the success rate of ERP projects we need to develop a robust, reliable and accurate predictor. This will help us to redirect the projects far better in advance. The success of ERP systems depends on many factors. US is one of the important factor among them. Hence, we develop an efficient predictor of US using hybrid of ANFIS and KNN. We were used this method first time in literature related to prediction of US in ERP. The Hybrid method increases the prediction accuracy more comparatively than previous reported techniques ANN, ANFIS and KNN. The RMSE using Hybrid method is 0.167629 and for KNN, ANFIS and ANN is 0.5, 0.486185, and 0.590329 respectively.**

**Index Terms— ANFIS, ANN, Critical Success Factors, ERP, KNN classifier, User Satisfaction**

## I. Introduction

Enterprise Resource Planning (ERP) is the core business process and management software for enterprises. It provides efficiency, centralized storage/backup, workflow, visibility, collaboration and modular design etc. benefits to the organizations [1]. Recently, almost all organizations are using ERP system software to improve business communication and efficiency. Hence, the success of ERP implementation is very important for organizations to improve their business plans. But it seems not easy and many types of failure have to face the organizations during implementation process. The latest CHAOS-2015 study report [2] of Standish group shows the track record of successful IT projects of which ERP is a subset projects remains poor. This report shows the marked decrease in success rate of IT projects and 29% projects were in the rate of successful projects, delivered on time and on budget with required features. And, 52% projects were in the category of challenged projects and delivered late and over budget with less than the required features. 19% projects were cancelled before completion and never used further. Hence this report shows that instead of rapidly adoption of ERP projects the failure rate is more reporting in day by day implementation of ERP. With this failure the organizations have to face many failures for example manpower, economical power, resources and time etc. So, to reduce this failure we need a proactive approach to redirects the projects in better direction. This proactive control can only be possible by some predictive capabilities. Therefore, this paper introduces a robust, reliable and easy predictor which would be a red-flag for impending failures in ERP projects. The success of ERP system depends on many factors. US is an important factor among them. According to peslak et al. [3] people are one of the important variables for winning ERP strategy. So, there satisfaction for used ERP is also very important criteria. Molonado et al. [4] also reported that US significantly influence ERP business improvement success. Because if an information system adds some value to the organization than it will also have some impact on the behavior of user. Hence we consider in this paper US for prediction which plays most important role in the success of ERP. US is also a multidimensional variable. It depends on some other factors which influence US and ERP system success. So, we consider in this paper the three most important factors for ERP system success that are Human, Organizational and Technological which is reported in literature by Bernal et al. [5]. Considering these three factors we proposed first time a method for prediction of US that is hybrid of ANFIS and KNN. And, also compare this method with existing methods ANN, ANFIS and KNN based on RMSE and in result hybrid method gives high predication accuracy and low error. So our approach is novel for US prediction and it can be useful to make ERP successful.

## II. Literature review

This section presents a review of the related research work, models, frameworks, prediction methods and analysis approach from literature. US is a term occurred several times in literature relevant to the success of ERP projects. Some researchers also focus on factors which influence the ERP success and affect US. Gupta et al. [6] reported a comparative study of ANFIS membership functions for US prediction and proposed trimf as a best function. This predicts the value closer the observed value of US and reported error of 0.5. Venugopal et al. [7] gave an Adaptive Neuro Fuzzy Inference System (ANFIS) approach for prediction of US. They reported RMSE value for ANFIS method is 0.277 which is comparatively less than ANN and MLRA. Dezdar et al. [8] identified the key factors that create ERP user satisfaction. They also discovered that ERP users how varied among different user's profiles. Frejik et al. [9] reported a literature review about the US with ERP implementation. They proposed a conceptual model based on literature for US prediction. They identified four factors user, innovation, organization and environment as predictors of US in the success of ERP implementation. Lotfy et al. [10] proposed a conceptual model to measure ERP user value. They considered technological, environmental and organizational as a predictor factors for US. Authors also measured these factors' impact on the overall ERP benefits for the organization. Gupta et al. [11] gave a comparative study of institute based ERP based on ANFIS, ANN and MLRA. They proved that ANFIS is efficient method for US prediction based on RMSE. The reported RMSE by them is 0.2945 which is comparatively less than previous reported prediction methods. Bhawarkar et al. [12] proposed a framework for the implementation of ERP to improve the performance of business. They considered approach related, culture, communication, and support related, project management related and vision, scope, goals and infrastructure related factors as input factors that affect ERP success. System quality, information quality, individual impact, organizational impact, and workgroup impact as output variables for performance measures of ERP system. Roses et al. [13] gave an end user computing satisfaction model in the context of an ERP system in a transnational bank. They consider five factors for predict US. Those factors were format, accuracy, content, easy to use and less time consuming. They used CFA for prove of further reliability and validity of the model with the observed data. Rouhani et al. [14] reported an artificial neural network approach for prediction of ERP success. They considered organizational profiles as input variables and ERP performance variables as output variables and by considering them they checked the correct prediction rate of ANN for ERP success and they reported 0.685 classification rate. Jenatabadi et al. [15] developed a logit regression model for ERP US prediction. They considered eight factors as input variables. Those factors were age, gender, marital, educational level, experience, computer, constant and income. The output variable was US which is binary outcome. They validate their model by conducting case study. Tsaur et al. [16] investigated the success of ERP systems by conducting the case studies in Taiwanese high tech industries. They considered service quality, information quality, system quality, behaviour intention, user satisfaction, benefit of use from end user's view and net value from business' view. After data collection they used principle component analysis and proved that system quality, service quality and information quality are most important successful factors. Kumawat et al. [17, 18, 23] reported some review about the US in ERP implementation success, use of Support Vector Machine (SVM) Classifier for US prediction and an analytical approach for US prediction.

## III. Proposed Methodology

In this paper the data collected based on the questionnaires. The questionnaire is designed by considering ex ante critical success factors. Those CSF's reported by Bernal et al. [5] in literature and they logically mapped them into three factors human, technological and organizational. These three factors are considered as input variables or predictor of US, which is output variable.

### *A. Sample Dataset*

The data is collected based on survey questionnaire using online Google survey form. Online survey is conducted in two organizations separately. And, separate further analysis approach and prediction methods applied on two collected set of responses. 40 responses collected from one organization and 28 responses collected from other organization. Out of 40, 7 were incomplete and with missing data so, 33 were used for further analysis and prediction. Out of 28, 4 were incomplete and with missing data so, 24 were used for further analysis and prediction. Hence the factor score is calculated based on the variance explained by each critical success factor with US response. A weight is assigned according to variance and using weight and corresponding responses, factor scores are calculated. The both sample datasets with calculated factor scores are shown in below tables.

TABLE I
US AND COMPUTED FACTOR SCORES: DATASET 1

| **S.No.** | **Factor Scores** | | | |
|---|---|---|---|---|
| | **Human** | **Tech.** | **Org.** | **US** |
| 1 | 2.08 | 2.21 | 1.86 | 1 |
| 2 | 6.37 | 5.79 | 6 | 6 |
| 3 | 5.35 | 5.58 | 5.26 | 5 |
| 4 | 4.18 | 4.07 | 4.6 | 4 |
| 5 | 5.08 | 3.86 | 4.07 | 5 |
| 6 | 6.12 | 5.87 | 6.2 | 6 |
| 7 | 4.43 | 2.9 | 4.31 | 3 |
| 8 | 6.33 | 5.53 | 6.32 | 6 |

TABLE II
US AND COMPUTED FACTOR SCORES: DATASET 2

| **S.No.** | **Factor Scores** | | | |
|---|---|---|---|---|
| | **Human** | **Tech.** | **Org.** | **US** |
| 1 | 5.08 | 5.2 | 5.49 | 5 |
| 2 | 6.33 | 6.15 | 5.63 | 6 |
| 3 | 6.08 | 4.98 | 6.23 | 5 |
| 4 | 5.83 | 4.74 | 4.55 | 5 |
| 5 | 2.33 | 1.79 | 2.03 | 1 |
| 6 | 5.58 | 4.34 | 5.3 | 5 |
| 7 | 6.33 | 6.15 | 5.99 | 6 |
| 8 | 3.74 | 4.42 | 5.43 | 4 |

### *B. Artificial Neural Networks (ANN)*

Artificial Neural Networks are multi layer and feed forward neural networks [19]. They work on the principle of adaptive learning capability. Training data is used for learning process and based on back propagation procedure. We used in this paper trainlm algorithm for training the network which is proposed by Dasgupta et al. [20]. During training using the back propagation method neural networks adjust their connection weights. By adjusting connection weights they tried to calculate prediction nearest to actual output. ANN contains five layer networks: layer1-input layer, layer2-hidden layer, layer3-output layer. Connections between these layers are known as synapses. The ANN architecture is shown in below figure.

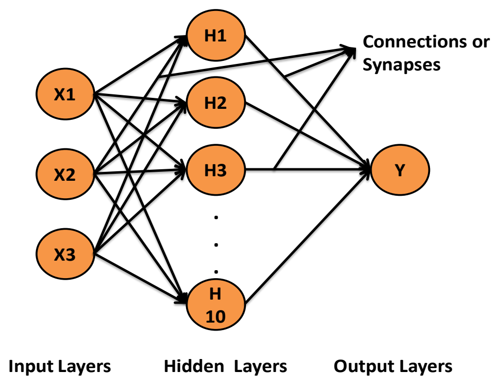


Fig. 1. ANN Architecture

### *C. Adaptive Neuro Fuzzy Inference System (ANFIS)*

ANFIS is a combination of adaptive learning capability and fuzzy logic [21]. It is a mix of ANN and Fuzzy Inference System (FIS). ANN learns by adjusting the weights between different connections in the network. FIS based on the principle of fuzzy reasoning, fuzzy rules and fuzzy sets theory. It gives good prediction results comparatively than ANN and gives less prediction errors in all cases. ANFIS is a five layer architecture given in below figure.

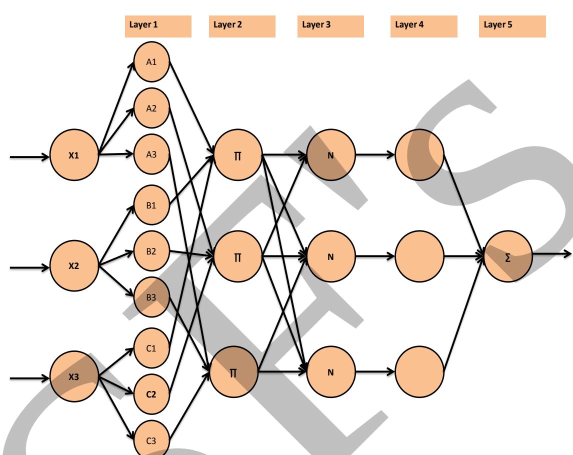


Fig. 2. ANFIS Architecture

### *D. K-Nearest Neighbor Classification (KNN)*

KNN is a simple and efficient algorithm and classifier. It stores all the available cases and classifies new dataset on the basis of a similarity measure (i.e. distance function). A new case is classified by a majority votes of its neighbors (here k represent the nearest number of neighbors). The predicted class of a corresponding new case is the most common class among its nearest neighbors using distance function. The default distance function used in KNN is Euclidean distance. We used in this paper the Euclidean distance for distance calculation [22].

### *E. Hybrid of ANFIS and KNN (Proposed Prediction Method)*

We proposed a hybrid of ANFIS and KNN. This combines the advantages of both methods ANFIS and KNN. We compared our prediction methods on the basis of Root Mean Squared Error (RMSE) reported by Venugopal et al. [7]. On the basis of RMSE hybrid method gives best prediction results among all prediction methods. ANFIS gives error in all the cases which is a disadvantage of it, but it gives less error in all the cases which is an advantage of it. KNN gives correct predictions in most of the cases which is an advantage of it, but it gives error in few cases which is a big error. Hence, we proposed a hybrid method which overcomes the disadvantage of both methods and combines the advantage of both methods. The prediction is calculated on the basis of weighted average. We assigned more weight (out of 100 percent) to that method which gives less error and assign less weight (remaining out of 100 percent) to that method which gives more error. On the basis of RMSE hybrid prediction method gives best prediction results among all methods. It increases the prediction accuracy by decreasing the error level more comparatively.

### *F. Modeling of Hybrid of ANFIS and KNN*

The mathematical model of hybrid of ANFIS and KNN classification approach is the weighted average approach of both the mathematical models. The steps to calculate the prediction results from hybrid of ANFIS and KNN are given as following:

1. Assign weight to ANFIS approach between 0-100 percent.

2. The weight assigned to KNN approach is calculated as:
   KNN_weight = 100-ANFIS_weight

3. ANFIS network loaded using readfis as:
   fismat = readfis('ANFIS_TRAINED.fis')
   where, ANFIS_TRAINED.fis file is made using anfisedit command and load training data and generate fuzzy inference system for trained data and also loaded test data in it.

4. Load input and output train data for KNN Classification using following commands
   load('Inputs_KNN.mat');
   load('Outputs_KNN.mat');

5. KNN Classification model is prepared using function as:
   mdl = ClassificationKNN.fit(Inputs,Outputs)
   where, mdl is a KNN classification model which contains types of all classes, distance function, number of nearest neighbors, number of cases, number of input and output parameters.

6. Enter input data values $X_i$ as Human, Technological and Organizational for prediction of their corresponding user satisfaction and taken collectively in xnew as.
   xnew = [Human, Technological, Organizational]
   where, i=1,2,3 number of input parameters

7. Prediction result of ANFIS is given using the function as
   results = evalfis(xnew',fismat)
   where, evalfis perform fuzzy inference calculation on xnew' using fismat and this calculation is done in five layers functions as given.

8. ANFIS calculate layer1 output as

$O_{1,i} = \mu_{Ai}(X1)$ where i=1,2,3

$O_{1,i} = \mu_{Bi}(X2)$ where i=1,2,3

$O_{1,i} = \mu_{Ci}(X3)$ where i=1,2,3

where, $\mu(X_i) = 1/1+e^{-aXi}$ is an activation function where a is the corresponding weight of the connection between input nodes/neurons and layer1.

9. The output of layer1 is given as input to layer2 and output of layer2 is calculated as

$O_{2,i} = W_i = \mu_{Ai}(X1)*\mu_{Bi}(X2)*\mu_{Ci}(X3)$

where, i=1,2,3

10. Layer2 output is given as input to layer3 and output of layer3 is calculated as

$O_{3,i} = \hat{W}_i = W_i / \sum W_i$ where, i=1,2,3

11. Layer3 output is given as input to layer4 and output of layer4 is calculated as

$O_{4,i} = \hat{W}_i f_i = \hat{W}_i(p_i X1+q_i X2+r_i X3)$

where, i=1,2,3

and $p_i$, $q_i$ and $r_i$ are the assigned weights from input to layer1 nodes between 0 to 1.

12. Layer4 output is given as input to layer5 and output of layer5 is calculated as the predicted US of ANFIS is.

Predicted US = results_anfis = $O_{5,i} = \sum \hat{W}_i f_i$

where, i=1,2,3 and

$f_i = (p_i X1+q_i X2+r_i X3)$

i=1,2,3 and $p_i$, $q_i$ and $r_i$ are weights from input to layer1 connections between 0 to 1.

13. Prediction of US using KNN Classification is given as

Predicted US = results_knn = predict(mdl,xnew);

and the calculation in KNN prediction is given in below steps.

14. Calculate the Euclidean distance of input values $Q=X_j$=xnew with all other cases $P_i$ in samples in training dataset as given.

Euclidean Distance

(1)

where, $P_i=X_j$ and j=1,2,3 number of factors or inputs

$Q=X_j$ where j=1,2,3 and i=1,2,3,……,n number of responses in training sample.

15. Now rank all the sample dataset according to minimum distance (minimum distance, highest rank).

16. Make prediction for the xnew input parameter Q according to the selection of k nearest number of neighbors.

17. Prediction outcome or results_knn is the majority vote between the selected neighbors.

18. Combined weighted result of hybrid of both ANFIS and KNN Classification is given using equation as:
result_Combined=((results_anfis*weight_anfis)+(results_knn *weight_knn))/100;

## IV. Results and Discussion

### *A. Predicted US*

Predicted US using all prediction methods ANN, ANFIS, KNN and Hybrid is given in Table III & IV

TABLE III
PREDICTED US: DATASET 1

| S.No. | Predicted US | | | |
|---|---|---|---|---|
| | **ANN** | **ANFIS** | **KNN** | **Hybrid** |
| 1 | 0.8654 | 0.9999 | 1 | 1 |
| 2 | 5.4733 | 5.891 | 6 | 5.9782 |
| 3 | 4.9956 | 5.0105 | 5 | 5.0011 |
| 4 | 4.4374 | 4.0126 | 4 | 4.0013 |
| 5 | 5.2766 | 4.0303 | 4 | 4.7443 |
| 6 | 6.8303 | 5.0494 | 6 | 5.9899 |
| 7 | 3.9938 | 2.9977 | 4 | 3.5981 |
| 8 | 5.8224 | 6.1868 | 6 | 6.0187 |

TABLE IV
PREDICTED US: DATASET 2

| S.No. | Predicted US | | | |
|---|---|---|---|---|
| | **ANN** | **ANFIS** | **KNN** | **Hybrid** |
| 1 | 5.6766 | 4.9012 | 5 | 4.98024 |
| 2 | 5.301 | 6.0868 | 6 | 6.01786 |
| 3 | 4.2073 | 5.7098 | 5 | 5.14196 |
| 4 | 5.2353 | 4.8213 | 4 | 4.65704 |
| 5 | -1.1052 | 0.6213 | 1 | 0.92426 |
| 6 | 4.8945 | 4.3565 | 5 | 4.8713 |
| 7 | 6.1963 | 6.0388 | 5 | 5.83104 |
| 8 | 3.968 | 3.4321 | 4 | 3.88642 |

### *B. Comparative Results*

The We were calculated error for ANN, ANFIS, KNN and hybrid prediction methods using following equation[6].

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n_i}(A_i - P_i)} \qquad (2)$$

wherein, $A_i$ is the actual US and $P_i$ is the predicted US, n is the number of test cases.

Table V & VI shows error calculated results for all methods for two data set using equation (2). This calculation suggest that hybrid method can be superior from other existing methods for US in ERP.

TABLE V
COMPARISION OF ALL METHODS IN TERMS OF RMSE: DATASET 1

| Method | ANN | ANFIS | KNN | Hybrid |
|---|---|---|---|---|
| **RMSE** | 0.590329 | 0.486185 | 0.5 | 0.167629 |

TABLE VI
COMPARISION OF ALL METHODS IN TERMS OF RMSE: DATASET 2

| Method | ANN | ANFIS | KNN | Hybrid |
|---|---|---|---|---|
| **RMSE** | 0.87412 | 0.423463 | 0.5 | 0.158994 |

Figure 3 & 4 shows RMSE calculation of data set1 and data set2, respectively for ANN, ANFIS, KNN and hybrid.

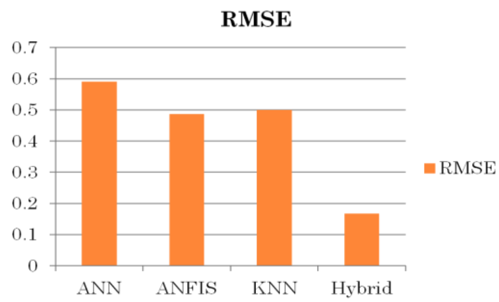


Fig. 3. Comparison of all prediction methods based on RMSE: Dataset1

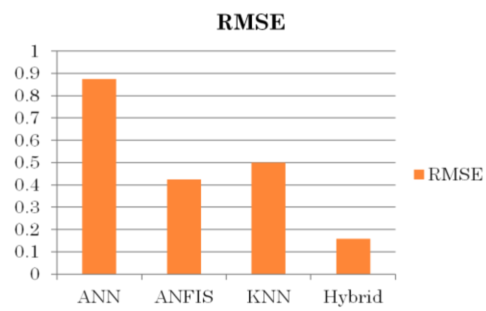


Fig. 4. Comparison of all prediction methods based on RMSE: Dataset2

## V. Conclusion

This paper proposed a novel approach for predicting ERP US using hybrid of ANFIS and KNN. Using this method for prediction of ERP US, the prediction accuracy can increase and the error level decrease more comparatively than existing methods. We also tried to consider important CSF's for ERP US prediction which influence ERP success and US.. This proposed prediction method will be very helpful for organizations in predicting ERP US because it gives more accurate and efficient prediction with less error. By using this method ERP implementing organizations can reduce the possibility of failure in advance and redirects the projects in better direction. We tried to propose a programming framework which combines the existing methods with proposed method in a single program for prediction. This method of prediction can be used as a decision making tool to support the management of organizations when taking decisions regarding the implementation of ERP.

## References


[1] Hossain, L. Patrick, J. David and Rashid, Mohammad. A., "Enterprise Resource Planning: Global Opportunities & Challenges", *ISBN: 193070836x*, Idea Group Publishing, 2002.

[2] https://www.infoq.com/articles/standish-chaos-2015.

[3] A. R. Peslak and T. A. Boyle, "An exploratory study of the key skills for entry-level ERP employees", *International Journal of Enterprise Information Systems (IJEIS)*, Vol. 6, No. 2, pp. 1-14, 2010.

[4] Maldonado, M., and Sierra, "User satisfaction as the foundation of the success following an ERP adoption: An empirical study from Latin America", *International Journal of Enterprise Information Systems (IJEIS)*, Vol. 9, No. 3, pp. 77-99, 2013.

[5] N. Garcia-Sanchez & L. E. Perez-Bernal, "Determination of critical success factors in implementing an ERP systems: A field study in Mexican enterprises", *Information Technology for Development*, Vol. 13, No. 3, pp. 293-309, 2007.

[6] N. Gupta, G. Sharma, and R. S. Sharma, "A Comparative Study of ANFIS Membership function to Predict ERP User Satisfaction using ANN and MLRA", *International Journal of Computer Applications*, Vol. 105-No. 5, pp. 0975-8887, November 2014.

[7] C. Venugopal, S. P. Devi and K. S. Rao, "Predicting ERP User Satisfaction-an Adaptive Neuro Fuzzy Inference System (ANFIS) Approach", *Intelligent Information Management*, Vol. 2, pp. 420-430, 2010.

[8] S. Dezdar, "User Satisfaction Issues in ERP Project", World Academy of Science, Engineering and Technology, *International Journal of Social, Behavioral, Economic, Business and Industrial Engineering*, Vol. 6, No. 8, pp. 2277-2280, 2012.

[9] M. D. Frejik, and A. Powell, "User Satisfaction with ERP Implementations: A Literature Review", *Association of Information Systems, AIS Electronic Library (AISeL)*, MWAIS 2015 Proceedings, Paper 25, pp. 1-4, 2015.

[10] M. A. Lotfy, and L. Halawi, "A Conceptual Model to Measure ERP User-Value", *Issues in Information Systems*, Vol. 16, Issue III, pp. 45-63, 2015.

[11] R. Gupta, S. Chouhan, and N. Bhuria **"**Comparative study of Institute based ERP based on ANFIS, ANN and MLRA", *International Journal of Engineering and Technical Research (IJETR)*, Vol. 2, Issue-5, ISSN: 2321-0869, May 2014.

[12] R. M. Bhawarkar, and L. P. Dhamande, "A Framework for the Implemenation of Enterprise Resource Planning (ERP) to Improve the Performance of Business", *International Journal of Research in Advent Technology*, Vol. 1, Issue 5, E-ISSN: 2321-9637, December 2013.

[13] L. K. Roses, "Antecedents of End-User Satisfaction with an ERP System in a Transactional Bank", *Journal of Information Systems and Technology Management (IJSTEM)*, Vol.-8, no.-2, pp. 389-406, 2011.

[14] S. Rouhani, and A. Z. Ravasan, "ERP success prediction: An artificial neural network approach", *Scientia Iranica, Transactions E: Industrial Engineering*, Vol.-20, no.-3, pp. 992–1001, October 2012.

[15] H. S. Jenatabadi, and A. Noudoostbeni, "End-User Satisfaction in ERP System: Application of Logit Modeling", *Applied Mathematics Sciences*, Vol.-8, no.-24, pp. 1187-1192, February 2014.

[16] Shih-Wen Chien, and S. M. Tsaur, "Investigating the Success of ERP Systems: Case Studies in three Taiwanese High-Tech Industries", *Computers in industry*, Vol.-58, no. 8-9, pp. 783-793, March 2007.

[17] P. Kumawat, and G. Kalani, "A Review on Prediction of ERP Outcome Measurement and User Satisfaction by Use of AI (Fuzzy Logic and Neural Networks) ", *International Journal of Engineering Research and General Science*, Vol. 3, Issue 4, ISSN 2091-2730, pp. 59-64, July-August, 2015.

[18] P. Kumawat, G. Kalani, N. K. Kumawat "Prediction of ERP Outcome Measurement and User Satisfaction by using Adaptive Neuro Fuzzy Inference System and SVM Classifiers approach" *Proceedings of the International Congress on Information and Communication Technology (ICICT) in Springer*, Vol. 1, pp. 229-237, (**DOI** 10.1007/978-981-10-0767-5_25), 2015.

[19] Haykin, S., "Neural Networks: A Comprehensive Foundation", *ISBN: 0132733501*, 2nd Edition, Prentice-Hall, Upper Saddle River, NJ, 1999.

[20] C. G. Dasgupta, G. S. Dispensa, and S. Ghose, ''Comparing the Predictive Performance of a Neural Network Model with Some Traditional Market Response Models", *International Journal of Forecasting*, Vol.-10, no.-2, pp. 235–244, September 1994.

[21] J. S. R. Jang, "Adaptive-Network-Based Fuzzy Inference system", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol.-23, no.-3, pp. 665-685, June 1993.

[22] G. Guo, H. Wang, D. Bell, Y. Bi, and K. Greer, "KNN Model-Based Approach in Classification", *Lecture Notes in Computer Science*, Vol.-2888, pp. 986 – 996, November 2003.

[23] P. Kumawat and N. K. Kumawat, "Investigating End User Satisfaction in ERP Systems: An Analytical Approach", *International Journal on Emerging Trends in Technology (IJETT)*, Vol. 4, ISSN: 2455 – 0124, Issue 3, pp. 9020-9024, Sept 2017.