User Satisfaction Prediction in ERP using KNN Classifier for high Prediction Accuracy

INTERNATIONAL JOURNAL ON EMERGING TRENDS IN TECHNOLOGY

User Satisfaction Prediction in ERP using KNN Classifier for high Prediction Accuracy

Pinky Kumawat[1],

kumawat.pinky3@gmail.com[1]
Computer Science and Engineering
Rajasthan College of Engineering for Women,
Jaipur, RTU, Kota

Abstract— ERP (Enterprise Resource Planning) systems are widely used in organizations; because, ERP provides a single platform to manage all the processes and functions of organizations. This single platform improves their productivity, business performance, decision making capabilities and efficiency. However, to achieve a proper level of ERP success depends on various factors e.g. organization, technology, environment and User Satisfaction etc. 'User Satisfaction' (US) is most important factor to make ERP successful. US refer the user's comfort and acceptability of ERP system during the use and interaction with the ERP system. This paper deploys the conceptual model for US prediction by considering Human, Technological and Organizational factors as predictors. In this report, we proposed K-Nearest Neighbor (KNN) Classification method first time in literature to predict the US and we compare it with ANFIS and ANN. We calculated average error for all test cases and demonstrate that KNN gives high predication accuracy in most of the cases and low average error (0.25) in comparison ANFIS (0.3378) and ANN (0.6053) methods. So our approach is novel and using KNN, prediction accuracy can be further improved for US to make successful ERP.

Keywords— *ERP, User Satisfaction, KNN classifier, Critical Success Factors, ANFIS, ANN.*

I. INTRODUCTION

Enterprise Resource Planning (ERP) systems are defined as a single software system which integrates or manages all departments and functions across an organization into a single computer system that can serve all those different department's particular needs [1]. These systems have been popularly adopted by organizations in developed and developing countries. ERP systems make enable to the organizations to improve their operational efficiency and business efficacy [2]. ERP provides many benefits to organizations e.g. integrated information, efficiency, customer service, visibility, workflow, advanced e-commerce, scalability, modular software capability etc. [3]. But, if the implementation of ERP systems is not successful then it will be a big failure, because the ERP implementation process is very time consuming and challenging task for the organizations. The implementation process is also with very high cost and the additional cost may also require during implementation. One key measure to gauge the success of ERP implementation is User Satisfaction and it also a measurement of effectiveness of information systems [4]. US defined as the degree to which the objective of any organization and system to utilizing the system is achieved. Hence the main success of ERP systems depends on the US [5]. US is also depends on many Critical Success Factors (CSF's) which contribute in the success of ERP implementation and affects US. The fourteen CSF's were considered from prior research for this study [1]. These fourteen CSF's were logically mapped into three factors that are human, technological and organizational and used as predictors for US. US is a term occurred many times in literature for successful ERP implementation. Lotfy *et al*. [6] identified the Technological, Organizational and Environmental factors that maximize the ERP user value in the upward and onward phase of ERP post implementation process. Frejik *et al*. [7] developed a framework to identify the concern areas in existing research on US and ERP implementation that are user, innovation, organization and environment. Gupta *et al*. [8] given a comparative study of ANFIS membership functions to predict ERP US. Rajeev Gupta *et al*. [9] reported a comparative study of institute based ERP using ANFIS, ANN and MLRA to predict ERP US. Jenatabadi *et al*. [10] proposed a logit model for End User Computing Satisfaction (EUCS) and validating it by case study considering eight items gender, age, educational level, marital, experience, income, computer and constant. Bhawarkar *et al*. [11] identified the critical success factors that impacts ERP systems success and performance of organizations. Authors also statistically validate their model for checking the existing correlation between them. Dezdar *et al*. [12] investigated factors that creates US and discovered that how ERP US varies among profiles of user's considering age, gender, education and IT experience. Rouhani *et al*. [13] gave an ERP success prediction model using ANN approach considering organizational profiles as input factors and ERP implementation success factors as output factors. Roses *et al*. [14] identified the antecedents of end user satisfaction with an ERP system in the context of a transactional bank. They developed an end user satisfaction model using Confirmatory Factor Analysis (CFA). Venugopal *et al*. [5] used the ANFIS approach for prediction of US based on RMSE and compared it with ANN and MLRA. Kumawat *et al*. [15, 16, 17] reported a review about the US in ERP implementation success, use of Support Vector Machine (SVM) Classifier for US prediction and an analytical approach. We collected data using online Google form from institute with live ERP implementation. And, we used KNN classification for US predication first time in the literature for ERP system and this method gives high prediction accuracy and low average error in comparison to ANFIS and ANN.

II. Proposed Methodology

In this paper, we used K-Nearest Neighbor (KNN) classification approach to predict US in ERP system success. KNN is compared with existing methods ANFIS and ANN for US prediction in ERP system success. US is a multidimensional variable which is affected by many factors in ERP implementation success. This paper considers the Human, Technological and Organizational factors as the predictor variables from earlier research [18] based on 14 CSF's and US as the dependent variable. The conceptual model for US prediction is given in figure 1 based on these factors.

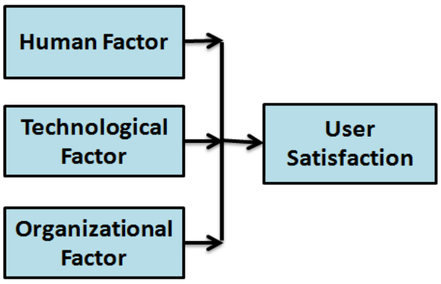

Fig. 1. Conceptual Model for US prediction

Data is collected by conducting the ERP user survey using the online Google form. Survey is based on the questionnaires which are designed using the 14 CSF's from earlier research. Total 28 responses were collected and out of them 4 were incomplete so, 24 responses were used for further analysis. And, 16 were used as a training dataset cases and 8 were used for test dataset cases. Analysis is done on the basis of variance explained by each CSF's to the output variable US. Based on the variance weights are assigned and factor score was calculated for each response.

A. *Prediction Model*

The prediction model used in this paper for US prediction by using existing methods ANN, ANFIS and proposed method KNN into a single programming model using MATLAB represented in figure 2.

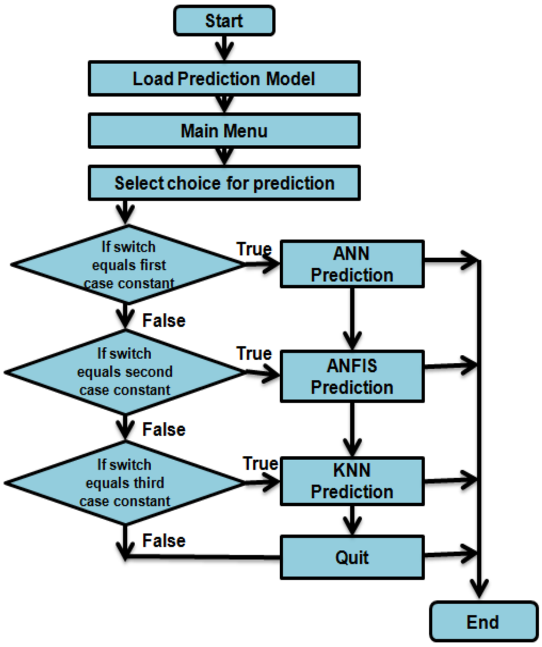

Fig. 2. Prediction Model for US prediction

B. *Artificial Neural Networks (ANN)*

This method mostly useful in classification problems and it is feed forward multi layer fully connected neural network [19]. The neural network contains three layers; input layer, hidden layer and output layer, so, it is a multilayer network. Each node is connected to all other nodes of other layers; so, it is a fully connected network. ANN is fundamentally non linear model that differentiate patterns and classify variables. For this, some training methods have been developed by researchers. Dasgupta *et al*. [20] gave the Levenberg Marquardt algorithm (trainlm) for training the network.

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

ANFIS is a hybrid technique which combines the adaptive learning capability of ANNs with the fuzzy logic of human reasoning and represented as a feed forward neural network. Therefore the benefits of the fuzzy system can be combined with a learning algorithm [21]. An ANN learns by adjusting the weights of interconnections between different layers. FIS uses fuzzy set theory, fuzzy reasoning, and fuzzy rules. ANFIS is a combination of five layers which performs various actions in it. All the five layers calculation is done as in the following steps:

Step1: Layer1

The output at layer1 for each membership function node is calculated using the function

$$O_{1,i} = \mu_{Ai}(X1) \quad \text{where, } i=1,2,3$$
$$O_{1,i} = \mu_{Bi}(X2) \quad \text{where, } i=1,2,3$$
$$O_{1,i} = \mu_{Ci}(X3) \quad \text{where, } i=1,2,3$$

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

Step2: Layer2

The output of this layer2 is calculated by multiplying all the incoming outputs from layer1 to corresponding nodes of layer2 as:

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

where, i=1,2,3.

Step3: Layer3

The output of the nodes of this layer is the function of inputs comes from the layer2 as following:

$O_{3,i} = \hat{W}_i = W_i / \Sigma W_i$

where, i=1,2,3.

Step4: Layer4

The output of the nodes of the layer4 is the function of all the inputs from layer3 and calculated as:

$O_{4,i} = \hat{W}_i f_i = \hat{W}_i(p_iX1+q_iX2+r_iX3)$

where, i=1,2,3 and, p_i, q_i, r_i are assigned weights by internal processing of it between 0 to 1 from input nodes to layer1.

Step5: Layer5

The output of this layer is the summation of all the outputs from layer4 which is calculated as:

$O_{5,i} = \sum\hat{W}_i f_i$ where, i=1,2,3.

D. K-Nearest Neighbor Classification (KNN)

KNN is an effective and simple algorithm that stores all available cases or sample dataset (training cases) and classifies the new cases (test cases) based on a similarity measure (e.g. distance functions). A case is classified or predicted by a majority vote of its neighbours (k is an integer to represent number of neighbour), with the case being predicted to the class most common amongst its k nearest neighbours measured by a distance function. The default distance function is used in KNN is Euclidean distance which is given in equation 1. If k=1 then the case is simply assigned to the class of its nearest neighbour [19].

Euclidean Distance

$$d_i(P_i,Q) = \sqrt{[\Sigma(P_i - Q)^2]} \quad (1)$$

where, $P_i = X_j$ and i=1,2,3,…..,n number of cases in sample dataset and j=1,2,3 is the number of factors or inputs $Q=X_j$ and j=1,2,3, number of input parameters.

E. Modelling of KNN

The steps for classification or prediction using MATLAB are as follows:

1. Load input and output train data for KNN Classification using following commands

   load('Inputs_KNN.mat');

   load('Outputs_KNN.mat');

2. 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.

3. 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

4. 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.

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

$$d_i(P_i, Q) = \sqrt{\left[\sum(P_i - Q)^2\right]}$$

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

$Q=X_j$ where j=1,2,3 and

i=1,2,3,……,n number of responses in training sample.

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

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

8. Prediction outcome is the majority vote between the selected neighbors.

9. Total error and average error are calculated using following equations.

$$\text{Total Error} = \sum_{i=1}^{n}(A_i - P_i) \quad (2)$$

$$\text{Average Error} = \frac{1}{n}\sum_{i=1}^{n}(A_i - P_i) \quad (3)$$

where, A_i is the actual US value and P_i is the predicted US value, n is the number of test cases in test dataset.

III. RESULTS AND DISCUSSION

The dataset used in this paper is calculated on the basis of variance explained by each CSF's to the output variable US. After, according the variance the weights are assigned to them and factor scores are calculated corresponding to the each response and observed US by responses is given in table 1.

TABLE I
FACTOR SCORE CALCULATED: SAMPLE DATASET

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

A. *Predicted US*

Predicted US using all prediction methods ANN, ANFIS and KNN is given in below tables.

.

TABLE II
PREDICTED US IN ANN

	Factors				
S.No.	Human	Tech.	Org.	US	Predicted US
1	5.08	5.20	5.49	5	5.6766
2	6.33	6.15	5.63	6	5.3010
3	6.08	4.98	6.23	5	4.2073
4	5.83	4.74	4.55	5	5.2353
5	2.33	1.79	2.03	1	1.1052
6	5.58	4.34	5.30	5	4.8945
7	6.33	6.15	5.99	6	6.1963
8	3.74	4.42	5.43	4	3.9680

TABLE III
PREDICTED US IN ANFIS

	Factors				
S.No.	Human	Tech.	Org.	US	Predicted US
1	5.08	5.2	5.49	5	4.9012
2	6.33	6.15	5.63	6	6.0868
3	6.08	4.98	6.23	5	5.7098
4	5.83	4.74	4.55	5	4.8213
5	2.33	1.79	2.03	1	0.6213
6	5.58	4.34	5.3	5	4.3565
7	6.33	6.15	5.99	6	6.0388
8	3.74	4.42	5.43	4	3.4321

TABLE IV
PREDICTED US IN KNN

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

B. *Model Comparison*

The comparisons of prediction of US using ANN, ANFIS and KNN is given based on the total error and average error calculated for all the test cases taken for prediction. Also comparing by predicted US in all the 8 test cases KNN given correct prediction in 6 cases. But in ANN and ANFIS there is error in all 8 predicted cases. So, when comparing based on the predicted outcome US on the test cases, KNN gives better and efficient prediction results as shown in above tables. Total error and average error is calculated using equation (2) and (3). Lower value of error indicates the good prediction model with efficient prediction accuracy. Also on the basis of total and average error KNN gives better results than ANN and ANFIS comparatively.

TABLE V
COMPARATIVE RESULTS

Methods	ANN	ANFIS	KNN
Total Error	4.8426	2.703	2
Average Error	0.605325	0.337875	0.25

IV. CONCLUSION

This study proposed the better prediction method for US in the process of ERP implementation. The prediction model considered the most important factors in the ERP system success which also affect US. These factors are Human, Technological and Organizational as predictor variables which are based on 14 CSF's taken from earlier study. US is taken as output variable which is dependent on predictor variables. Data for modelling and analysis were taken by conducting survey in the organization which implementing ERP system, using the questionnaire based on ex ante 14 CSF's. The collected data from respondents further used for study, analysis and prediction. We trained ANN, ANFIS and store the training dataset in KNN using 18 training sample dataset cases out of 24 sample dataset cases. Among the three techniques KNN outperforms well based on test cases, total error and average error. Our approach is novel in research related to prediction of US in ERP implementation process and its success. This will be very helpful for organizations to predict US in advance and to check the possibility of success or failure. This can be used as a decision making tool to support organizations to take corrective measures in advance and redirect the projects far in better direction. The future scope of this is to propose the efficient prediction technique than KNN with good prediction accuracy and less error. The utility of this prediction model can be further enhanced if it is supplemented with decision support system for organizations.

REFERENCES

[1] N. Garcia-Sanchez & Luis 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.

[2] W. E. and Chen, J. "Effects of internal support and consultant quality on the consulting process and ERP system quality", *Decision Support Systems*, Vol. 42 No. 2, pp. 1029-1042, 2006.

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

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] 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, 2014.

[11] 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.

[12] 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.

[13] S. Rouhani, and A. Z. Ravasan, "ERP success prediction: An artificial neural network approach", *Scientific Iranica, Transactions E: Industrial Engineering*, pp. 1-10, 2012.

[14] Luis Kalb 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.

[15] 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.

[16] 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.

[17] 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.

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

[19] 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, 1994.

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

[21] 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, 2003.