

PREDICTING THE EFFECTIVENESS OF WEB INFORMATION SYSTEMS USING NEURAL NETWORKS MODELING: FRAMEWORK & EMPIRICAL TESTING

Dr. Kamal Mohammed Alhendawi

Assistant Professor in Information Management, Al-Quds Open University, Gaza,
Palestine

Alhendawi.kamal@gmail.com or Hindawi.kamal@yahoo.com

ABSTRACT

The information systems (IS) assessment studies have still used the commonly traditional tools such as questionnaires in evaluating the dependent variables and specially effectiveness of systems. Artificial neural networks have been recently accepted as an effective alternative tool for modeling the complicated systems and widely used for forecasting. A very few is known about the employment of Artificial Neural Network (ANN) in the prediction IS effectiveness. For this reason, this study is considered as one of the fewest studies to investigate the efficiency and capability of using ANN for forecasting the user perceptions towards IS effectiveness where MATLAB is utilized for building and training the neural network model. A dataset of 175 subjects collected from international organization are utilized for ANN learning where each subject consists of 6 features (5 quality factors as inputs and one Boolean output). A percentage of 75% of subjects are used in the training phase. The results indicate an evidence on the ANN models has a reasonable accuracy in forecasting the IS effectiveness. For prediction, ANN with PURELIN (ANNP) and ANN with TANSIG (ANNTS) transfer functions are used. It is found that both two models have a reasonable prediction, however, the accuracy of ANNTS model is better than ANNP model (88.6% and 70.4% respectively). As the study proposes a new model for predicting IS dependent variables, it could save the considerably high cost that might be spent in sample data collection in the quantitative studies in the fields science, management, education, arts and others.

Keywords: Artificial Neural Network (ANN); Business Intelligence; Artificial Intelligence (AI); Forecasting; IS Effectiveness; Learning; MathLAB;

INTRODUCTION

Employment of WBIS as a tool for performing the managerial and task activities has a significant influence on the effectiveness, productivity of the organizations as well as the image of the organization itself (Alhendawi & Baharudin, 2013a; Panigyrakis & Chartzipanagiotou, 2006; Gichoya, 2007; Love & Irani, 2004). Despite the organizations spent too much cost in IT investment, organizations has adopted web based management information system in order to integrate the significant advantages of internet-based systems with organization functions, which leads to significant contributions to the organization effectiveness and individual performance (Alhendawi & Baharudin, 2013a, Abugabah & Sanzogni, 2010; Wang et al., 2007). In this context, it is essential to indicate that IS effectiveness can be defined as the IS value that could influence the users performance where is affected by the quality factors (i.e. the quality of technical, information and interaction characteristics) of the information system (Alhendawi & Baharudin, 2013a, 2013b; Al-adaileh, 2009).

However, the information systems studies have traditional used the questionnaires as the only tool in the evaluation of IS effectiveness measures such as questionnaires in evaluating the effectiveness of systems. As an important point, the full dependency on the use of the questionnaires in the IS assessment studies is considered as one of the significant limitations because the reliability of the results are highly correlated with the accuracy of respondents (Field, 2009; Hair et al., 2010).

In terms of AI tools, the Artificial Neural Network (ANN) is one of the most effective tools that could be used in forecasting and predicting the target values based on the learning process. ANN is a learning system based on a computational technique that can simulate the inputs of the upcoming sample subjects in order to forecast the target vale or the predicted result (Zhang et al., 2007). Such type of AI techniques is also has the ability to quantify the relationship between causal factors and users responses through the iterative training of the data obtained by experiment (Engelbrecht, 2007).

ANN can provide reasonable estimates on the predicted values within a considerably short time, and also, it can be utilized as a validation tool to ensure the accuracy of the findings of the quantitative studies in general, and IS studies particularly. Multilayer Layer Perceptron (MLP) is the most useful neural networks in function approximation. Also, this study is one of the fewest to discuss prediction of IS effectiveness using ANN towards enhancing the accuracy of decision making and IS developmental policy.

Therefore, there is a need for employing the data mining and Artificial Intelligence (AI) approaches, specially MLP, in the evaluation IS effectiveness (Alhendawi & Baharudin, 2013a) in order to respond to the existing significant gaps of cost, time, and accuracy of findings in the current IS studies.

IS EFFECTIVENESS MODEL

Based on the literature review (Alhendawi & Baharudin, 2013a; Petter et al., 2008; Petter & Mclean, 2009; Delone & Mclean, 2003), it can be concluded that quality factors such as system quality, information quality and service quality are the most important features in specifying the user satisfaction. Also, the AI tools such as ANN is considered as the most proper technique in the prediction of target value, i.e. user perception toward the IS effectiveness (Engelbrecht, 2007). Consequently, the input

dataset should include the quality factors as input features or determinants while the user perception should be received as a supervised target function. Figure 1 reveals the proposed conceptual model used in the prediction of the user satisfaction with IS.

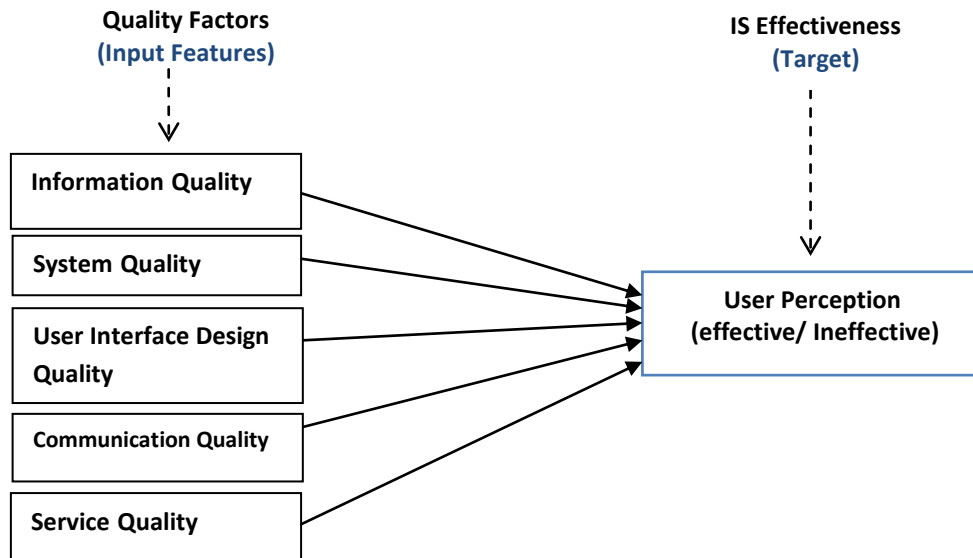


Figure 1. The Proposed Theoretical Framework for Predicting User Satisfaction

An extensive literature review was conducted to ensure that a comprehensive list of items was generated to assess the study variables included in the proposed model. Therefore, the researcher generated a scale with a reasonable reliability and validity level. The items are adapted from standard scales. Thus, it is essential to develop an instrument (i.e. questionnaire) to include all assessment measures. The questionnaire has six dimensions: five quality factors as independent variables and user perception as a dependent variable. Accordingly, there are causal factors or classifiers and one target. Regarding the values of object or target function (i.e. whether it is effective or ineffective) it depends on the mean of user perception towards IS value whether it is greater than 4.5 or not because the scale used in the data collection is 7-point Likert scale. If the value of user response greater than 4.5 this means “the system somewhat accepted” which means the system is effective. Otherwise the system is ineffective.

Artificial Neural Network

Artificial Neural Networks (ANN) have been considered as an effective AI tool for modeling the non-linear system and it is efficiently used for forecasting. Many studies indicated that ANN modeling is superior with respect to the traditional linear model (Yun et al., 1998). Furthermore, using ANN in system modeling and forecasting has many advantages (Yun et al., 1998; Zhang et al., 1998): (1) ANN has the ability to learn from examples, (2) it can predict and forecast even in case of noisy dataset, and (3) ANN can approximate multivariate functions with high accuracy, (4) ANN is a self-adaptive method and so it includes few pre assumptions, and (5) ANN can be utilized whenever there is a limited dataset or the relationship between input features and targets are vague. There are several types of ANNs such as Generalized Regression Neural Network (GRNN), Back Propagation (BP), and Radial Bias Function (RBF). Additionally, ANN may include input, hidden and output layers where the neurons in

one layer are connected fully or partially to the neurons in the next layer (Engelbrecht, 2007). The ANN Output (N_i) can be expressed as follows:

$$N_i(x_1, x_2, \dots, x_{m_i}) = a_i \left(\sum_{1 \leq j \leq m_j} (w_{ji} * X_j + b_j) \right);$$

Where x_j is the input signal, w_{ji} is the weight, a_j is the activation function and b_j bias weight.

Concerning the learning type, this study focuses on the learning situation, namely supervised learning, in which a set of input/output data patterns is available. Therefore, the ANN has to be trained to produce the desired output according to the examples. To conduct a supervised training we need a way of evaluating the ANN output error between the actual and the expected output. A popular measure is the Mean Squared Error (MSE) or Root Mean Squared Error (RMSE)

$$MSE = \frac{\sum (y_i - O_i)^2}{n}$$

$$RMSE = \sqrt{MSE}$$

Where y_i is the predicted value, o_i the observed value, and n is the number of data set. Multilayer Layer Perceptron (MLP) is the most useful neural networks in function approximation (Graupe, 2007). Figure 2 shows the multi-layer architecture of neural network.

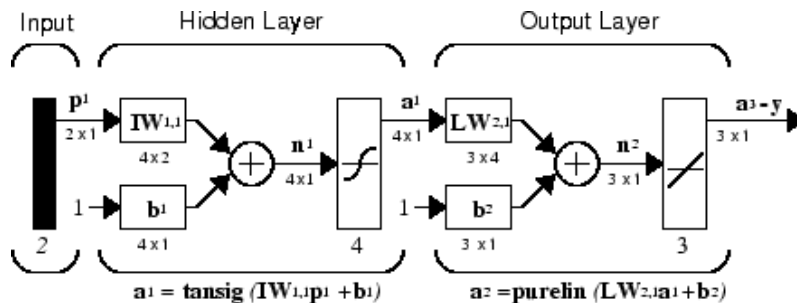


Figure 2. The multiple-layer networks

For multiple-layer networks the layer number determines the superscript on the weight matrix. As shown in Fig.2, the appropriate notation is used in the two-layer tansig/purelin network where tansig and pueline are the transfer functions used in ANN model for prediction purposes. This type of ANNs can be used as a general function approximator. It can estimate any function with a finite number of discontinuities arbitrarily well, given sufficient neurons in the hidden layer. Also, it is indicated by many researchers that this type of network is efficient estimating the non-linear relationship (Graupe, 2007). Therefore, in this study, the 2-layer ANN is employed in order to forecast the outputs based on the input data set.

For ANN training and design, there are many algorithms, which can be used for training ANN such as Back Propagation (BP). In terms of ANN training, Back Propagation algorithm is one of the most commonly used training algorithms in the field

of neural networks. According to Graupe (2007), BP algorithm is commonly used for the training of multi-layer Perceptrons (MLP).

The weights can be adjusted by back propagation gradient descent-based algorithm (GBA). GBA aims at measuring the output error, calculating the gradient of this error, and adjusting the ANN weights (and biases) in the descending gradient direction. Hence, this algorithm is a gradient descent local search procedure (Rumelhart et al., 1986).

THEORETICAL FRAMEWORK & METHODOLOGY

For clarity purpose, the author dedicates a section to explain the theoretical framework which is needed to fill the time cost and validation gaps. It also discusses the method that actually followed for forecasting IS effectiveness using ANN.

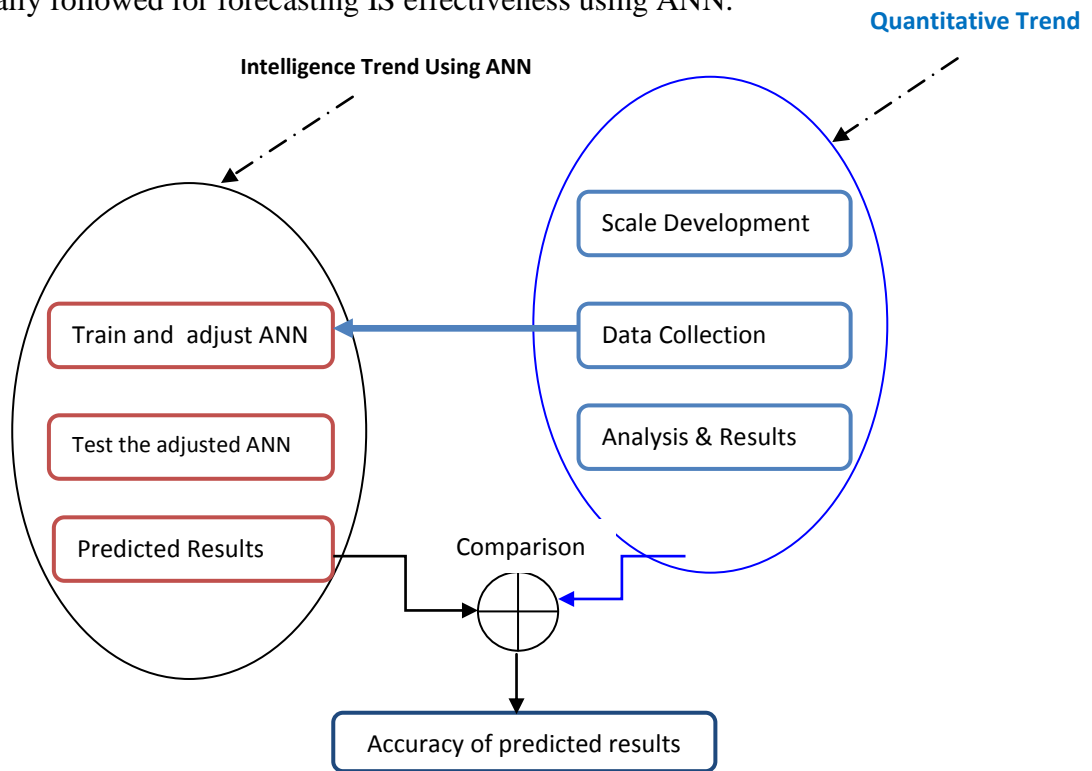


Figure 3. Theoretical Framework of ANN Prediction & Validation

As shown in Figure 3, the method of using ANN consists of several steps. At the beginning the users' responses toward the information system quality (i.e. system quality, information quality, service quality, user interface quality and communication quality) and effectiveness have to be collected through questionnaires for a small sample of 131 users to store their responses as a function of quality factors. The following equation explains the value of IS in terms of the five quality factors which are mentioned previously, and therefore, IS Effectiveness (ISE) = Function (5 Quality factors).

$$ISE = f(SQ, IQ, SERQ, UIQ, ComQx)$$

, where $0 \leq ISE \leq 1$

In the second step, ANN learning is started where the dataset of 175 subjects is divided into two datasets: training and testing dataset with percentages of 75% and 25% respectively. This means 131 subjects are used in the ANN training and 44 subjects are used in the testing phase. Figure 4 shows the steps used during the ANN training process.

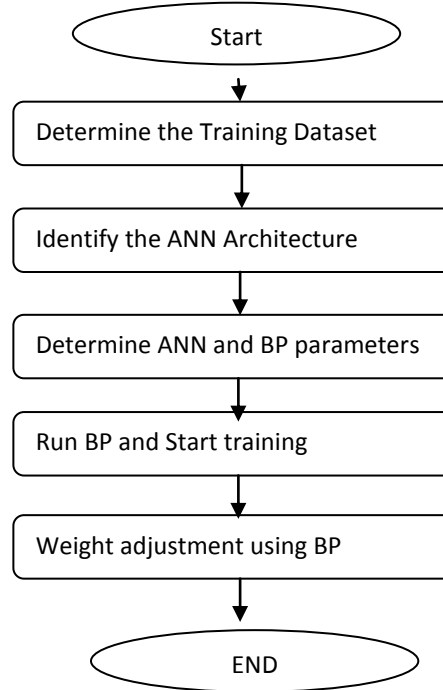


Figure 4. The steps of ANN Training

In the third step, after training the network the test phase is used for prediction to ensure the accuracy of ANN forecasting. The actual user perception toward quality and effectiveness is used to feed the ANN and to get the output regarding the upcoming users perception of the IS effectiveness. Finally, the predicted results or outputs should be compared with the collected values of the targets (i.e. IS effectiveness) in order to ensure the accuracy of using ANN in the prediction of the future users' perceptions.

EXPERIMENTAL RESULTS AND DISCUSSION

In this study, it is essential to identify the input / training dataset in addition to other parameters such as training algorithm and learning type because these variables influence the structure of ANN model, coefficients' weights and the results of ANN model. These parameters also affect the accuracy of the ANN forecasting. A model with so many parameters could result in data over fitting, and there, the early stop criteria is used to avoid the over fitting problem. In this study, six variables are used including five causal factors(system quality, information quality, service quality, user interface quality and communication quality) and one target variable called IS effectiveness. To investigate the best ANN model that could provide a reasonable forecasting, it is essential make comparison between the ANN models. Therefore, it seems that it is appropriate to divide the section into 4 sub sections including architecture, performance, correlation coefficients and accuracy of ANN prediction.

ANN Performance

For the ANN architecture, the multi-layer architecture is adopted and BP algorithm is used as a training algorithm. Concerning the learning type, it is more appropriate to use supervised learning as the measurement and the observation of the target function are known. IS effectiveness is suggested to be the target function and it has two values including “effective” or “ineffective”. MATLAB is used for building, training and testing the ANN model. Figure 5 reveals the architecture of the proposed multi-layer ANN.

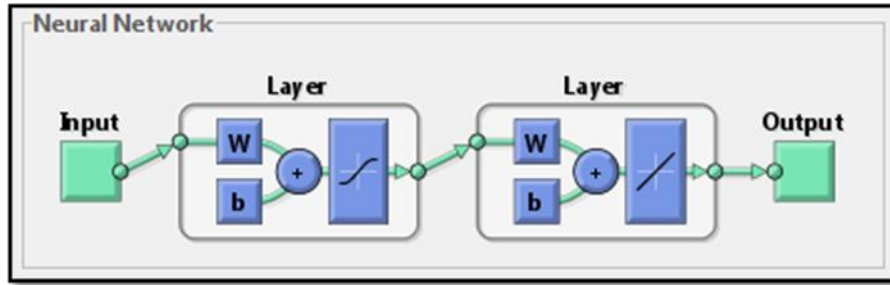


Figure 5. Neural Network (BK propagation approach with 2 Layers)

Generally, the prediction of outputs consists of two steps which are learning and testing steps. In the learning step, a model describing a predetermined set of concepts and parameters is created through analyzing a set of subjects or instances. An instance is supposed to be belong some predetermined group (i.e. effective or ineffective). The results of BP training algorithm indicates a reasonable level of accuracy in training, validation and testing. Figure 6 shows the mean square error of the three sets is relatively acceptable as the MSE of the training is the smallest one (Hudson et al., 2012).

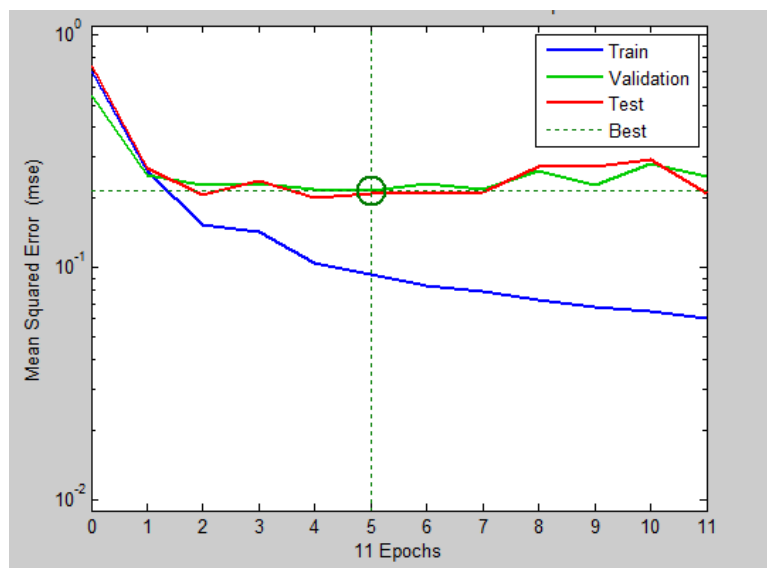


Figure 6. The performance of BK algorithm (PureLin Transfare).

As shown in Figure 6, the blue line which represents MSE decrease rapidly as ANN trained more. The green line which represents the validation process shows that validation error decreased as training increased. The test error is represented by red line where this line is closed to the validation error line, and therefore, there is evidence on the accuracy of using ANN with BP in predicting the future perceptions of system users. Since the test error and validation error are almost closed, this indicates that there is a reasonable division for the dataset. Figure 7 shows the performance of ANN in case of using the TANSIG as a transfer function.

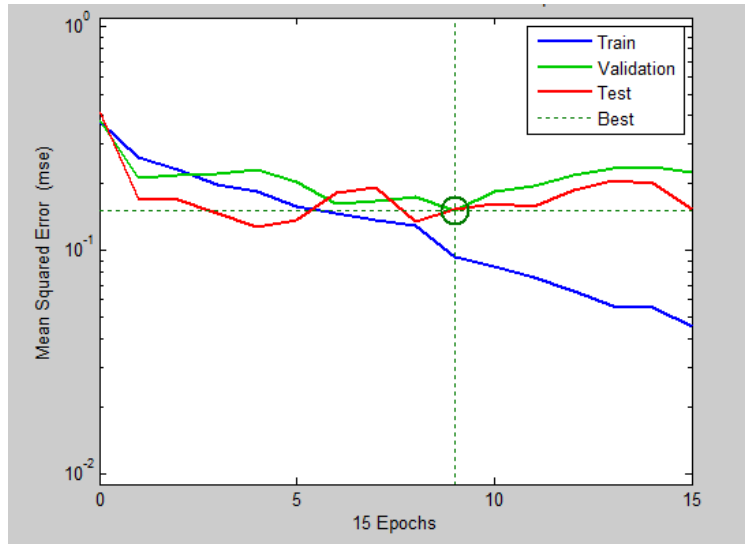


Figure 7. The performance of BK algorithm (TANSIG Transfare).

Based on Figures 6 and 7, it is clearly seen that the test error for the model of ANN with TANSIG (ANNTS) is lower than that of ANN model with PURELIN (ANNP).

Correlation coefficients of ANNTS & ANNP

Based on Figure 9, ANN with TANSIG transfer function has also a reasonable accuracy in forecasting the future targets as approximately there is a small variance between the validation and test errors. In order to decide on the optimal ANN parameters, it is important to find out the regression plot of the ANN models including BP with PURELIN and TANSIG transfer function. Figure 8 reveals the regression plot of ANN with PURELIN function.

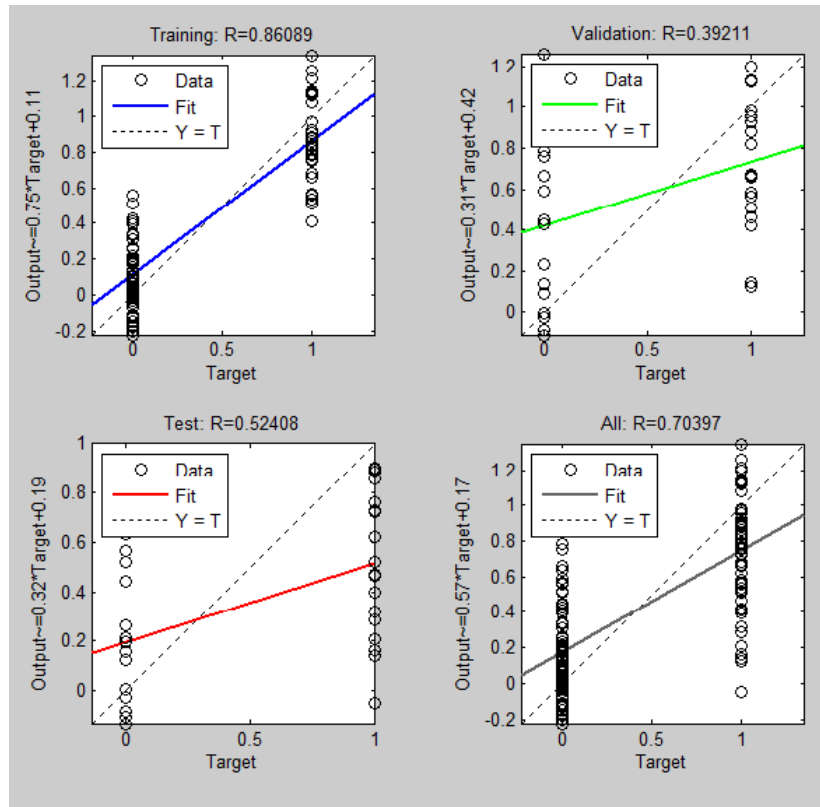


Figure 8. The regression Plot of BP with PURELIN function

Based on Figure 8, the correlation coefficients of training, validation and testing are 0.860, 0.396 and 0.524 respectively while the over correlation is 0.704. However, with regard to TANSIG the correlation coefficients of training, validation, testing and overall are 0.904, 0.502, 0.650 and 0.768 respectively Figure 9 demonstrates the regression plot of ANN with TANSIG.

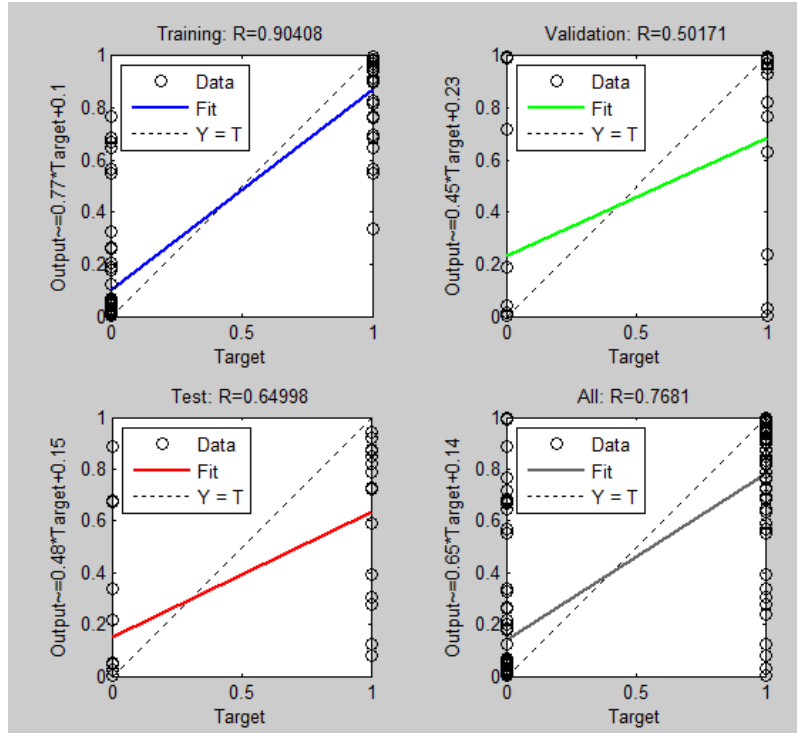


Figure 9. The Regression Plot of BK with Tansig Transfer

Hence, in terms of correlation coefficients, the ANN with PURELIN transfer function is better than ANN with TAGSIN, and consequently, ANNTS model provide better identification for the target values than ANNP.

ANN Foresting for IS Dependent Variables

After training the ANN models (ANNTS and ANNP), 25% of dataset of 175 is used in the testing phase. This means 44 out 175 subjects are used for testing the ANN model accuracy and predicting the user performance regarding the effectiveness of IS. Table 1 shows the accuracy of prediction for ANNTS and ANNP.

Table 1. Accuracy of IS effectiveness Prediction

Algorithm	Transfer Method	Total Dataset	Training	Testing	Correctly Identified	Correct Percentage
BP	PURELIN	175	131	44	31	70.4%
BP	TANSIG	175	131	44	39	88.6%

The predicted outputs of ANNP and ANNTS match 70.4% and 88.6% of the actual surveyed outputs (i.e. user perceptions), and therefore, the ANNP and ANNTS correctly forecasts 31 and 39 out of 44 subjects of the testing dataset. Therefore, the accuracy of ANNTS is also better than ANNP which indicates that ANNTS has the ability to correctly identify and forecast IS effectiveness more than ANNP. Table 2 demonstrates a sample of predicted output; where IQ, SERQ, UIQ, ComQ are the abbreviations of information quality, service quality, user interface quality and communication quality respectively.

Table 2. Sample of predicted outputs (IS Effectiveness)

SQ	4	6	6	6	5	6	4	6	4	4
IQ	4	5	4	6	5	5	4	6	5	6
SERQ	1	1	1	5	5	5	4	6	7	3
UIQ	5	6	6	5	6	5	4	6	6	5
ComQ	2	1	1	2	5	5	4	5	5	4
Actual Collected Values (Effectiveness)	0	0	0	1	1	1	0	1	1	1
Predicted Effectiveness Using ANNTS	0	1	0	1	1	1	0	1	1	1
Predicted Effectiveness Using ANNP	1	0	0	0	1	1	0	1	1	0

It can be concluded from Table 2 that the artificial neural network has the capability to predict at least a reasonable percentage of 70% of the unknown future targets (i.e. IS effectiveness). Also, it is found that results indicate that ANNTS has better forecasting than ANNP.

Implications

As this study attempts to employ the artificial intelligence and data mining approaches in the field of information system effectiveness and measures, there are several implications that could be drawn in the light of the obtained findings.

First, the artificial intelligence tools including ANN has the capability of correctly modeling the attributes of the information systems (SQ, IQ, SERQ, UIQ, ComQ, and effectiveness). Therefore, the results emphasized the existence of the relationships between the study factors of the proposed model in Fig1 which is consistent with Alhendawi and Baharudin(2013a). Additionally, the five quality factors of the input data set have successfully identified the IS effectiveness as a supervised target and this is in agreement with many IS researchers (Balaban et al., 2013;Alhendawi and Baharudin, 2013a, b; c).

Second, for the ANN modeling, ANNTS and ANNP scored a reasonable percentage of prediction (70.4% and 88.6% respectively) which means that ANN successfully forecasting the data occurrences of effectiveness as a dependent variable in IS field. Consequently, it can be concluded that ANNTS could provide an acceptable prediction to the effectiveness of Web based information systems such as E-learning and E-Government systems.

Third, the findings of this research indicate the suitability of ANNTS in building the ANN model for a supervised target (user satisfaction) with two values (effective and ineffective). The findings also provide evidence on the suitability of TANSIG as a transfer function because there is a considerable difference in the efficiency of ANN forecasting between PURELIN and TANSIG.

Fourth, since a model has been built that can predict the Web based information system effectiveness inside non-profit organization using ANN, it is

expected that the ANNTS model has the ability to provide a correct identification for the future values of the other dependent variables such as profitability, user satisfaction, and performance. In this respect, it seems that ANNTS could provide prediction for the dependent variables in the field of E-learning such as E-learner satisfaction and E-learner achievement.

Fifth, as this study is conducted on the prediction of Web based information system effectiveness in terms of IS value and contribution, this study will effectively assist in assessing and predicting two critical effectiveness factors: quality and value of the systems. Therefore, this study will attain a priority from the decision makers, IS managers and developers sides. Also, it will help the top management and IS developers in setting the development policy in the light of the assessment and forecasting results.

Sixth, as a new model is proposed for predicting IS dependent variables in general, this study contributes to decision makers and IS managers in various organizational contexts as it assists in the prediction of user perceptions and development.

Finally, for a large sample, it is worth mentioning that this model could provide a full perception regarding the future values of the targets such as IS value and satisfaction. Therefore it will save the high cost that might be spent in case of collecting the whole samples in the quantitative studies in different fields such as science, management, social, education and arts.

LIMITATION & FUTURE WORK

In terms of limitations, even ANN reveals the data pattern and associations between attributes or study variables; it does not sufficiently identify the causality relationship between them. Second, this model is applied within a non-profit organization so the encouraging findings need to be checked again in other contexts.

As a new research direction, it can be suggested to examine the ANNTS model in the prediction of other dependent variable such as user satisfaction with the Web system, loyalty, and innovation. Also, it is recommended to use this model in other contexts such as E-learning and E-government. Moreover, it is suggested to use the classification and clustering approaches in the prediction of other IS measures such as IS value and the impact of IS on the individuals and groups within organizations.

CONCLUSION

Most of IS studies followed the traditional approach (i.e. questionnaire) in the assessment of the dependent variables such as IS effectiveness. This costs too much time in questionnaire design, scale development, data collection and analysis. The main purpose of this study is to explore the applicability and capability of ANN for forecasting the user perceptions towards the IS effectiveness. In this study, the supervised learning is used as the measurement and the observation of the target function are known. The user perception towards IS effectiveness is suggested to be the target function and it has two values including “effective” or “ineffective”. Practically, two ANN models are experimentally tested: ANN with PURELIN and TANSIG transfer functions (ANNP and ANNTS). The findings indicate that ANNTS is more effective than ANNP in terms of correlation coefficients and accuracy of prediction. As

another important point, the results provide evidence on the efficiency of using ANN in the prediction of information system and management variables. Hence, we can derive the approximate value for the IS effectiveness based on quality factors including system, information, service, user interface and communication quality. This calculated effectiveness can be used to explore IS effectiveness for next users' perceptions. It is expected that this study would contribute to IS empirical studies as it presents a new approach for data validation and prediction. Also, this study will help the decision makers and IS developers to minimize the time, sample size and analysis efforts in order to rapidly respond to the IS development needs inside organizations.

REFERENCES

- Abugabah A, Sanzogni L (2010). Re-conceptualizing Information Systems Models: An Experience from ERP Systems Environment", *International Journal for Infonomics*, Vol. 3, No. 4, pp. 414-421.
- Al-adaileh RM (2009). An Evaluation of Information Systems Success: A User Perspective - the Case of Jordan Telecom Group". *European Journal of Scientific Research*, Vol.37, No. 2, pp. 226-239.
- Alhendawi K, Baharudin A. (2013b). Evaluating the Effectiveness of Web-based Management Information System from the Perception of Educationalists: An Exploratory Study. *Information Technology Journal*, Vol. 12, No. 6, pp. 1068-1078.
- Alhendawi K., Baharudin A. (2013a). The Impact of Interaction Quality Factors on the Effectiveness of Web-based System: The Mediating Role of User Satisfaction. *Journal of cognition technology & Work. Springer (DOI) 10.1007/s10111-013-0272-9*.
- D. Graupe, "Principle of Artificial Neural Networks", World Scientific Publishing Co. Pte. Lte., 2007.
- Engelbrecht, A. P. (2007). *Computational intelligence: an introduction*: Wiley. com.
- Field, A. (2009). *Discovering Statistics Using SPSS (3 ed.)*. London: Sage Publications Limited.
- Gichoya D. (2005). Factors affecting the successful implementation of ICT projects in government. *Journal of e-Government"*, Vol. 3, No. 4, pp. 175-184.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate Data Analysis: A Global Perspective (7 ed.)*. NJ: Prentice Hall.
- Hudson, B. M., Hagan Martin, T., & Demuth Howard, B. (2012). *Neural Network Toolbox™ User's Guide: MathWorks*.
- Love PED, Irani Z (2004). An exploratory study of information technology evaluation and benefits management practices of SMEs in the construction industry", *Information & Management*, Vol. 42, No. 1, pp. 242-2477.

- Panigyrakis GG, Chartzipanagiotou K (2006). The Impact of Design characteristics and support services on the effectiveness of Marketing Information systems: An empirical Investigation. *Review of Business Information systems*, Vol. 10, No. 2, pp. 91-100.
- Rumelhart, D. E., Hintont, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), 533-536.
- Wang Y.-S., Wang H.-Y., Shee DY (2007). Measuring e-learning systems success in an organizational context: Scale development and validation”, *Computers in Human Behavior*, Vol. 23, No. 4, pp. 1792-1808.
- Yun, S.-Y., Namkoong, S., Rho, J.-H., Shin, S.-W., & Choi, J.-U. (1998). A performance evaluation of neural network models in traffic volume forecasting. *Mathematical and Computer Modelling*, 27(9), 293-310.
- Zhang, G., Eddy Patuwo, B., & Y Hu, M. (1998). Forecasting with artificial neural networks:: The state of the art. *International journal of forecasting*, 14(1), 35-62.
- Zhang, J.-R., Zhang, J., Lok, T.-M., & Lyu, M. R. (2007). A hybrid particle swarm optimization–back-propagation algorithm for feedforward neural network training. *Applied Mathematics and Computation*, 185(2), 1026-1037.

Author Biography

KAMAL MOHAMMED ALHENDAWI has **Ph.D. degree in Management Information Systems (MIS)** in school of Computer Sciences, Universiti Sains Malaysia (USM).



He received Master degree from Islamic University, Gaza in Data Envelopment Analysis (DEA). He had received the B.Sc. in computer sciences from Abu Dies - Faculty of science and technology, Jerusalem University, Palestine. He has many publications (more than 15 indexed publications) during last 2 years in the field of management information system, information technology, artificial information systems, business intelligence, applied data mining, intelligent computing, and advanced statistical analysis. He has been working as a reviewer in several international journals with good reputation and index such as ELSEVIER, SPRINGER AND TAYLOR & FRANCIS. Prior to joining USM, He has

an academic experience in the fields of system analysis, information management, software engineering, object oriented programming, and network administration. During his work at United Nations for Relief and Works Agency (UNRWA), he got several rewarding and meritorious steps. Recently, he has assigned as HEAD ASESSMENT UNIT at UNRWA. His current research interests include information systems, information management, business information systems, knowledge management, applied data mining & business intelligence, E-Learning, mobile Learning, and intelligent wireless Network computing.