Attribute Importance Measure Based on Back-Propagation Neural Network: An Empirical Study

Boonyarat Phadermrod*, Richard M. Crowder, Gary B. Wills

Dept. of Electronics and Computer Science, University of Southampton, Southampton, United Kingdom.

* Corresponding author. Tel.: +44 61784159; email: bp6g12@soton.ac.uk

Abstract: Over the years, many different importance-performance analysis (IPA) variations have emerged as it is a primary tool for analyzing customer satisfaction. One of the recent IPA variations is back-propagation neural network based importance-performance analysis (BPNN based IPA) that utilizes BPNN to measure Importance. To investigate the performance of the BPNN based IPA, the authors compared two types of BPNN models that have one and multiple output neurons referred as BPNN (regression) and BPNN (classification) respectively, with multiple linear regression (MLR). This comparison demonstrates that the BPNN (regression) does not outperform MLR in term of model accuracy and training time, yet BPNN (classification) is superior to MLR and BPNN (regression) in term of model accuracy and predictive power. This finding leads to a reconsideration of the BPNN model used in the present BPNN based IPA.

Key words: Back-propagation neural network, empirical comparison, importance-performance analysis.

1. Introduction

Importance-performance analysis (IPA) is a technique proposed by Martilla and James in 1977 [1] for analyzing customer satisfaction towards a company's product or service. For a considerable period of time, IPA has been used as a tool for understanding customers' needs and desires so as to develop marketing strategies to respond to them. IPA is widely used in many areas that customer satisfaction is a key to a thriving business including higher education [2], tourism [3], government service [4] and health service [5]. IPA measures the satisfaction based on two components: importance and performance. The intersection of these two components creates a two-dimensional matrix that helps a company to identify improvement opportunities by discovering the attributes of company's product or service that should be maintained or improved based on the customers' viewpoints [6].

Generally, data regarding customer perceptions towards a product or service gathered via customer satisfaction surveys are examined for measuring importance and performance of company's product/ service attributes. Typically, method for measuring performance is well-established; by using direct rating from customers survey in which the customers are asked to rate the performance of the attribute ranging from "very dissatisfied" to "very satisfied" in a 5-point or 7-point Likert scale. Whereas importance can be measured either a rating scales (self-stated importance) or estimated on the basis of performance (implicitly derived importance) using statistical methods such as multiple linear regression analysis (MLR) [7] and partial correlation [8]. Recently, back-propagation neural network (BPNN) has become an alternative method for implicitly deriving importance which is called BPNN based IPA. It is reported that the

BPNN based IPA approach yields significant improvement in estimating of importance compared to conventional IPA and regression-based IPA approaches [9], [10], thereby several researchers [11]–[14] have applied and extended BPNN based IPA for a customer satisfaction improvement.

Different techniques for measuring the importance are likely to result in identifying dramatically different attributes for improvement. Therefore it is essential to determine the technique that provides the most accurate representation of consumers' perspective. The consequence of selecting the correct technique to measure importance motivated us to conduct an empirical comparison of BPNN (as the promising data mining techniques) with MLR (as the established statistical method) for measuring importance. The performance of each technique is examined using predictive accuracy, elapse time, predictive power, and diagnosticity.

Additionally, this paper suggests an alternative BPNN model with multiple output neurons referred as BPPN (classification) and compares its performance with the performance of BPNN model with one output neuron referred as BPPN (regression) that currently used to determine importance. Although, the BPPN (regression) is used in most BPNN based IPA publications as shown in Table 1 which describes structure of model and activation function exploited in each publication. The BPPN (regression) requires that input attributes and target attribute are continuous, which is not compatible with the discrete data type of the customer satisfaction surveys. In contrast, the BPPN (classification) is suitable for predicting a categorical (discrete) target attribute which is more compatible with the customer satisfaction surveys than the BPNN (regression). The comparison between these two types of BPNN model and MLR will reveal the differences between techniques and facilitates the importance measure techniques selection.

This paper is structured as follows. Section 2 identifies criteria for evaluating techniques for determining importance. Section 3 illustrates the empirical comparison of two different BPNN models and MLR. Section 4 concludes the paper.

Table 1. Publications That Implicitly Derived Importance Using BPNN

Author(s)	BPNN model (Input-Hidden-Output)	Activation Function (Hidden/Output-layer)
Deng, Chen, and Pei (2008) [9]	20-31-1	Hyperbolic tangent/ Hyperbolic tangent
Deng and Pei (2009)[11]	20-26-1	Hyperbolic tangent/ Hyperbolic tangent
Chen, Lin and Lin (2010) [12]	7-15-1	Sigmoid/ Sigmoid
Mikulić and Prebežac (2012) [13]	8-10-1	Hyperbolic tangent/ Hyperbolic tangent
Hosseini and Bideh (2013) [14]	7-7-1	Hyperbolic tangent/ Hyperbolic tangent
Krešić, Mikulić, and Kožić (2013) [10]	11-11-1	Hyperbolic tangent/ Identity

2. Evaluation Criteria

The results of different approaches including BPNN (regression), BPNN (classification) and MLR for deriving importance will be empirically compared by various criteria in order to identify possible significant differences between the approaches. The evaluation metrics are described as follows:

Predictive accuracy: This is a measure of how well the technique correlates the overall customer satisfaction (outcome) with the performance of product/service attributes in the data that has been provided. The predictive accuracy is the most important metric because the model with high predictive accuracy will generate the accurate importance and consequently produces the most accurate IPA result. The root mean squared error (RMSE) is used as a common measure for estimating predictive accuracy of regression and classification techniques.

Elapsed time training: This is the amount of time that each technique takes to construct the model on the training dataset. The elapsed time training is measured in seconds in which the minimum training time

is desirable. In cases where the input data is large, the elapsed time training is the one metric that has to be measured in order to determine the trade-off between the performance and time spent for training.

Predictive power of overall customer satisfaction: This predictive power is measured by correlating between the observed overall customer satisfaction (dependent variable) and the predicted overall satisfaction (independent variable) from the models of BPNN (regression), BPNN (classification) and MLR.

Generally, the value of correlation coefficients fall between −1 and +1 and the closer correlation coefficient is to 1 or to −1, the stronger correlation it is as noted by field [15]. It is expected that the overall customer satisfaction has positive correlations with the predicted overall satisfaction which means the value of observed overall customer satisfaction is moved in the same direction of the value of predicted overall satisfaction. Therefore, the model with higher correlation coefficient is superior at predicting overall customer satisfaction than the model with lower correlation coefficient.

Diagnosticity: This is an ability of the technique to identify just which company's attributes are most important in affecting customer satisfaction [16] in other word it is the ability of the technique to distinguish the most important attribute from the less important attributes. The test of diagnosticity is conducted following to Doyle *et al.* [17]'s approach for investigating characteristic of two directed (self-stated) importance measured methods. Diagnosticity of each technique is assessed as the regression coefficients of the regression equation tested on the basis of relationship between Importance and rank [17], as shown in (1).

$$I = \beta_1 R_1 + \beta_2 R_2 + \text{constant} \tag{1}$$

where I is Importance derived from a particular technique, R_1 is rank order importance of attribute ranged from 1 (most importance) to the number of attributes (least importance), β_1 is a linear coefficient for rank order (R_1), R_2 is residual of squared rank, and β_2 is a quadratic coefficient for residual of squared rank (R_2). Note that R_2 is computed by subtracting the squared rank with the resulting from regressing squared rank on rank order. R_2 is used instead of squared rank order for the purpose to reduce the collinearity between the rank and squared rank order which provided stable regression coefficients [17].

The linear coefficient (β_1) and quadratic coefficient (β_2) will be examined to compare the diagnosticity of each technique. The linear coefficient indicated linear relationship between the rank and importance measure since the Importance is decrease as the rank order of attributes increase, the linear coefficient should be a negative coefficient. The bigger linear coefficient, the steeper slope of the function relating importance to rank. The quadratic coefficient indicated non-linear relationship between the rank and importance which consist of concave (positive quadratic) and convex (negative quadratic) relationship [17]. The first suggests that the technique is able to distinguish the most important attribute and the technique with the larger quadratic coefficient has higher ability to diagnostic the attribute that most importance than the technique with a lower one. The latter suggests that the technique is lack of diagnosticity of the most important attribute in contrast it has more diagnostic the least important attribute.

3. Empirical Study

In this section, the results of different approaches including BPNN (regression), BPNN (classification) and MLR for deriving importance are investigated in order to identify possible significant differences between the approaches.

3.1. Dataset

2010 customer satisfaction survey of the media and telecommunication service provider company, known as ABC, dataset [18], is used as a case study. The dataset contains customer feedback from a

questionnaire with 81 questions and the number of customers responding to this questionnaire is 266. Two parts of data in the dataset are selected for conducting IPA in this case study including (1) assessment of overall satisfaction which is measured on a 5-point Likert scale; (2) assessment of overall satisfaction level of six company's attributes — equipment and system, sales support, technical support, supplies and orders, purchasing support, and contracts and pricing which measured based on a 5-point Likert scale.

3.2. Methodology

First, BPNN (regression), BPNN (classification) and MLR are implemented using Waikato Environment for Knowledge Analysis (WEKA) and a 10-fold cross validation is used for the purpose of model evaluation. The importance is then derived from the outcomes of these prediction models. Subsequently, the model performance of each technique and its importance measures are compared against the evaluation criteria.

Regarding the model training, the satisfactions of six company's attributes of the input dataset namely — equipment and system, sales support, technical support, supplies and orders, purchasing support, and contracts and pricing are selected as input attributes (independent variables) and the overall customer satisfaction is class attribute (dependent variable). Methods related to the model training and importance computation of each technique are given below:

Multiple linear regression: MLR is a simple technique for which a small number of parameters are required. Most of the parameters are set at their default value except parameter named *attributeSelectionMethod* where the value "*no attribute selection*" is assigned in order to display coefficients of all input attributes. Generally, the regression coefficient indicates how much a one unit increase in the independent variable results in an increase or decrease in the dependent variable with all other variables held constant [19]. Therefore, the regression coefficients can be referred to as implicit importance which expresses the influence of a company's attributes on the overall satisfaction. The importance derived from MLR is shown in Table 2.

Back-propagation neural network: The BPNN model is constructed for discovering the company's attribute that has the major influence on overall satisfaction. The BPNN contains three parts, including one input layer, one hidden layers, and one output layer. The performance of each of six attributes is the neurons in the input layer. For the hidden layer, the number of hidden-layer neurons is assigned by a training network with different configurations of hidden-layer neurons and selecting the best performing network. The output layer corresponds to the overall customer satisfaction.

There are two possible numbers to be assigned as the number of output-layer neurons depended on the type of problem: BPNN (regression) or BPNN (classification). The first has one output neuron whereas the latter has six output neurons representing possible class values of overall customer satisfaction {0-5}. Both BPNN structure models are trained using the *MultilayerPerceptron* classifier in WEKA. Logistic (sigmoid) function and identity function are used as the activation functions for hidden and output neurons in BPNN (regression) respectively, whereas logistic functions are used as the activation functions for all neurons in BPNN (classification).

The bound of neurons in the hidden layer are {7, 13}[1] and {13, 42}[1] for BPNN (regression) and BPNN (classification) respectively. Additionally, two formula for calculating number of hidden layer proposed by Shibata and Ikeda [20] and Sheela and Deepa [21] are also considered which yield 2 and 5 hidden neurons for BPNN (regression) and yield 6 and 5 hidden neurons for BPNN (classification) respectively. These several configurations of hidden-layer neurons are trained on the dataset in order to determine the number of hidden neurons of the two BPNN structure models using *Experimenter* application in WEKA. Specifically,

---

[1]Maren, Harston, and Pap (1990) cited in [9] demonstrated that the bound of neurons in first hidden layer was between $2N+1$ and $O(N+1)$, where N is the number of input neurons and O is the number of output neurons.

each given BPNN model is tested using 10-fold cross validation repeated 10 times. Subsequently, the paired T-test is conducted (with 5% significance level) to check the null hypothesis that the mean difference of performance measures between the BPNN models is zero.

BPNN (regression) is trained with the following parameters setting (learning rate = 0.7, momentum = 0.3 and number of cycles = 10000). Then the performances of several network configurations (with 2, 5, 7, 10, 12-13 hidden neurons) are measured by three indicators: the mean absolute error (MAE), the RMSE and goodness-of-fit (R^2). Note that, the MAE and RMSE approach to 0 indicate that BPNN model has precise prediction ability whereas R^2 close to 1 indicates that BPNN model has excellent goodness-of-fit [9]. The best performing network is the network with the two hidden neurons (Training–MAE = 0.588, RMSE = 0.770 and R^2 = 0.442; Testing–MAE = 0.617, RMSE = 0.802) since its RMSE value in training being the lowest one and its R^2 value being the highest one. The 6-2-1 network architecture is shown in Fig. 1(a) and its neural weights are used as input for calculating importance based on the extended Garson's algorithm [22]. Subsequently, the importance is produced and shown in Table 2.

BPNN (classification) is trained (with learning rate = 0.3, momentum = 0.2 and number of cycles = 500). Then the performances of several network configurations (with 5, 6, 13, 18, 22, 25, 28, 33, 35, 42 hidden neurons) are measured by three indicators: percentage of accuracy, RMSE and area under the Receiver Operating Characteristic (ROC) curve (AUC). With regard to RMSE and AUC in testing phase, two models with 5 and 6 hidden neurons are statistically different from other network structures at the 5% level of statistical significance. Consider the training performance, the neural network composed of 6 hidden neurons is selected (Training–accuracy = 65.53%, RMSE = 0.279; Testing–accuracy = 48.30%, RMSE = 0.342, AUC = 0.685) since its RMSE value is lower than the RMSE value of the 5 hidden neurons network and its accuracy is slightly greater than the accuracy of the 5 hidden neurons network. The 6-6-6 network architecture is shown in Fig. 1(b) and the importance is derived from its neural weights based on the extended Garson's algorithm [22], shown in Table 2.

Table 2. The Comparison of Importance Measured by Using Different Techniques

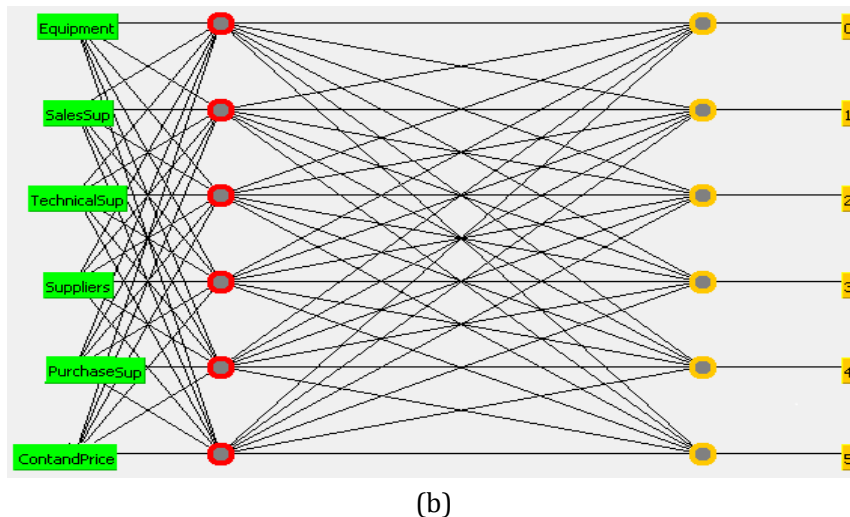| Company Service Attribute	MLR	BPNN (regression)	BPNN (classification)
F1-Equipment and System	0.529	0.790	1.000
F2-Sales Support	0.133	0.476	0.467
F3-Technical Support	1.000	1.000	0.660
F4-Supplies and Orders	0.413	0.560	0.304
F5-Purchasing Support	0.014	0.364	0.267
F6-Contracts and Pricing	0.466	0.716	0.763

(a)

(b)
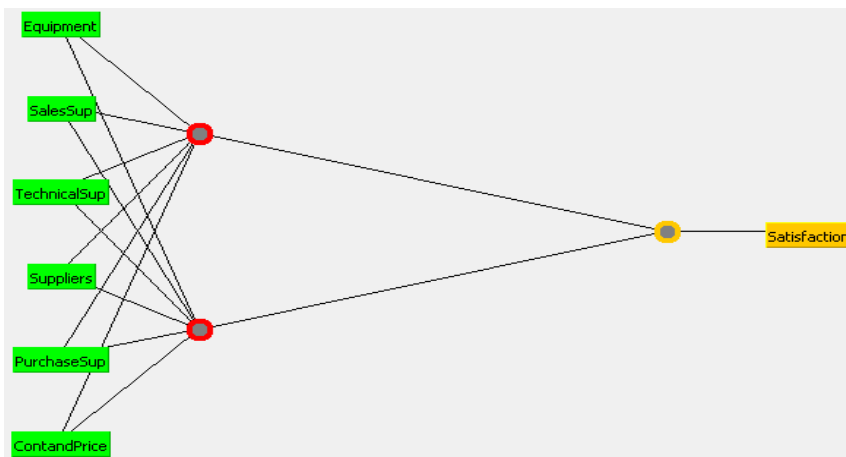
Fig 1. (a) The 6-2-1 neural network and (b) the 6-6-6 neural network on ABC company customer satisfaction dataset.

For a purpose of comparison of the three derived importance measures, all importance are expressed as a proportion of the highest value within each type of measure as shown in Table 2. Regarding this table, both MLR and BPNN (regression) identified the same ranking of importance and attribute "F3" is the most important attribute. All three techniques consistently identified attribute "F5" as the less important attribute. However, BPNN (classification) identified the different ranking of the top five important attributes from the first two techniques. It identified attribute "F1" as the most and attribute "F6" as the second most important attribute followed by attribute "F3".

3.3. Result

Predictive accuracy: The predictive accuracy of each technique was compared through RMSE measured in training and testing mode using 10-fold cross validation repeated 10 times, shown in Table 3. Additionally, the paired *T*-test on RMSE measured by using 10-fold cross validation was performed to check significant differences between RMSE of the techniques. As shown in Table 3, the predictive accuracy of BPNN (classification) is far better than predictive accuracy of MLR as well as BPNN (regression) since BPNN (classification) yields the lowest RMSE which is close to 0. The statistical comparison of RMSE shown in Table 4 confirms that RMSE of BPNN (classification) is statistically significant different from MLR and BPNN (regression), and the RMSE of these two techniques is not statistically significantly different.

Table 3. Comparison of RMSE

Technique	Training	Testing
MLR	0.775	0.792
BPNN (regression)	0.770	0.802
BPNN (classification)	0.279	0.342

Table 4. Result of Paired *T*-test on RMSE

Tests	*t* Stat	*t* Critical two-tail	*P(T<=t)* two-tail	Conclusion
MLR vs BPNN (regression)	-0.450	1.972	0.654	μMLR = μBPNN (regression)
MLR vs BPNN (classification)	28.581	1.983	> 0.0001 *	μMLR ≠ μBPNN (classification)
BPNN (regression) vs BPNN (classification)	-29.649	1.983	> 0.0001 *	μBPNN (regression) ≠ μBPNN (classification)

* Significantly difference at 5% significance level

Elapsed time training: The mean of elapsed training time was calculated by averaging elapsed time of each repetition of 10-fold cross validation shown in Table 5. Table 5 shows that BPNN takes a longer time to build model than MLR due to its approach that repeats network learning until the terminated condition is reached. Specifically, BPNN (regression) takes a longer time than BPNN (classification) since the number of training cycles of BPNN (regression) is greater than that of BPNN (classification).

Table 5. Comparison of Elapsed Time Training

Technique	Training time (seconds)
MLR	0.0008
BPNN (regression)	1.8115
BPNN (classification)	0.4062

Predictive power of overall customer satisfaction: With regard the correlation coefficient shown in Table 6, all techniques produced a positive correlation coefficient and the BPNN (classification) yielded the highest correlation coefficient followed by BPNN (regression) and the lowest correlation coefficient belongs to MLR. Thus it can be reasonably concluded that BPNN (classification) has a superior predictive power to BPNN (regression) and MLR.

Table 6. Comparison of Correlation Coefficient

Technique	Correlation coefficient
MLR	0.661
BPNN (regression)	0.666
BPNN (classification)	0.786

Diagnosticity: Regarding Table 7, linear coefficients of all techniques are statistically significant negative which indicate significant linear relationship between the ranks and the importance measures. Quadratic coefficients of all techniques are positive that means all techniques are able to distinguish the most important attribute however all quadratic coefficients are not significant. Unexpectedly, MLR has the largest linear and quadratic coefficients which means MLR has highest ability to diagnostic the attribute that most importance among techniques. Additionally, it can be concluded that BPNN (classification) has higher ability to diagnostic the attribute that most importance than BPNN (regression) since the linear and quadratic coefficients of BPNN (classification) are larger than that of BPNN (regression).

Table 7. Regression Coefficients for Rank and $Rank^2$ Residual

Technique	Rank (t)	$Rank^2$ Residual (t)
MLR	-0.176 (-6.094*)	0.016 (0.803)
BPNN (regression)	-0.122 (-16.803*)	0.008 (1.615)
BPNN (classification)	-0.150 (-15.724*)	0.014 (2.091)

* Indicates significant at $p < 0.05$

4. Summary and Discussion

This paper investigated the performance of the BPNN based IPA, the latest variation of IPA that applies BPNN to measure Importance. For this purpose, two types of BPNN models that have one and multiple output neurons referred as BPNN (regression) and BPNN (classification) respectively, and MLR were examined and compared against four evaluation metrics. The results are summarized in Table 8. in which performance in each criterion of each technique are ranked from 1 (best performance) to 3 (worst performance).

Table 8. Summary of Comparison Results across Evaluation Metrics

Evaluation metrics	MLR	BPNN (regression)	BPNN (classification)
Predictive accuracy	2	2	1
Elapsed time training	1	3	2
Predictive power of overall customer satisfaction	3	2	1
Diagnosticity	1	3	2

Based on this summary finding, it indicated that there is no clear winning technique since no one technique outperforms all the others. While MLR takes the shortest training time and yields the highest diagnosticity, it has the least predictive power of overall customer satisfaction and has low predictive accuracy. BPNN (classification) yields the best predictive accuracy and predictive power of overall customer satisfaction, but BPNN (classification) importance measures are less diagnostic (less able to identify customers' most important attributes) than that of MLR. Besides, it takes longer training time than MLR. However, our results show that the performance of BPNN (classification) is outperforms BPNN (regression) in all criteria. Especially, predictive accuracy of BPNN (classification) measured by RMSE is statistically significantly different and lower than of predictive accuracy of BPNN (regression) (see Table 3 and Table 4). Additionally, the predictive power of overall customer satisfaction measured by correlation and diagnosticity of BPNN (classification) is greater than that of BPNN (regression) and the BPNN (classification) takes shorter training time than the BPNN (regression).

In conclusion, this paper illustrates the methodology of BPNN based IPA and provides a set of evaluation metrics that can be used to determine technique that provides the most accurate representation of consumers' perspective thereby offers a sound direction for the business planning. Moreover, the empirical comparison of this paper demonstrates that the BPNN (classification) is superior to BPNN (regression). This finding raises a new research area of IPA, however it cannot strongly confirm that it is also true in different survey datasets. Thus, the comparison should be conducted across two or more survey datasets which is our future work.

References

[1] Martilla, J., & James, J. (1977). Importance-performance analysis. *Journal of Marketing*, *41(1)*, 77–79.

[2] Silva, F., & Fernandes, P. (2012). Empirical study on the student satisfaction in higher education: Importance-satisfaction analysis. *Management*, *293(42)*, 1192–1197.

[3] Taplin, R. H. (2012). Competitive importance-performance analysis of an Australian wildlife park. *Tourism Management*, *33(1)*, 29–37.

[4] Wong, M. S., Hideki, N., & George, P. (2011). The use of importance-performance analysis (IPA) in evaluating Japan's E-government services. *Journal of Theoretical and Applied Electronic Commerce Research*, *6(2)*, 5–6.

[5] Gonçalves, J. R., Pinto, A., Batista, M. J., Pereira, A. C., & Ambrosano, G. M. B. (2014). Importance-performance analysis: Revisiting a tool for the evaluation of clinical services. *Health*, *6(5)*, 285–291.

[6] Garver, M. S. (2003). Best practices in identifying customer-driven improvement opportunities. *Industrial Marketing Management*, *32(6)*, 455–466.

[7] Matzler, K., & Sauerwein, E. (2002). The factor structure of customer satisfaction: An empirical test of the importance grid and the penalty-reward-contrast analysis. *International Journal of Service Industry Management*, *13(4)*, 314–332.

[8] Matzler, K., Sauerwein, E., & Heischmidt, K. (2003). Importance-performance analysis revisited: The

role of the factor structure of customer satisfaction. *The Service Industries Journal*, *23(2)*, 112–129.

[9] Deng, W., Chen, W., & Pei, W. (2008). Back-propagation neural network based importance–performance analysis for determining critical service attributes. *Expert Systems with Applications*, *34(2)*, 1115–1125.

[10] Krešić, D., Mikulić, J., & Kožić, I. (2013). Artificial neural network-based applications in travel and tourism research : A review and case study. *Proceedings of International Critical Tourism Studies Conference*.

[11] Deng, W.-J., & Pei, W. (2009). Fuzzy neural based importance-performance analysis for determining critical service attributes. *Expert Systems with Applications*, *36(2)*, 3774–3784.

[12] Chen, L., Lin, S., & Lin, Y. (June 2010). Application of IPA and back-propagation neural network to build a new service quality decision making model. *International Journal of Information and Management Sciences*, *21*, 315–330.

[13] Mikulić, J., & Prebežac, D. (2012). Accounting for dynamics in attribute-importance and for competitor performance to enhance reliability of BPNN-based importance–performance analysis. *Expert Systems with Applications*, *39(5)*, 5144–5153.

[14] Hosseini, S. Y., & Bideh, A. Z. (2013). A data mining approach for segmentation-based importance-performance analysis (SOM–BPNN–IPA): A new framework for developing customer retention strategies. *Service Business*, *1977*.

[15] Field, A. (2009). *Discovering Statistics Usinsg SPSS* (3rd ed.). Sage Publications.

[16] Gustafsson, A. (2004). Determining attribute importance in a service satisfaction model. *Journal of Service Research*, *7(2)*, 124–141.

[17] Doyle, J., Green, R., & Bottomley, P. (1997). Judging relative Importance: Direct rating and point allocation are not equivalent. *Organizational Behavior and Human Decision Processes*, *70(1)*, 65–72.

[18] Kenett, R., & Salini, S. (2011). The ABC annual customer satisfaction survey. *Modern Analysis of Customer Satisfaction Surveys: With Applications Using R* (pp. 45–62).

[19] Nathans, L., Oswald, F., & Nimon, K. (2012). Interpreting multiple linear regression: A guidebook of variable importance. *Practical Assessment, Research & Evaluation*, *17(9)*, 2–19.

[20] Shibata, K., & Ikeda, Y. (2009). Effect of number of hidden neurons on learning in large-scale layered neural networks. *Proceedings of ICCAS-SICE* (pp. 5008–5013).

[21] Sheela, K. G., & Deepa, S. N. (2013). Review on methods to fix number of hidden neurons in neural networks. *Mathematical Problems in Engineering*, *2013*, 1–11.

[22] Glorfeld, L. (1996). A methodology for simplification and interpretation of backpropagation-based neural network models. *Expert Systems with Applications*, *10(1)*, 37–54.

Boonyarat Phadermrod was born in Bangkok, Thailand in 1982. She received her BSc degree in computer science from the Prince of Songkla University, Thailand with a second class honour in 2003 and a MEng degree in computer engineering from the Kasetsart University, Thailand in 2008. She is awarded full scholarship by the Royal Thai Government to study in a graduate program to earn the PhD degree in computer science, beginning in 2012 until completion.

Currently, she is a PhD student in computer science at the University of Southampton. She is also a lecturer at the Department of Computer Engineering, Kasetsart University (Kamphaengsaen campus), Thailand. Boonyarat's research interests are focus on data mining and database, business intelligent and knowledge management.

Richard M. Crowder was born in Macclesfield, Cheshire, UK in 1953. He received his BSc degree in engineering and PhD degree in electrical engineering from the University of Leicester in 1974 and 1977 respectively.

Following a period in machine tool manufacturing, he joined the academic staff of the University of Southampton in 1982 and is currently an associate professor in the Agents, Interaction and Complexity Research Group, within electronics and computer science.

His research interests are in the application of information technology to manufacturing industry and robotics, and have published over 150 papers in this area.

In addition Dr. Richard is the lead lecturer for a number of modules taught primarily to electrical and electromechnical engineers, as well as the examinations officer for electronics and computer science. He is also a member of the Institution of Engineering and Technology.

Gary B. Wills graduated from the University of Southampton with an honor's degree in electromechanical engineering, and then a PhD degree in industrial hypermedia systems.

Currently, he is an associate professor, in computer science at the University of Southampton. Gary's research interests are focus on system engineering and he works across disciplines with colleagues from industry, medicine, psychology, nursing, social worker and education.

Dr. Gary is a chartered engineer and a member of the Institute of Engineering Technology. He is also a principal fellow of the Higher Educational Academy. He is also an adjunct professor at the Cape Peninsular University of Technology.