Exploring Hidden Relationships within Students' Data Using Neural Network and Logistic Regression

Fadzilah Siraj, Nooraini Yusoff and Noorlin Mohd Ali

Faculty of IT
Universiti Utara Malaysia, 06010 Sintok
Kedah, Malaysia
Tel: +604-9284672, Fax: +604-9284753, E-mail: fad173@uum.edu.my, nooraini@uum.edu.my, s85468@ss.uum.edu.my

ABSTRACT

Considerable attention has been given to the development of sophisticated techniques for exploring data sets. One of the most commonly used techniques is neural networks that have the abilities to detect nonlinear effects and/or interactions. Due to the reduced interpretability of the output model of neural networks, the same data set has been analyzed using logistic regression. In this study, both techniques have been applied to education data set. The study aims to provide some insight into first year students undertaking undergraduate programs, namely Bachelor of Information Technology (BIT), Bachelor of Multimedia (BMM) and Bachelor in Management of Technology (BMoT) at Universiti Utara Malaysia. The Holland Personality Model was used to indicate the students' personality traits in conjunction with students' academic achievement. Although there is some differences in percentage of accuracies in both methods, the methods were used in this exploratory study in a complementary manner.

Keywords

Neural Network, Logistic Regression Holland Model, Education, Personality Traits

1.0 INTRODUCTION

Considerable attention has been given to the development of sophisticated techniques for exploring data sets. One of the most commonly used techniques is neural networks (NNs) that have the abilities to detect nonlinear effects and/or interactions.

NNs have been used extensively in pattern recognition, speech recognition and synthesis, medical applications (diagnosis, drug design), fault detection, problem diagnosis, robot control, electric load forecasting (Lopes *et al.*, 2003; Pai and Hong, 2005), electricity demand forecasting (Mamun & Nagaska, 2004), stock forecasting (Pérez-Rodriguez *et al.*, 2005), inflation forecasting (Nakamura, 2005) and computer vision (Zhang, *et al.*, 1998; Hong and Weiss, 2001). One major application areas of NN is forecasting (Liao & Fildes, 2005; Yu *et al.*, 2005) and these techniques have been widely touted as solving many forecasting problems (Marquez *et al.*, 1992; Heravi *et al.*, 2003; Lam, 2004; De Andrés *et al.*, 2005; Huang *et al.*, 2005).

NN models can learn from experience, generalize and "see through" noise and distortion, and able to abstract essential characteristics in the presence of irrelevant data (Wasserman, 1989). NN model is also described as a 'black box' approach, which has great capacity in predictive modelling, i.e. all the characters describing the unknown situation must be presented to the trained NNs, and the identification (prediction) is then given (Lek and Gue´gan, 1999). Lippman (1987) indicated that NN models provide a high degree of robustness and fault tolerance because of the number of processing nodes, each having primarily local connections. NNs techniques are also advocated as a replacement for statistical forecasting methods because of its capabilities and performance (Marquez *et al.*, 1992; Siraj and Asman, 2002; Siraj and Mohd Ali, 2004).

Badri (1999) used a neural network based decision support system to advise the business school student to select their major (marketing, finance, management information system, and general business). Three layers of backpropagation system were applied to predict the final GPA of all courses taken by a student after the student has completed their study from the college level. From the experiments, the study shows that NNs model performed adequately on the basis of the statistical criteria that had been identified in the study. In addition, NNs models operate under the assumption that current and future students' performance will be not differ from past trends given various characteristics of students. NNs system also has its reliability and validity compared to other traditional statistical models.

In another study, Barker *et al.* (2004) explore the student data from the University of Oklahoma in classifying student graduation behaviour (either students obtained undergraduate degree within 6 years or fail to do so) by considering academic, demographic, and attitudinal variables. The study introduces the use of NNs and support vector machines, which both are nonlinear discriminant methods in order to overcome the classification problem.

NNs have many theoretic properties, specifically their ability to detect nonlinear effects and/ or interactions. Due to the reduced interpretability of the model output of neural networks, logistic regression was also employed in this study. The regression analysis model is also known as one of the most useful tools in quantitative analysis phase of the decision-making process (Marquez *et al.*, 1992; Thomas & Galambos, 2004). It is generally used to predict future values based on past values by fitting a set of points to a curve (Dunham, 2003; Kleinbaum *et al.*, 1998). It is most often used when the independent variables cannot be controlled, as when they are collected in a sample survey or other observational study. In other words, it is equally applicable to more controlled experimental situations. The regression analysis model is also known as one of the most useful tools in quantitative analysis phase of the decision-making process (Marquez *et al.*, 1992).

The concept of personality psychology is basically extended from the meaning and the concept of personality (Funder, 2001). The understanding of the concept begins with the assumption that each individual is different and these differences reflect an underlying organization (Mischel *et al.*, 2001). As defined by Pervin (1996), personality is described as the complex cognitions organization, affects and behaviours that give direction and pattern (coherence) to the person's life. As an analogy from the human body, personality consists of both structures and processes and therefore reflects both nature (genes) and nurture (experience), including the effects of memories of the past, as well as constructions of the present and the future.

The main objective of personality psychology is to understand the individuality; while Funder (2001) describes the mission of personality psychology as the requirement in order to account for the individual's characteristic patterns of thoughts, emotions, and behaviour together with the psychological mechanisms. Personality psychology has been grouped into six levels which are Trait Level, Psychodynamic-Motivational Level, Phenomenological Level, Behavioral-Conditioning Level, Social Cognitive Level, and Biological Level (Mischel *et al.*, 2003). Since this study emphasizes on the use of Holland Model, therefore the concentration of this study is on trait theories.

The suitability of Holland Model in specifying individual personality is proved through studies in many domains. Adib-Hajbaghery and Dianati (2005) conduct a study to indicate a student's personality for the admission into nursing schools. The study was conducted because of the personality problem among the nursing students that should have in order to fulfill the nurse profession requirement. As a result, it has caused the nursing leaders to think the necessity of considering student's personality during the admission process. The main objective of the study is to assess freshman nursing student's personality characteristics and their compatibility with the demands of the nursing profession. The study was conducted at Tehran and Kashan medical universities and one of the branches of Azad University. The respondents of the study are 52 freshman nursing students that have been assessed using Holland's Vocational Interests Inventory. From the assessment, the result shows that 44% of the respondents did not have appropriate personality characteristics for the nursing profession. The study concludes that Holland personality model can help select the appropriate student for the nursing schools.

In the tourism domain, Frew and Shaw (1999) apply Holland's personality theory to effectively explain individuals' visitation of, and likelihood of visiting, tourist to 500 sample of adults that have been collected which covers the Holland's Self-Directed Search, past visitation of specific attractions, interest in visiting such attractions and the visitation plan in future. The study discovered that there was significant association between the respondents' Holland personality types, their gender and their tourism behaviour, based on some attractions and some measures of behaviour. Thus, the result gives some implication for tourism marketing and management levels.

The validity of Holland Model as a tool for facilitating personality performance is agreed by researchers in psychology domain (Prediger, 2000; Prediger and Vansickle, 1992). A study conducted by Prediger and Vansickle (1992) of 3612 4-year college alumni who had taken a RIASEC based interest inventory approximately 8 years earlier, plotted 51 career groups on the hexagon by applying theory-based (i.e. hexagon-based) weights to mean RIASEC interest scores for the groups. As a result, career groups were generally located near Holland type to which they belong.

The Holland's hexagon typology of personalities and work environments which is also known as RIASEC theory displays a classification model of people and work environments using six types; Realistic (R), Investigative (I), Artistic (A), Social (S), Enterprising (E), and Conventional (C).

2.0 METHODS

Several experiments will be conducted in this study in order to explore the relationship of student personality and the course undertake. DeLurgio (2000) and Kaastra and Boyd (1996) concluded the steps involved in evaluating Neural Network performance as shown in Figure 1.

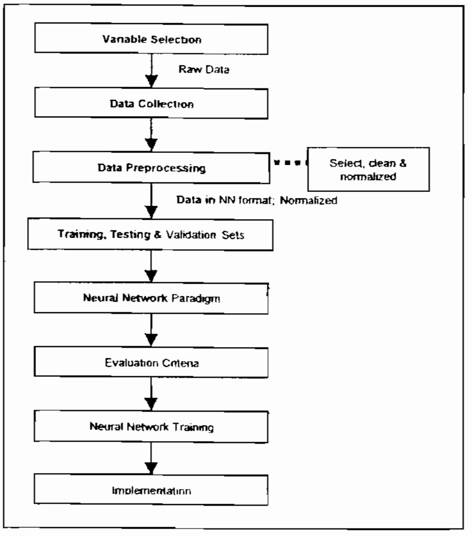

Figure 1: The Steps in Performing Neural Network Experiments

The raw data set for this study was collected and the respondents of the study were the first semester students from Faculty of Information Technology and Faculty of Technology Management of Universiti Utara Malaysia. In this study, the instrument used is the questionnaire and convenient sampling is adopted. The questionnaire contains the questions on student's background, academic achievement and personality traits questions based on Holland personality test.

According to Stone and Thorton (1995), the preprocessing of the raw input data may appear to be a trivial task, but can have a major impact on the performance of the learning. The preprocessing stage is mainly to ensure that the statistical distribution of values for each input and output is roughly uniform; otherwise, NN will not produce accurate forecasting task with incomplete, noisy, and inconsistant data.
Data normalization is important to ensure that the distance measure accords equal weight to each variable, as well as to speed up the learning phase. It also helps to avoid computational problems (Lapedes and Farber, 1988), to meet the algorithm requirement (Sharda and Patil, 1992) and to facilitate network learning (Srinivasan *et al.*, 1994). The process is performed on a single data input to distribute the data evenly and scale it into an acceptable range for the network. In this study, the most common normalization technique is used; where the values are scaled such that the minimum value goes to 0, the maximum value goes to 1, and other values are scaled accordingly. This type of scaling is known as linear scaling. The output or the dependent variable is classified as 1 for Bachelor of Information Technology (BIT), 2 for Bachelor of Multimedia and 3 for Bachelor of Technology Management.

3.0 RESULTS

For the initial analysis, the *Sijil Tinggi Persekolahan Malaysia's* (*STPM*) results which are equivalent to A level examination results were used as independent variables and the program undertaken at Universiti Utara as the dependent variable. Each *STPM*'s subject was coded as a single attribute. A total of 210 questionnaires were collected, however some questionnaires were not included in the analysis since some information was missing. Out of these questionnaires, 168 were identified to be used as datasets for building the model (refer to Table 1 and also Fig. 2).

Table 1: The Total number of respondents based on the selected undergraduate program

Program	Total	Percentage
BIT	80	47.62
BMM	20	11.91
MOT	63	37.50
BEDU IT	5	2.97
Total	168	100

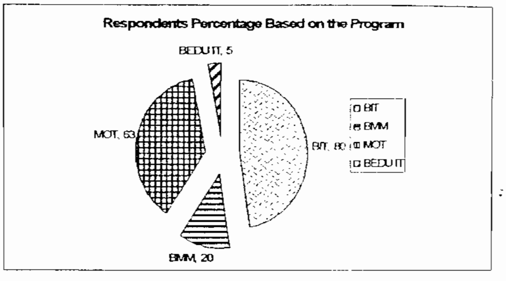

Figure 2: The Percentage of Respondents by program

Initially, the number of BIT students is 85, the *STPM*'s results of 5 students were not completed. Therefore, those questionnaires were not considered in this study.

The class distribution of the students undertaking the Bachelor of Education in IT is very small (2.97%), hence these respondents were also not considered. Therefore, a total of 163 respondents were only chosen in this study. This indicates that the composition of the respondents was 50.30%, 11.66% and 38.04% for BIT, BMM and MOT respectively. Further analysis using descriptive techniques indicate that none of the respondent took Physics and Chemistry subjects at STPM's level. Hence, the subjects considered are *Bahasa Melayu (BM), Pengajian Am (PA), Sejarah (Sej), Geografi (Geo), Ekonomi (Eko), Matematik (Maths), Perkomputeran (Perkomp), Sastera, Syariah, Usuluddin, Seni, Seni Visual (SV), Bahasa Cina (Cina), Bahasa Tamil (Tamil),* and *Bahasa Arab*.

In this study, the subjects are considered as independent variables and the program undertaken by the students at the university is considered as dependent variable. Prior to obtaining prediction model from the data set, further investigation was carried out on each independent variable and the results are shown in Fig. 3.

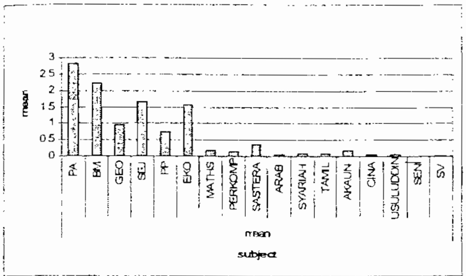

Figure 3: Mean Value of STPM examination subjects

The bar chart indicates that there are two subjects that have mean greater than 2 (*PA* and *BM*), two subjects have mean between 1 and 2 (*SEJ* and *EKONOMI*), two subjects have mean between 0.5 and 1 (*GEO* and *PP*), and the rest of the subjects have mean lower than 0.5. In other words, there are only six subjects that have mean greater than 0.5.

Initial attempt was to understand the behaviour of the dependent variable with respect to the independent variable. For this purpose, the experimental results are shown in Table 2.

Table 2: The percentage of accuracy of NN and Logistic Regression

Method	
NN	Logistic regression
56.25%	44.20%

Based on Fig. 3, 11 out 17 subjects have mean lower than 0.5. Due to this reason, subjects were grouped to reduce the effect of having a lot of zeros as input for independent variables. Subjects such as *bahasa Arab, Tamil* and *Cina* are combined as one entity whereas *Seni* is combined with *Visual*. Prior to this decision, the raw data was explored to ensure that none of the students took up the subjects at the same time. The mean distribution of the subjects after some of them are combined is shown in Fig. 4.

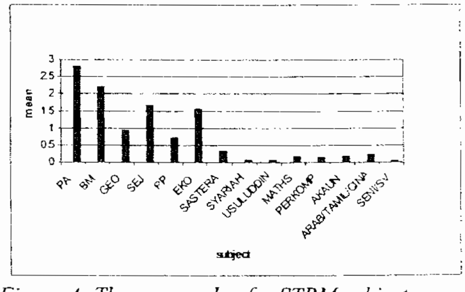

Figure 4: The mean value for STPM subject after combining

The impact of combining some subjects do not increase the number of subjects whose mean is 0.5. Further analysis on the prediction modeling methods were carried out and the results are exhibited in Fig. 5.

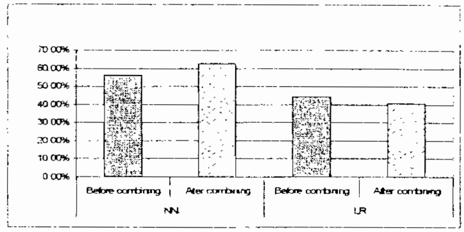

Figure 5: The percentage of before and after combining subject

Clearly the impact of reducing zeros in the datasets has affected the accuracy results of both methods. For neural networks, the accuracy increases by 6.25%, on the other hand the logistic result decreases by 3.7%.

Based on logistic regression results, subjects such as *PA, BM, GEO, SEJ* and *Perkomp* have significant contribution to BIT program since their significant values are less than 0.05. For BMM program, only *PA* has significant contribution to this program. Subjects such as Geo and Sej are only significant at 10% level.

The highest percentage of accuracy obtained so far is 62.5%. One possible way to increase the performance of NN is by excluding variables that have small input values. The mean values for each STPM's subject with respect to the three programs are illustrated in line graph shown in Fig. 6.

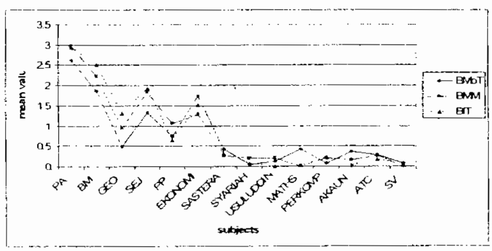

Figure 6: Mean value of BMOT student

The graphs indicate that subjects such as Maths and *Seni/ Seni Visual* have values zeros for *BMM*, and subject such as *Usuluddin* for BIT has also zero value. Among the subjects that have mean less than 5, only *PERKOM* has a significant contribution to BIT program. For this reason, this subject will be considered as one of the subjects to be included in the prediction model apart from *PA, BM, GEO, SEJ, PP, EKONOMI* and *SASTERA*. The results of considering these values in the prediction model are depicted in the following Table 3.

Table 3: The result of NN and Logistic Regression with and without the combination of Perkom subject

Method			
NN		Logistic Regression	
With *PERKOM*	Without *PERKOM*	With *PERKOM*	Without *PERKOM*
56.25%	54.17%	32.6%	30.3%

The results shown in Table 3 indicate that the NN performance increases by 6.25% when the PERKOM is excluded in the model. On the other hand, the accuracy produces by Logistic Regression reduces by 2.6%. Further analysis on the Likelihood Ratio Tests shows that the effect of *PA, BM, GEO* and *SEJ* are significant on the prediction model with significant values less than 0.05.

To get some insight into the information captured using Holland Model, a total of 35 questions were tested against the program undertaken by the student. The pseudo R square of the logistic regression model obtained 49.2% accuracy whereas NN model achieved 50.0%. Further investigation was carried out to determine the effect of Holland model on the program when combined with the results (refer to Table 4).

The results displayed in Table 5 reveal the effect of considering *Perkomp* as on of the attribute in the datasets. When *Perkomp* is included in the datasets, the accuracy of NN decreases by 5% and 6.25% respectively. However, the logistic regression result increases by 4.5% when the average psychological test for Holland model was considered. One interesting finding is that the logistic regression obtained higher accuracy when the STPM's results were combined with the Holland model. Thus, the highest accuracy obtained so far is 87.9% with logistic regression model.

Table 4: The comparison of both method with the combination of result and Holland Model

Method (COMBINING WITHOUT PERKOMP SUBJECT)			
NN		Logistic Regression	
STPM's Results with all questions with respect to Holland model	STPM's Results with average personality type with respect to Holland model	STPM's Results with all questions with respect to Holland model	STPM's Results with average personality type with respect to Holland model
68.75%	56.25%	86.8%	35.3%
Method (COMBINING WITH PERKOMP SUBJECT)			
NN		Logistic Regression	
STPM's Results with all questions with respect to Holland model	STPM's Results with average personality type with respect to Holland model	STPM's Results with all questions with respect to Holland model	STPM's Results with average personality type with respect to Holland model
62.5%	62.5%	87.9%	39.8%

Based on the results discussed so far, both NN and Logistic Regression results can be improved by integrating the Holland model with the STPM's results. As a conclusion, the Neural Network model that achieved the highest accuracy (62.5%) with an architecture of 42-4-3 with learning rate 0.1, momentum rate 0.3 and epoch 100. For logistic regression, the percentage of 87.9% can be further investigated by considering the significant values of the independent variables.

Based on the results in Table 6, most BIT students have arts background, except a few who have sat for *PERKOM* subject. This shows none of the respondent is from the *AGAMA* school. The earlier analysis shows that some students who took up *Syariah* and *Usuluddin* are currently undertaking the MOT program.

As for the Holland Model, it appears that BIT students are more artistic since 50% of the questions that measure the personality type is significant (significant value = 0.029, 0.037 and 0.031). In addition, the BIT students are realistic (33.33%) and investigative (33.33%). On the other hand, the results also reveal that none of questions on social personality type is significant.

Table 6: The result of Logistic Regression to the selected dataset

SIGNIFICANT VALUES <0.05		BIT
PA	.003	0.011
BM	.000	0.07
GEO	.002	0.03
SEJ	.000	0.01
EKONOMI	.031	0.018
0.05 <SIGNIFICANT VALUES <0.10		
PERKOMP	.057	.049
ARTISTIC		**BIT**
I like to see arts exhibition, theatre and film.	.000	0.029
I like to express my feelings and thoughts on paper through drawing, musics or building things.	.000	0.037
I like to use my imagination.	.053	0.031
REALISTIC		
I like to repair damage items.	.000	
I like to do things with my hands.	.000	
I like to do outdoor activities.	.000	0.013
I am comfortable working with machine tools.	.027	0.016
SOCIAL		
I am happy when I am with other people.	.000	
I like to be opened with people.	.000	
I like to pay attention to what other people needs.	.000	
INVESTIGATIVE		
I like solving problems.	000	0.036
I like to learn something new.	.046	0.026
I like to find a solution to the problem by my own.	.000	
ENTERPRISING		
I like when other people carry out the tasks I assign to them.	.000	
I like to take risk.	000	
I like to carry out heavy tasks that involve the use of energy and able to overcome pressure.		0.065
CONVENTIONAL		
I like to manage project, idea and people		0.04

4.0 DISCUSSION

Based on the experiment, the result on prediction accuracy using NN is 62.5% and logisctic regression achieves 87.9%. Apart from NN, the validity of Holland Personality Model is also identified as a tool in indicating student personality and their interest. The model obtain from this study could be used by the university's management for providing better education plans.

Mining students' data are not easy. It is very tedious and requires a lot of effort in performing the analysis on the data. However, with a Data Mining tool can assist and accelerate data mining activities.

In this study, the data mining techniques have been applied to education data. The findings from the study indicate that by using data mining, the hidden information within the students' data was uncovered. The information covered may be useful for the management to plan strategically for offering special courses to cater for the students who do not have sufficient background to undertake specific programs.

REFERENCES

Adib-Hajbaghery, M., & Dianati, M. (2005). Undergraduate nursing students' compatibility with the nursing profession. BMC Medical Education, 5 (25), pp. 1-5.

Badri, M.A. (1999). A neural network based DSS for advising students in the school of business Neural Networks. *IJCNN '99. International Joint Conference*. 5, pp. 3533 -- 3538.

Barker, K., Trafalis, T., & Rhoads, T.R. (2004). Learning from student data *Systems and Information Engineering Design Symposium, 2004. .Proceedings of the 2004 IEEE.* pp. 79 - 86.

De Andrés, J., Landajo, M., & Lorca, P. (2005). Forecasting business profitability by using classification techniques: A comparative analysis based on a Spanish case.*Electric Power Engineering, PowerTech Budapest 99.*

DeLurgio, S. A. (2000). *Forecasting Principle and Applications.* McGraw-Hill.

Dunham, M. H. (2003) *Data Mining: Introductory and Advanced Topic.* Upper Saddle River, NJ: Prentice Hall.

Frew, E. A., & Shaw, R.N. (1999). The relationship between personality, gender, and tourism behavior *Tourism Management*, 20 (2), pp. 193-202.

Funder, D. C. (2001). Personality . *Annual Review of Psychology*, 52, pp. 197-221.

Heravi, S., Osborn, D. R., & Birchenhall, C. R. (2004). Linear versus neural network forecasts for European industrial production series. *International Journal of Forecasting*, 20 (3), pp. 435-446.

Hong, S. J., & Weiss, S. W. (2001). Advances in Predictive Models for Data Mining. *Pattern Recognition Letters*. 22, pp. 55-61.

Huang, W., Nakamori, Y., & Wang, S.-Y. (2005). Forecasting stock market movement direction with support vector machine. *Computers & Operations Research*, 32 (10), pp. 2513-2522.

Kaastra, I., & Boyd, M. (1996). Desigining a Neural Network for Forecasting Financial and Economic Time Series. *Neurocomputing*, 10, pp. 215-236.

Kiang, M. Y. (2003). A comparative assessment of classification methods. *Decision Support Systems*, 35(4), pp. 441-454.

Kleinbaum, D. G., Kupper, L. L., & Nizam, A. (1998). Applied Regression Analyisi and Other Multivariate Methods. Pacific Grove, California US: Thomson Publishing Inc.

Lam, M. (2004). Neural network techniques for financial performance prediction: integrating fundamental and technical analysis. *Decision Support Systems*, 37 (4), pp. 567-581.

Lapedes, A., & Farber, R. (1988). How neural nets work. In: Anderson, D.Z., (Ed.). *Neural Information Processing Systems, American Institute of Physics, New York*, pp. 442–456.

Lek, S., & Guégan, J. F. (1999). Artificial Neural Networks as a Tool in Ecological Modeling, An Introduction. *Ecological Modeling, 120*, pp. 65-73.

Liao, K. P., & Fildes, R. (2005). The accuracy of a procedural approach to specifying feedforward neural networks for forecasting. *Computers & Operations Research*, 32 (8), pp. 2151-2169.

Liao, S. H. (2003). Knowledge management technologies and applications-- literature review from 1995 to 2002. *Expert Systems with Applications*, 25 (2), pp. 155-164.

Lippmann, R.P. (1987). An introduction to computing with neural neural network. *IEEE Transactions on nets, IEEE ASSP Magazine*. April, pp. 4–22.

Lopes, M.L.M., Lotufo, A.D.P., & Minussi, C.R. (2003). A fast electric load forecasting using adaptive neural networks. *Power Tech Conference Proceedings 2003*, 1, pp. 23-26.

Mamun, M. A. & Nagasaka, K. (2004). Artificial neural networks applied to long-term electricity demand forecasting. *Proceedings of HIS 2004 Fourth International Conference*. pp. 204 – 209.

Marquez, L., Hill, T., O'Connor, M., & Remus, W. (1992). Neural network models for forecast a review. *IEEE proceedings of the 25th Hawaii International Conference on System Sciences*, 4, pp. 494–498.

Minaei-Bidgoli, B., A. Kashy, D., Kortemeyer, G., & F.Punch, W. (2003). Predicting Student Performance: An Application of Data Mining Methods with the Educational Web-Based System LON-CAPA. *Proceedings of 33rd ASEE IEEE Frontiers in Education Conference, November*, pp. 5-8.

Mischel, W., Shoda, Y., & Smith, R. E. (2001). *Introduction to personality*. Hoboken, NJ: John Wiley & Sons Inc.

Nakamura, E. (2005). Inflation forecasting using a neural network. *Economics Letters*, 86 (3), pp. 373-378.

Pai, P. F., & Hong, W. C. (2005). Support vector machines with simulated annealing algorithms in electricity load forecasting. *Energy Conversion and Management*, 46 (17), pp. 2669-2688.

Pérez-Rodríguez, J. V., Torra, S. & Andrada-Félix, J. (2005). STAR and ANN models: forecasting performance on the Spanish "Ibex-35" stock index. *Journal of Empirical Finance*. 12 (3), pp. 490-509.

Pervin, L. A. (1996). *Personality: theory and research*. Canada, US: John Wiley & Sons, Inc.

Prediger, D. J. (2000). Holland's Hexagon Is Alive and Well—Though Somewhat out of Shape: Response to Tinsley. *Journal of Vocational Behavior*, 56 (2), pp. 197-204.

Prediger, D. J., & Vansickle, T. R. (1992). Locating occupations on Holland's hexagon: Beyond RIASEC. *Journal of Vocational Behavior*, 40, pp. 111-128.

Sharda, R., & Patil, R. B. (1992). Connectionist Approach to Time Series Prediction: An Empirical Test. *Journal of Intelligent Manufacturing*, 3, pp. 317-323.

Siraj, F., & Asman, H. (2002). Predicting Information Technology Competency Using Neural Networks. *Proceedings of the 7th Asia Pacific Decision Sciences Institute Conference*, pp. 249 – 255.

Siraj, F. & Mohd Ali, A. (2004). Web-Based Neuro Fuzzy Classification for Breast Cancer. *Proceedings of the Second International Conference on Artificial Intelligence in Engineering &* Technology, pp. 383 – 387.

Srinivasan, D., Liew, A. C., & Chang, C. S. (1994). A Neural Network Short-Term Load Forecaster. *Electric Power Systems Research*, 28, pp. 227-234.

Stone, J. V., & Thorton, C. J. (1995). Can Artificial Neural Networks Discover Useful Regularities? *IEE Artificial Neural Networks Conference*, 409, pp. 201-205.

Thomas, E. H., & Galambos, N. (2004). What Satisfies Students? Mining Student-Opinion Data with Regression and Decision Tree Analysis. *Research in Higher Education*, 45 (3), pp. 251-260.

Wasserman, P. D. (1989). Neural Computing: Theory and Practice, Van Nostrand-Reinhold, New York.

Yu, L., Shouyang Wang, S., & Lai, K. K. (2005). A novel nonlinear ensemble forecasting model incorporating GLAR and ANN for foreign exchange rates. *Computers & Operations Research*, 32 (10), pp. 2523-2541.

Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks:: The state of the art. *International Journal of Forecasting*, 14 (1), pp. 35-62.