Full reference: Abdou, H., Alam, S. & Mulkeen, J. (2014) ' Would credit scoring work for Islamic finance? A neural network approach', *International Journal of Islamic and Middle Eastern Finance and Management, 7* (1), pp. 112-125.

WOULD CREDIT SCORING WORK FOR ISLAMIC FINANCE? A NEURAL NETWORK APPROACH

Hussein A. Abdou*
The University of Huddersfield, Huddersfield Business School, Huddersfield, West Yorkshire, UK, HD1 3DH

Shaair T. Alam
The University of Salford, Salford Business School, Salford, Greater Manchester, UK, M5 4WT

James Mulkeen
The University of Salford, Salford Business School, Salford, Greater Manchester, UK, M5 4WT

Abstract

Purpose – The main aim of this paper is to distinguish whether the decision making process of the Islamic financial houses in the UK can be improved through the use of credit scoring modeling techniques as opposed to the currently used judgmental approaches. Subsidiary aims are to identify how scoring models can reclassify accepted applicants who later are considered as having bad credit and how many of the rejected applicants are later considered as having good credit; and highlight significant variables that are crucial in terms of accepting and rejecting applicants which can further aid the decision making process.

Design/methodology/approach – A real data-set of 487 applicants are used consisting of 336 accepted credit applications and 151 rejected credit applications make to an Islamic finance house in the UK. In order to build the proposed scoring models, the data-set is divided into training and hold-out sub-set. The training sub-set is used to build the scoring models and the hold-out sub-set is used to test the predictive capabilities of the scoring models.70 percent of the overall applicants will be used for the training sub-set and 30 percent will be used for the testing sub-set. Three statistical modeling techniques namely Discriminant Analysis (DA), Logistic Regression (LR) and Multi-layer Perceptron (MP) neural network are used to build the proposed scoring models.

Findings – Our findings reveal that the LR model has the highest Correct Classification (CC) rate in the training sub-set whereas MP outperforms other techniques and has the highest CC rate in the hold-out sub-set. MP also outperforms other techniques in terms of predicting the rejected credit applications and has the lowest Misclassification Cost (MC) above other techniques. In addition, results from MP models show that monthly expenses, age and marital status are identified as the key factors affecting the decision making process.

Research limitations/implications – Although our sample is small and restricted to an Islamic Finance house in the UK the results are robust. Future research could consider enlarging the sample in the UK and also internationally allowing for cultural differences to be identified. The results indicate that the scoring models can be of great benefit to Islamic finance houses in regards to their decision making processes of accepting and rejecting new credit applications and thus improve their efficiency and effectiveness.

Originality/value –Our contribution is the first to apply credit scoring modeling techniques in Islamic Finance. Also in building a scoring model our application applies a different approach by using accepted and rejected credit applications instead of good and bad credit histories. This identifies opportunity costs of misclassifying credit applications as rejected.

Key words Islamic finance; decision-making processes; credit scoring techniques; neural networks
Paper type Research paper

*** Correspondence Author**
Dr. Hussein A. Abdou
Professor of Finance & banking
The Huddersfield Business School, University of Huddersfield
Huddersfield, West Yorkshire, HD1 3DH, UK
Tel.: +44 1484473872
Fax: +44 1484473148
Email: h.abdou@hud.ac.uk

1. Introduction

Islamic finance can be referred to as a banking system or it can be related to any other fiscal activity that is consistent with *Shari'ah* or Islamic law. *Shari'ah* law is derived from The Holy *Qur'an* which was revealed to Prophet Mohammed (pbuh) by Allah over 1400 years ago. The Holy *Qur'an* is a book of guidance to all of mankind and within it are certain laws which mankind has to follow in order to protect them from harm. An example of this is that Muslims are not allowed to invest in businesses which are contrary to Islamic values such as alcohol, pork or pornography.

One of the most notable factors of *Shari'ah* law in an Islamic economy is the prohibition of interest (*Riba*) which can be translated into 'excess' or 'additional' which is also considered to be a form of usury. Interest rates is an example of *Riba* which can neither be paid nor consumed by Muslims and is considered to be a serious offence. In chapter 2, Surah al-Baqrah, verse no. 275 of The Holy *Qur'an* states:

"*Those who eat Riba (usury) will not stand (on the Day of Resurrection] except like the standing of a person beaten by Shaitan (Satan) leading him to insanity. That is because they say: "Trading is only like Riba (usury)," whereas Allah has permitted trading and forbidden Riba (usury).*

This is one of the reasons why many Muslims stay away from conventional banking as it is run solely on interest. In Islam, money cannot be created from money; instead money must be backed by an asset. Thus, an Islamic economy is built on sharing profit and loss where there are buyers and sellers and not borrowers and lenders. In contrast conventional banks use interest rates which are both socially and morally undesirable within *Shari'ah* law. The embargo of interest rates in Islam has enabled Islamic banking and financial institutions to develop over the last couple of decades. This has been evident in the recent financial crisis where various Islamic financial institutions escaped the full brunt of the recession whereas conventional banks faced a very uncertain and bleak trading period amid huge losses and government bailouts. Although the UK is predominantly dominated by conventional banks, which often means that many Muslims do not have an alternative way to obtain capital; there is an increasing number of 'Islamic-like' financial houses that have been established around Muslim communities in the UK. This has enabled individuals and businesses to borrow money without having to incur any interest rates or fees. This type of activity is also popular in the Middle East and in particular Asia and Africa where there are numerous financial houses set up to meet the needs of Muslims and Non-Muslims.

Due to the lack of resources and funds within Islamic finance houses in the UK, only a limited number of applications are able to join financial schemes. Acceptance to such schemes is often based on the applicant's credit report which is designed to identify those applicants who are most creditworthy and those who pose the least risk. Whilst the current scoring technique certainly helps the financial house to minimize their risk, there is still room for improvements. The rest of this paper is organized as follows: section 2 reviews relevant literature; section 3 outlines the research methodology; section 4 discusses the results; and section 5 concludes the research findings and suggests areas for future research.

2. Literature review

The lack of currently available literature on credit scoring in the Islamic banking and finance sector has resulted in the application and review of general literature on credit scoring in cconventional banks and other financial institutions. Although there is no universally accepted definition of credit scoring, many definitions highlight a credit score as being an essential appraisal tool that uses different statistical techniques to distinguish between good and bad creditworthy customers. For example, whilst Hand and Jacka (1998) defined credit scoring as the model of assessing the creditworthiness of an applicant, implying that a 'good' customer is expected to pay on time whereas a 'bad' customer is expected to fail to pay on time. Anderson (2007) defined credit scoring in two parts: firstly, credit basically means to buy now and pay later and secondly, scoring means to numerically rank something and then have the ability to discriminate between the numbers as stated by Abdou and Pointon (2011). As such, the resulting credit score reflects the applicant's credit history and report.

The importance of credit scoring has become more critical within the banking sector due to the fast growth of the credit industry, the need to manage loan portfolios, the need to reduce the cost of the credit evaluation and the need to enhance the efficiency and effectiveness of the credit decision-making process. Decision-making involving accepting or rejecting a client's credit can be supported by both judgemental techniques and credit scoring models.

Many theorists including Al Amri (2002) have highlighted the importance of using credit scoring models in evaluating credit risk. However, Al Amri (2002) also argues that there is no optimal method. This indicates that one type of scoring model might work for specific financial institutions but fails to work in others. He also reflects on other factors used in determining the creditworthiness

of a customer, for example to what extent is a customer classified as good or bad which can measured via statistical techniques. This approach is much more efficient and consistent in comparison to the judgemental approaches which are currently utilized by the finance house. Abdou and Pointon (2011) stated that the judgemental approach relies heavily on the past and present experiences of the credit analyst who have to judge whether the customers has the ability to repay a loan within a certain time. Consequently, judgemental approaches are associated with "subjectivity, inconsistency and individual preferences motivating decisions". Judgemental approaches have some strengths, such as "taking account of qualitative characteristics and having a good track record in evaluating past credit by utilizing the wealth of the credit analyst's past experience" (Abdou & Pointon, 2011, p. 61). Lee *et al* (2002) and Ong *et al* (2005) explained that an effective credit scoring model can have a positive impact on a bank as it can reduce the costs involved in the credit process and also reduce the prospect of bad loans. They also suggested that it improves the decision making process and can save the bank a lot of time and money as better decisions can be made.

Applications of credit scoring have been widely used in different fields, including: accounting and finance (Altman and Narayman, 1997; Pendharkar, 2005; Landajo et al., 2007), marketing (Kumar et al., 1995; Thieme et al., 2000; Chiang et al., 2006), and general applications (Walczak and Sincich, 1999; Usha, 2005; Nikolopoulos et al., 2007). In the area of accounting and finance, credit scoring applications have been used for different purposes including: bankruptcy prediction (Zhang *et al.* 1999; Tsai and Wu, 2008; Etemadi et al., 2009; Nanni and Lumini, 2009); scoring applications (Huang et al., 2006; Crook et al., 2007; Huang et al., 2007; Erbas and Stefanou, 2008)) and classification problems (Ong et al., 2005; Laha, 2007; Ben-David and Frank, 2009).

There are a number of key determinants of credit scoring which including: characteristics of a client such as gender, age, marital status, educational level, occupation, time at present job and having a credit card (Banasik et al., 2003; Sustersic et al., 2009; Abdou & Pointon, 2011); other loan characteristics such as loan amount, loan duration, monthly income, bank accounts, purpose of loan and guarantees (Sarlija et al., 2004; Ong et al., 2005; Abdou & Pointon, 2011); statistical techniques such as logistic regressing, CART and neural networks (Lee & Chen, 2005; Abdou & Pointon, 2011); scoring methodology and the predictive ability of the model (Al Amri, 2002; Abdou & Pointon, 2011) and different sampling methods (Banasik et al., 2003; Lee & Chen, 2005; Abdou & Pointopn, 2011). Whilst some determinants are common between different models, there is no universally accepted set of determinants and therefore no optimal credit scoring model procedure

which can be applied to different banks in different countries (see for example, Abdou & Pointon, 2011).

Various statistical techniques such as discriminant analysis, logistic regression and neural networks are widely used in building scoring models (Abdou and Pointon, 2011). Altman (1968) developed the first operational model which predicted corporate bankruptcy. According to Lawson (1980) discriminant analysis is designed to look for social and economic differences between two sets of segments which can be used to classify applications as being either accepted or rejected. Chijorija (2010) stated that if the Tanzanian banks had used multiple discriminant analysis to classify and evaluate customers then the probability of failure could have been known two years prior to actual failure. Therefore, the misclassification costs could have been calculated and consequently reduced the credit risk exposure of the banks. Zuccaro (2010) also used discriminant analysis to evaluate 323 observations of which 168 individuals were classified as good credit and 155 individuals were classified as bad credit. They concluded that the most significant variable identified by discriminant analysis model was salary payment. Abdou and Pointon (2011) suggest that multiple discriminant analysis has overall better classification ability in comparison to its predictive ability. This is in contrast to the logistic regression model which has superior predictive ability. On the other hand, Coats and Fant (1993) reported that neural networks are more accurate than discriminant analysis in terms of predicting firms in financial distress. Mittal et al (2011) used multilayer perceptron neural network to classify credit risk in three categories namely bad risk, foreclosed risk and good risk in micro finance enterprises in India. Abdou and Pointon (2011) stated that neural networks have a better representation of data in comparison to both discriminant analysis and logistic regression.

This paper uses a binary variable of accepted and rejected applicants as a dependent variable in building the proposed scoring models. Where an application is accepted, the data-set suggests that all loans were repaid to the finance house. Where an application is rejected, analysis concentrated on opportunity cost i.e. determining how many clients would have fully repaid their loan(s). This distinguishes this paper from previous studies as it focuses on using credit scoring models in Islamic finance; investigates whether credit scoring can improve the decision-making process; and enhance the currently used scheme in the UK finance house.

3. Research Methodology

The data-set used in this paper consists of accepted and rejected applications made to an Islamic Finance House in the UK. Three statistical models used to analyze this data-set namely, Discriminant Analysis (DA), Logistic Regression (LR) and Multi-layer Perceptron (MP) neural network.

3.1. Statistical modeling techniques

3.1.1. Discriminant analysis

The DA model aims to identify which variables are the best predictors for accepting or rejecting new applicants. It vastly improves an administrator's decision making process as it takes into account and evaluates the pre-historic data which has already been established. This would be virtually impossible for a human being to manually do as it would be too complex and extremely time consuming. The formula for the discriminant analysis is as follows:

$$Z = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots.. + \beta_n X_n$$

where,

Z symbolizes the discriminant Z-score, α is the intercept term, $\beta_1 \ldots \beta_n$ stand for coefficients in the linear combination of the explanatory variables for X_i for i = 1 … *n*.

3.1.2. Logistic regression

LR uses categorical and numerical variables to predict if a certain event may occur. Where the logistic regression model is used to determine whether or not to accept someone for credit the outcome is that they are either accepted or declined. Consequently, the LR model, which is a popular statistical technique in credit scoring models, aims to calculate the probability of an individual falling into one of the categories based on explanatory variables. The formula for logistic regression is as follows:

$$\text{Log}\,[p/(1-p)] = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots.. + \beta_n X_n$$

where,

p represents the probability of the outcome, α is the intercept term, $\beta_1 \ldots \beta_n$ correspond with the coefficients in the linear combination of explanatory variables, X_i, for i = 1… n. The dependant variable is of the logarithm with odds {1n [*p* / (1-*p*)]}.

3.1.3 Multi-layer Perceptron neural network

A MP is a neural network whereby there is one or more layers between the input and output layers. The model consists of layers of nodes that are connected with one another, as shown in Figure 1, which enables it to be used where there are complex relationships between variables.

FIGURE 1 HERE

The input layer of the MP model has a vector of predictor variables such as ($x_1 \ldots x_n$). The input layer distributes each of these values of the neuron to the hidden layer where they are multiplied by a weight ($w_1 \ldots w_n$), and the resulting weighted value are added together generating a combined value ($C_1 \ldots C_n$). Consequently this value is fed into a transfer function (σ), which output a value ($O_1 \ldots O_n$). The output of this hidden layer is then disseminated to the output layer. Each neuron is then multiplied by weight ($W_{i1} \ldots W_{in}$), and the resulting values are added together creating a combined value ($V_1 \ldots V_n$) and is again fed into a transfer function (σ) that produces a value (Y i.e. $Y_1 \ldots Y_n$) which is the output of this model (see for example, Abdou & Pointon, 2011; Bishop, 2005; Trippi & Turban, 1993).

3.2. Data collection and sampling

The data-set is provided by one of the UK Islamic Finance Houses. It comprises 487 cases of which 336 are accepted applicants and 151 are rejected applicants. This data-set is linked to 15 independent predictor variables, in addition to the dependent binary variable, which is explained by two values, 1 for accepted applicants and 0 for rejected applicants, as shown in Table 1.

TABLE 1 HERE

Because of the high correlation between some of the variables, and due to their potential importance for the models, an Orthogonalization test is carried out between them. After running the test, the revised correlations are all within an acceptable range, as shown in Table 2.

TABLE 2 HERE

In building various scoring models under DA, LR and MP, the data-set is divided into two sub-sets. The training sub-set which is used in building the proposed scoring models consists of 341 cases (representing 70% of the overall data-set) and the testing sub-set which is used to test the predictive capabilities of the fitted models consists of 146 cases (representing 30% of the overall data-set). For MP the model is repeated with different training sub-sets and testing sub-sets. The reason for building this model is to investigate whether different results in terms of prediction accuracy is being achieved due to the random selection of both training and testing sub-sets, as part of the software design.

4. Results

In building the proposed scoring models we use DTREG software[1]. These models consist of the DA, LR and MLF. The dependant variable in all the models is a categorical variable whereby 1 = accepted and 0 = rejected. The detailed scoring models results for the above three modeling techniques are summarized below.

4.1. Modeling techniques results

4.1.1 Discriminant analysis

DA model is designed to develop a discrimination function which can be used to predict the dependent variable. In building the model 15 predicted variables are entered. The overall model is statistically significant at the 99% confidence level with *P*-value of 0.000. Our DA classification results based on 0.50 cut-off score show that the Correct Classification (CC) rates for the training sub-set is 78% and 75.4% for the testing sub-set, as shown in Table 3.

4.1.2 Logistic regression

The LR model classification results using the 15 predictor variables and 0.50 cut-off score are shown in Table 3. The *P*-value for the overall model is less than 0.01 which indicate a relationship between the predictor variables at the 99% confidence level. The CC rate for the training sub-set is 83.6% while it is 72.6% for the testing sub-set. It is clear that LR classification results are higher than the DA model in the training sub-set, but is lower than those under the testing sub-set.

[1] In building our scoring models we also used SPSS 20 and similar results were obtained.

TABLE 3 HERE

4.1.3 Multi-layer Perceptron neural network

Table 3 shows the MP[2] model classification results based on 15 predictor variables. The overall CC rate for the training sub-set is 80.4% and for the testing sub-set is 76.7%. The MP model classification results are better than the other two models namely DA and LR in terms of prediction capabilities i.e. 76.7% for MP compared with 72.6% and 75.4% for LR and DA respectively. However, this is not the case when considereing the training sub-set as the CC rate of 80.4% for the MP model is lower than the CC rate of 83.6% for the LR model but higher than the CC rate of 78% for the DA model. When the MP model is rerun based on the random selection of the training and the testing sub-sets, namely MP_r, the classification results are superior to all other models. As shown in Table 3, under the testing sub-set the CC rate is 83.5% compasred with 75.4%, 72.6% and 76.7% for DA, LR and MP models respectively.

4.2. Comparison of results of various scoring models

Two criteria namely Correct Classification (CC) Rates and Misclassification Costs (MC) are used to compare different scoring models' results. The CC rate is crucial in determining the classification efficiency of the scoring models and the MC is important in terms of estimating the costs of misclassifying a client as being rejected (Type I error) or as being accepted (Type II error). The formula used in computing the EMC is as follows:

$$\text{MC} = C\,(\text{R/A})\,P\,(\text{R/A})\,\pi_2 + C\,(\text{A/R})\,P\,(\text{A/R})\,\pi_1$$

where,

C (predicted rejected/actually accepted) and *C* (predicted accepted/actually rejected) are both corresponding MCs of both Type I and Type II errors. *P* (predicted rejected/actually accepted) and *P* (predicted accepted/actually rejected) measure the probabilities of Type I and Type II errors. π_2 and π_1 are the prior probabilities of rejected and accepted.

[2]It should be emphasized that other neural network models, namely radial basis, were also examined. The classification results for the original training and testing sub-sets were 78.6% and 76% respectively. When the random selection procedures were applied as part of the software design the classification results of the radial basis were 83.2% and 79.1% for the training and testing sub-sets respectively.

It is a challenging task to have actual misclassification costs specific to the UK. Consequently, the ratio of MCs associated with Type II and Type I errors used in this paper is 5:1, as noted by Hofmann, who used German credit data-set in his research (West, 2000). This is justified given the similarities between the developments of the two countries.

TABLE 4 HERE

Tables 4 summarize the training and testing results of the CC rates, errors and the MCs for conventional techniques namely discriminant analysis and logistic regression and for the advanced technique namely multi-layer perceptron neural network. It can be noted that the DA scoring model has the lowest Type I error of 0.245, whilst the LR model has the lowest Type II error of 0.079 under the training sub-set. This shows that conventional techniques in our sample are better than sophisticated techniques in building the scoring models. This also supported by the fact that the LR model has the lowest MC of 0.373 under the training sub-set. On the other hand, under the testing sub-set the lowest Type I and Type II errors of 0.333 and 0.084 respectively are associated with the MP_r scoring model. This suggests that the MP_r model is superior to all other scoring models used in this paper in terms of testing the prediction accuracy of the fitted model which is useful for credit officers in assessing new applications. This is also supported by the fact that the MP_r scoring model has the lowest MC of 0.361 under the testing sub-set. Subsequently, decision makers in the Islamic finance field can successfully predict a new applicant more accurately i.e. in the original sample 31% (151, rejected applicants/487, total number of applicants) are rejected whereas only 16.5% (100 – 83.5% CC rate for MP_r model under testing sub-set) should be rejected. This means that there is an improvement of 14.5% (31% - 16.5%) compared to the current judgmental system used by the UK Islamic Finance House.

Whilst there are fundamental differences between Islamic finance and their conventional counterparts which may imply differences in expected results, our actual results reflect similar conclusions from conventional literature i.e. advanced techniques such as MP are better than conventional techniques such as LR in predicting new client's credit behaviour. In addition, conventional techniques (e.g. logistic regression) have higher classification rates than advanced techniques when building scoring models (see for example, Dimla & Lister, 2000; Chen & Lee, 2005; Chijoriga, 2010; Abdou & Pointon, 2011). Applying credit scoring modeling techniques

within Islamic finance can save time and effort; cut costs and improve the quality of financial services offered within this sector. Therefore our results can be used in practice as this can help increase the profitability of the finance house by considering the 14.5% rejected applicants as accepted applicants. This can enable their clients e.g. retail business to expand which will reflect in the economic cycle of the wider society.

4.3. Importance of dependent variables within multi-layer perceptron neural network

As shown in Table 5, the most important three predictor variables for the multilayer perceptron (MP) scoring model are MONEXP, AGE and MOW with contribution weightings of 0.258, 0.125 and 0.118 respectively. By contrast, the least important three predictor variables for this model are HSTA, MOINC and TERM with contribution weightings of 0.011, 0.022 and 0.028 respectively. For the MP_r scoring model, the key predictor variables are MONEXP, AGE and MSTA with contribution weightings of 0.235, 0.126 and .087 respectively. On the other hand, the least influential predictor variables for this model are MOINC, HSTA and TERM with contribution weightings of 0.016, 0.021 and 0.027 respectively, as shown in Table 5.

TABLE 5 HERE

Finally, it can be observed that the two neural network models treat the variables slightly different as they respectively attribute to them different levels of importance. Aggregating the ranking of the contribution weights of the two models allows us to establish the three most importantly ranked variables, as follows: MONEXP, AGE and MSTA. By contrast, the least important variables for these two modelling techniques are as follows: MOINC, HSTA and TERM. Thus we submit a case for the UK Islamic Finance House to pay more attention to the variables which we find to be important, even while they are not yet using scoring models. It is expected that, if implemented, credit scoring models in combination with the current approaches could help them to provide credit not only at lower cost to themselves but also more expeditiously accurate predictions. Whilst these findings have direct implications and benefits for the Islamic finance houses, they also present a number of benefits to the wider society. For example, they can increase the number of loans offered; where clients are businesses they can enable them to expand their operations and thus improve profitability; and they can improve the perception of the credibility of Islamic finance industry by those who had not previously considered them as a source of finance.

5. Conclusion

Developing and applying credit scoring models has become one of the main challenges for financial institutions with increasing emphasis being placed on whether a client is creditworthy. This paper presents the first application of credit scoring modeling techniques in assessing credit applications in Islamic Finance. It also highlights the effectiveness of using such scoring models as opposed to the judgmental approaches currently used by the UK Islamic Finance House.

Our investigation shows that multi-layer perceptron neural network models are superior to conventional techniques used in this paper. Multi-layer perceptron models namely MP and MP_r have the highest correct classification rates in the testing sub-set compared to other modeling techniques and they also have the lowest misclassification cost. These models suggest that the three most important predictors are MONEXP, AGE and MSTA and the least important predictors are MOINC, HSTA and TERM. However, there is a role for conventional technique namely logistic regression in terms of the lowest MC in the training sub-set. This provides an answer to the main concern in this paper namely that credit scoring can work for Islamic finance and specifically by applying a neural network approach. These findings have implications for those studying credit scoring applications and those who are responsible for establishing regulations under which conventional and Islamic Finance operate as they identify a better credit evaluation process.

Our results highlight specific changes in practice that can be adopted by those using credit scoring models in both conventional and Islamic financial institutions. For example, by using scoring models to predict the new applicants' creditworthiness, a further 14.50% of rejected applications could have been accepted by the finance house. This has implications for individuals, regulators and wider society in that making loans to those who are able to repay them have positive implications for the economic cycle; they stabilize the value of financial assets by reducing the challenge of fluctuating value of assets which may swing widely at times; help retailers to expand their business activities which can have potential economic gains for individuals and society who will successfully employ finance resource; and improve public perception of products and services offered by conventional and Islamic banks in the UK. A requirement and additional benefit is the use of more transparent information by the credit bureau. Clearly the use of credit scoring models can help Islamic Finance industry to further enhance their current operations, products and services and the way they are perceived by the public. Further research would aim to investigate the main reasons behind late payments in order to gain a better understanding of the characteristics of the current

clients. Furthermore, research could be undertaking using a larger data-set and by applying other statistical modeling techniques to other Islamic financial institutions.

References

Abdou, H. and Pointon, J. (2011), "Credit scoring, statistical techniques and evaluation criteria: A review of the literature", *Intelligent Systems in Accounting, Finance and Management*, Vol. 18, No. 2-3, pp. 59-88.

Al Amari, A. (2002), *The credit evaluation process and the role of credit scoring: a case study of Qatar*. PhD thesis, University College Dublin.

Altman, E.I. (1968), "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy", *The Journal of Finance*, Vol. XXIII, No. 4, pp. 589–609.

Altman, E.I. and Narayman, P. (1997), "An International survey of business failure classification models", *Financial markets, institutions and instruments*, Vol. 6, No.2, pp. 1-57.

Anderson, R. (2007), *The Credit Scoring Toolkit: Theory and Practice for Retail Credit Risk Management and Decision Automation*, Oxford University Press: New York.

Banasik, J., Crook, J. and Thomas, L. (2003), "Sample selection bias in credit scoring models", *Journal of the Operational Research Society*, Vol. 54, No. 8, pp. 822–832.

Ben-David, A. and Frank, E. (2009), "Accuracy of machine learning models versus 'hand crafted' expert systems – a credit scoring case study", *Expert Systems with Applications*, Vol. 36, No. 3, pp. 5264–5271.

Bishop, C.M. (2005), *Neural networks for pattern recognition*, Oxford University Press, New York, NY.

Chiang, W.K., Zhang, D. and Zhou, L. (2006), "Predicting and explaining patronage behavior toward web and traditional stores using neural networks: a comparative analysis with logistic regression", *Decision Support Systems*, Vol. 41, No. 2, pp. 514–531.

Coats, P.K. and Fant, F.L. (1993), R*ecognizing financial distress patterns using a neural network,* sage, London.

Chijoriga, M. M. (2010), "Application of multiple discriminant analysis (MDA) as a credit scoring and risk assessment model", *Faculty of commerce and management,* United republic of Tanzania.

Crook, J., Edelman, D. and Thomas, L. (2007), "Recent developments in consumer credit risk assessment", *European Journal of Operational Research*, Vol. 183, No. 3, pp. 1447–1465.

Dimla, D.E. and Lister, P.M. (2000), "On-line metal cutting tool condition monitoring: II: tool state classification using multi-layer perceptron neural networks", *International journal of machine tools & manufacturer*, Vol. 40, No. 5, pp. 769-81.

Erbas, B. and Stefanou, S. (2008), "An application of neural networks in microeconomics: input-output mapping in a power generation subsector of the US electrical industry", *Expert Systems with Applications*, Vol. 36, No. 2, pp. 2317-26.

Etemadi, H., Rostamy, A. and Dehkordi, H. (2009), "A genetic programming model for bankruptcy prediction: empirical evidence from Iran", *Expert Systems with Applications*, Vol. 36, No. (2/2), pp. 3199–3207.

Hand, J. and Jacka, D. (1998), *Statistics in Finance*, Arnold: London

Huang, J., Tzeng, G. and Ong, C. (2006), "Two-stage genetic programming (2SGP) for the credit scoring model", *Applied Mathematics and Computation*, Vol. 174, No. 2, pp. 1039–1053.

Huang, C., Chen, M. and Wang, C. (2007), "Credit scoring with a data mining approach based on support vector machines", *Expert Systems with Applications*, Vol. 33, No. 4, pp. 847–856.

Kumar, A., Rao, V.R. and Soni, H. (1995), "An empirical comparison of neural network and logistic regression models", *Marketing Letters*, Vol. 6, No. 4, pp. 251–263.

Laha, A. (2007), "Building contextual classifiers by integrating fuzzy rule based classification technique and k-nn method for credit scoring", *Advanced Engineering Informatics*, Vol. 21, No. 3, pp. 281–291.

Landajo, M., Andres, J.D. and Lorca, P. (2007), "Robust neural modelling for the cross-sectional analysis of accounting information", *European Journal of Operational Research*, Vol. 177, No. 2, pp. 1232–1252.

Lawson, R. (1980), "Discriminant analysis – an aid to market segmentation", *European of Journal marketing*, Vol. 14, No. 7, pp.422 – 435.

Lee, T., Chen, I. (2005), "A two-stage hybrid credit scoring model using artificial neural networks and multivariate adaptive regression splines", *Expert Systems with Applications*, Vol. 28, No. 4, pp. 743–752.

Lee, T., Chiu, C., Lu, C. and Chen, I. (2002), "Credit scoring using the hybrid neural discriminant technique", *Expert Systems with Applications*, Vol. 23, No. 3, pp. 245–254.

Mittal, S., Gupta, P. and Jain, K. (2011), "Neural network credit scoring model for micro enterprise financing in India", *Qualitative Research in Financial Markets*, Vol. 3, No. 3, pp. 224 – 242.

Nanni, L. and Lumini, A. (2009), "An experimental comparison of ensemble of classifiers for bankruptcy prediction and credit scoring", *Expert Systems with Applications*, Vol. 36, No. (2/2), pp. 3028–3033.

Nikolopoulos, K., Goodwin, P., Patelis, A, and Assimakopoulos, V. (2007), "Forecasting with cue information: a comparison of multiple regression with alternative forecasting approaches", *European Journal of Operational Research,* Vol. 180, No. 1, pp. 354–368.

Ong, C., Huang, J. and Tzeng, G. (2005), "Building credit scoring models using genetic programming", *Expert Systems with Applications*, Vol. 29, No. 1, pp. 41–47.

Pendharkar, P.C. (2005), "A threshold-varying artificial neural network approach for classification and its application to bankruptcy prediction problem", *Computers and Operations Research*, Vol. 32, No. 10, pp. 2561–2582.

Sarlija, N., Bensic, M. and Bohacek, Z. (2004), "Multinomial model in consumer credit scoring" *In 10th International Conference on Operational Research,* Trogir, Croatia.

Sustersic, M, Mramor, D. and Zupan, J. (2009), "Consumer credit scoring models with limited data", *Expert Systems with Applications*, Vol. 36, No. 3, pp. 4736–4744.

Thieme, R.J., Song, M. and Calantone, R.J. (2000), "Artificial neural network decision support systems for new product development project selection", *Journal of Marketing Research*, Vol. 37, No. 4, pp. 499–507.

Trippi, R.R. and Turban, E. (1993), *Neural Networks in Finance and Investing: Using Artificial Intelligence to Improve Real-World Performance*. Irwin: Chicago, IL.

Tsai, C. and Wu, J. (2008), "Using neural networks ensembles for bankruptcy prediction and credit scoring", *Expert Systems with Applications*, Vol. 34, No. 4, pp. 2639–2649.

Usha, A.K. (2005), "Comparison of neural networks and regression analysis: a new insight", *Expert Systems with Applications*, Vol. 29, No. 2, pp. 424–430.

Walczak, S. and Sincich, T. (1999), "A comparative analysis of regression and neural networks for university admissions", *Information Sciences*, Vol. 119, No. (1–2), pp. 1–20.

West, D. (2000), "neural network credit scoring models", *Computers & operations research,* Vol. 27, No. 11/12, pp 1131-52.

Zhang, G., Hu, M.Y., Patuwo, B.E. and Indro, D.C. (1999), "Artificial neural networks in bankruptcy prediction: general framework and cross validation analysis", *European Journal of operational research*, Vol. 116 No. 1, pp.16-32.

Zuccaro, C. (2010), "Classification and prediction in customer scoring", University of Quebec in Monteral, Canada.

Tables

Table 1: List of predictor variables used in building the scoring models

Variable	Code	Description
Age	AGE	Actual age
Gender	GDR	Male = 0 or Female = 1
Marital status	MSTA	Single = 0, Married = 1 or divorced = 3
Number of dependants	NDEP	Actual number of dependants
Housing status	HSTA	Home ownership = 0 or rented = 1
Own a car	CRWN	Yes = 0 or No = 1
Occupation	OCCUP	Call centre = 1 council = 2 office/admin = 3 retail = 4 sales advisor = 5 self-employed = 6 student = 7 taxi driver = 8
Mode of work	MOW	Full-time = 0 or part-time = 1
Mode of income	MOINC	Fixed income = 0 or variable income = 1
Type of credit	TCRED	None = 0 credit card =1, family/friends =2
Overdraft	OVDFT	Yes = 0 or No =1
Monthly income	MONINC	Actual monthly income
Monthly expense	MONEXP	Actual monthly expenses
Loan amount	LMOT	Actual loan amount
Loan duration	TERM	Maximum of two years

Table 2: Orthognalisation test results for highly correlated predictor variables

High correlation variables			Correlation before the Orthogonalization test	Correlation after the Orthogonalization test
Monthly expenses	versus	Monthly income	95.4 %	0.00 %
Monthly expenses	versus	Mode of work	82.9 %	22.9 %
Mode of work	versus	Monthly income	82.8 %	0.00 %
Mode of work	versus	Type of credit	82.2 %	0.02 %

Table 3: Classification results for the DA, LR and MP scoring models

Scoring Model	Training sub-set				Testing sub-set			
	A	R	Total	%	A	R	Total	%
DA								
A	78	24	102	76.5	28	21	49	57.1
R	60	179	239	74.9	24	73	97	75.3
Total			341	78.0			146	75.4
LR								
A	65	37	102	63.7	23	26	49	46.9
R	19	220	239	92.1	14	83	97	85.6
Total			341	83.6			146	72.6
MP								
A	67	35	102	65.7	25	24	49	51
R	32	207	239	86.6	10	87	97	89.7
Total			341	80.4			146	76.7
MP_r								
A	63	37	100	63.0	34	17	51	66.7
R	22	207	229	90.4	9	98	107	91.6
Total			339	82.1			158	83.5

Note: A refers to Accepted; R refers to Rejected; DA refers to Discriminant Analysis; LR refers to Logistic Regression; MP refers to Multi-layer Perceptron; MP_r refers to Multi-layer Perceptron based on random selection of training and testing sub-sets.

Table 4: Classification results, errors and MCs for different scoring models

Scoring model	ACC rates				Errors				Misclassification costs	
	Training sub-set		Testing sub-set		Training sub-set		Testing sub-set		Training sub-set	Testing sub-set
	A	R	A	R	Type I	Type II	Type I	Type II	MC (5:1)	MC (5:1)
DA	75.5	74.9	57.1	75.3	0.245	0.251	0.429	0.247	0.574	0.692
LR	63.7	92.1	46.9	85.6	0.363	0.079	0.531	0.144	0.373	0.593
MP	65.7	86.6	51	89.7	0.343	0.134	0.490	0.103	0.449	0.498
MP_r	63	90.4	66.7	91.6	0.370	0.096	0.333	0.084	0.405	0.361

Note: A refers to Accepted; R refers to Rejected and MC refers to Misclassification Cost.

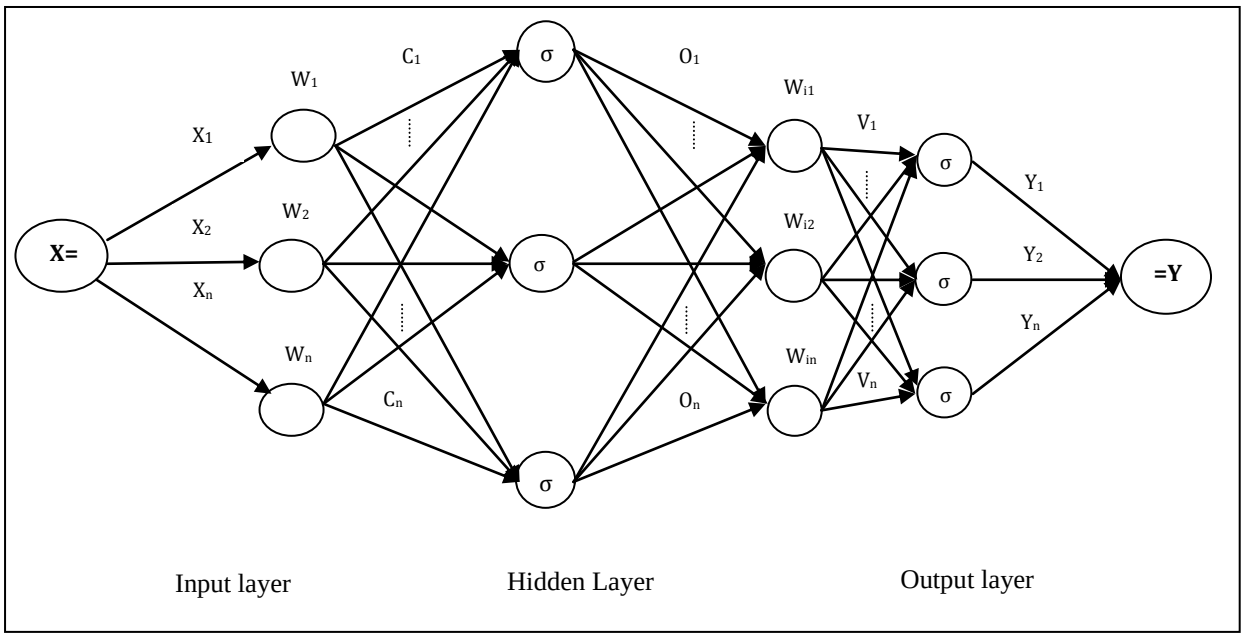
Table 5: Independent variables importance

Variable	Importance of dependent variables	
	MP	MP_r
MONEXP	0.258	0.235
AGE	0.125	0.126
MSTA	0.085	0.087
CRWN	0.041	0.084
OVDFT	0.073	0.073
MONINC	0.034	0.063
MOW	0.118	0.061
NDEP	0.038	0.053
TCRED	0.037	0.051
GDR	0.060	0.047
LMOT	0.039	0.030
OCCUP	0.030	0.027
TERM	0.028	0.026
HSTA	0.012	0.021
MOINC	0.022	0.016
$\sum$	100	100

Note: MP refers to Multi-layer Perceptron.

Figure

Figure 1: Architecture of Multi-Layer Perceptron Neural Network



Source: own figure