Available online at www.sciencedirect.com
ScienceDirect
Review of Development Finance 4 (2014) 20–28
Review of DEVELOPMENT FINANCE
www.elsevier.com/locate/rdf

Credit risk assessment model for Jordanian commercial banks: Neural scoring approach

Hussain Ali Bekhet*, Shorouq Fathi Kamel Eletter

Graduate Business School, College of Graduate Studies, Universiti Tenaga Nasional (UNITEN), 43000 Kajang, Selangor, Malaysia

Abstract

Despite the increase in the number of non-performing loans and competition in the banking market, most of the Jordanian commercial banks are reluctant to use data mining tools to support credit decisions. Artificial neural networks represent a new family of statistical techniques and promising data mining tools that have been used successfully in classification problems in many domains. This paper proposes two credit scoring models using data mining techniques to support loan decisions for the Jordanian commercial banks. Loan application evaluation would improve credit decision effectiveness and control loan office tasks, as well as save analysis time and cost. Both accepted and rejected loan applications, from different Jordanian commercial banks, were used to build the credit scoring models. The results indicate that the logistic regression model performed slightly better than the radial basis function model in terms of the overall accuracy rate. However, the radial basis function was superior in identifying those customers who may default.

JEL classification: C45; C52; C83

Keywords: Artificial neural network; Credit scoring; Logistic regression; Credit risk; Commercial bank; Jordan

1. Introduction

Credit loans constitute a cornerstone of the banking industry. The performance of credit department in good standing guarantees profitability and stability of a bank. Therefore, screening the customer's financial history and financial background is a very significant factor before making any credit decision and is a key determinant in reducing credit risk (Bekhet and Eletter, 2012).

Credit risk is the most critical and the biggest challenge facing banks' management. In fact, risk estimate is a major factor contributing to any credit decision, and the inability to precisely determine risk adversely affects credit management. In addition, risk affects both approved and unapproved financing decisions. When credit manager approves a loan, he/she risks the possibility that the customer may be unable to repay his/her obligation. Conversely, when loan is rejected, there is a risk of losing a potentially profitable customer to competitors and the risk of opportunity cost. Hence, credit risk evaluation is essential before making any lending decision.

Lahsasna et al. (2010) emphasized that credit risk decisions are key determinants for the success of financial institutions because of huge losses that result from wrong decisions. Poor evaluation of credit risk can cause money loss (Gouvea, 2007). Wu et al. (2010) stressed that credit risk assessment is the basis of credit risk management in commercial banks and provides the basis for loan decision-making. Furthermore, Angelini et al. (2008) stressed that risk continues to provide a major threat to successful lending despite advancements in credit evaluation techniques and portfolio diversification. Due to the significance of credit risk, a number of studies have proposed embracing data mining tools in banks to improve their risk assessment models and hence increase the prediction accuracy of existing models (Akkoc, 2012; Chen and Huang, 2003; Gao et al., 2006; Huang et al., 2006; Malhorta and Malhorta, 2003; Martens et al., 2007; Tsai and Wu, 2008; West, 2000). Artificial neural networks, genetic algorithms, genetic programming, support vector machines, and some hybrid models have been used to evaluate credit risk with promising results in terms of performance accuracy.

Commercial banks in Jordan are regarded as vitally important and competitive financial organizations that seek profit by

* Corresponding author. Tel.: +603 8928 7326.
E-mail addresses: profhussain@uniten.edu.my, drbekhet1953@hotmail.com (H.A. Bekhet).

Peer review under responsibility of Africagrowth Institute.

ELSEVIER Production and hosting by Elsevier

1879-9337 © 2014 Africagrowth Institute. Production and hosting by Elsevier B.V. Open access under CC BY-NC-ND license.
http://dx.doi.org/10.1016/j.rdf.2014.03.002

providing various financial services to households and business firms while managing different types of risk. Consequently, risk taking is often viewed as the basic driver for financial behavior and profitability (Bekhet and Eletter, 2012). In addition, the banking sector contributes significantly to the Jordanian GDP on average by 20% of the country's GDP during 2000–2010 (CBJ, 2002, 2006, 2010; Bekhet and Eletter, 2012; Khrawish, 2011). However, the ratio of non-performing loans to total loans has rapidly increased and reached 7.9% in 2010 compared to 4% in 2007 (Association of Banks in Jordan, 2010). This reflects a slowdown in the country's economic activities as a consequence of the global financial crisis. The rise of the ratio for non-performing loans indicates that some clients are trying to delay paying their financial obligations to banks. Such a situation will increase credit risk and might cause local financial crisis. In fact, bad loans and foreign currency speculations had led to the bankruptcy of Petra Bank (the third largest Jordanian bank) in 1989 and the bankruptcy of Trade and Credit Bank in 1991 (CBJ, 1988, 1989, 1990, 2010, 2011).

As a matter of fact, loan application evaluation at the Jordanian banks is subjective in nature. This entails reviewing each loan application manually, which imposes biases including personal insights, knowledge, and intuition of the credit manager. This method nevertheless has been replaced in a few banks by credit scoring models or a combination of objective and subjective reviews to make proper credit decisions. In fact, subjective decision-making in lending institutions might cause financial crisis or distress. Simultaneously, credit risk assessment is significant in reducing manual errors in credit decisions. On the other hand, banks store data about their customers in data warehouses which can be viewed as hidden knowledge assets that can be accessed and utilized through data mining tools. However, despite the increase in the number of non-performing loans and competition in the banking market, most of the Jordanian commercial banks are unwilling to use data mining tools to support credit decisions. Nevertheless, credit managers at Jordanian banks need to develop more effective models to improve the classification accuracy of credit risk decisions, and recently, artificial neural networks (ANNs) represent a statistical technique and an auspicious data mining tool that have been used successfully in classification problems in many domains.

Therefore, the aim of the current study is to explore the effectiveness of two credit scoring models in the Jordanian commercial banks. Radial basis function (RBF) and logistic regression model in evaluating credit applications were used. The paper also aims to investigate the superiority of the RBF model over logistic regression in screening out potential defaulters. Furthermore, using data mining techniques in application evaluation would improve credit decision effectiveness and control loan officer tasks, as well as save analysis time and cost.

The structure of the paper is as follows: Section 2 discusses the credit scoring models. Section 3 introduces the literature review. Section 4 defines the data sources and definitions of variables. Section 5 describes the methodology and models while Section 6 presents the experimental results and relevant discussion. Finally, Section 7 provides the conclusions and recommendations for further research.

2. Credit scoring

Credit scoring is a group of decision models and their underlying techniques which give support to lenders when providing credit to customers (Heiat, 2012; Thomas et al., 2002). In addition, credit scoring model is a decision support system that helps the managers in financial decision-making process. Chen and Huang (2003) stated that with the rapid development in credit industry, credit scoring models are used on decisions related to credit admission evaluation. These models are developed to classify credit applications as "accepted" or "rejected" with respect to applicants' characteristics such as age, income, and marital status. An application is accepted or rejected based on expectation that the applicant is able or not able to repay his financial obligation. Besides, they emphasized that creditors can build classification rules using previous accepted and rejected applications. Furthermore, they are used to predict borrower's credit risk (Thomas, 2000; Yap et al., 2011). The objective of the credit scoring model is to determine credit applicant's capacity to repay financial obligations by evaluating the credit risk of loan application (Emel et al., 2003; Lee et al., 2002). Credit scoring is a system that aims to classify loan applications; those that have high probability of fulfilling financial obligations are classified as "good" and those that have low probability of fulfilling financial obligations are classified as "bad" (Akkoc, 2012; Gao et al., 2006; Lahsasna et al., 2010; Lee et al., 2002; West, 2000). In addition, Khashman (2010) pointed out that application scoring is one of the two credit scoring tasks which use financial and demographic information of credit applicant in order to classify loan application into "good" or "bad" risk groups.

However, it is necessary to rely on models and algorithms rather than human judgment in consumer lending because of the vast number of decisions involved (Khandani et al., 2010). This highlights the need for accurate decision support model for credit admission evaluation and also for monitoring the ongoing health of credit customers (West et al., 2005). A small improvement in the accuracy of the credit decision might reduce credit risk and translate into important future savings (Chen and Huang, 2003; Hand and Henley, 1997; West, 2000; West et al., 2005; Tsai and Wu, 2008; Lahsasna et al., 2010). In the light of that, credit scoring has been studied widely in accounting and finance literature because of its impacts in lending decisions and profitability of financial institutions (Tsai and Wu, 2008).

Usually, a credit scoring model is built using statistical techniques such as linear discriminant analysis (LDA) and logistic regression (LR). However, artificial neural networks (ANNs) are introduced as promising data mining tools that provide an alternative to statistical techniques in building credit scoring models. Furthermore, artificial neural networks have recently been used successfully in different business applications (Akkoc, 2012; Chen and Huang, 2003; Eletter, 2012; Gao et al., 2006; Huang et al., 2004a; Khashman, 2010; Malhorta and Malhorta, 2003; Martens et al., 2007; Tsai and Wu, 2008; West, 2000).

3. Literature review

Many studies have been emphasizing on the use of data mining techniques such as statistical techniques, artificial neural networks, and many others in business applications. The current paper will survey the latest studies to highlight the importance of utilizing data mining techniques to build credit scoring models that support credit decision in the Jordanian commercial banks.

Artificial neural networks (ANNs) have been used in many business applications in problems such as classification, pattern recognition, forecasting, optimization, and clustering. ANNs are distributed information-processing systems composed of many simple interconnected nodes inspired biologically by the human brain (Eletter, 2012). Paliwal and Kumar (2009) asserted that ANNs have been applied widely in research focused on prediction and classification in a mixture of fields' applications. They viewed neural networks and traditional statistical techniques as competing model building tools. Angelini et al. (2008) pointed out that ANNs have emerged effectively in credit scoring because of their ability to model non-linear relationship between a set of inputs and a set of outputs. They viewed ANNs as black boxes because it is impossible to extort any symbolic information from their internal configurations. Khashman (2010) employed neural networks to credit risk evaluation using the German dataset. Three neural network models with nine learning schemes were developed and then the different implementation outcomes were compared. The experimental results showed that one of the learning schemes achieved high performance with an overall accuracy rate of 83.6%. Similarly, Angelini et al. (2008) developed two neural networks credit scoring models using Italian data from small businesses. The overall performance assured that neural networks can be applied successfully in credit risk assessment.

Witkowska et al. (2004) used multilayer perceptron and RBF networks to classify customers into "good" or "bad" credit risk. They stressed that ANNs are useful tools for supporting decision-making in financial institutions. Gao et al. (2006) used feed forward neural network with a structured tuning particle swarm algorithm to optimize the structure and weights for the network simultaneously. The training algorithm improved data handling efficiency and generalization ability of the neural network. The results showed that the fitting classification model has reduced the creditor's risk and thereby provides a promising alternative for credit analysis system. Additionally, Malhorta and Malhorta (2003) used a collective dataset of twelve credit unions to evaluate the ability of ANNs in classifying loan applications into "good" or "bad". The effectiveness of the ANNs model in screening loan applications was compared with multiple discriminant analysis (MDA) models. They found that neural network models outperformed the discriminant analysis model in identifying potential loan defaulters.

In their study, Jagric et al. (2011) emphasized that a bank's main challenge is to build up new credit risk models with higher predictive accuracy. They stressed on using ANNs to construct a credit scoring model because of its ability to capture non-linearity in financial data. They developed a credit decision model using learning vector quantization (LVQ) neural network for retail loans and logistic regression model for benchmarking. A real life dataset from Slovenian banks was used. The results showed that LVQ model outperformed the logistic model and achieved higher accuracy results in the validation set. Boguslauskas and Mileris (2009) further asserted that ANNs and logistic regression are the most efficient, widely used methods for credit risk measurement. They described rates of credit risk estimation models and their calculation for the analysis of Lithuanian enterprises credit risk. They stressed that neural networks models achieved higher rates of classification accuracy.

Recently, Blanco et al. (2013) used the multilayer perceptron neural network (MLP) to develop a specific microfinance credit scoring model. They compared the performance of the MLP model against three statistical techniques: linear discriminant analysis, quadratic discriminant analysis, and logistic regression. The MLP model achieved higher accuracy with lower misclassification cost. The findings confirmed the superiority of the MLP over the parametric statistical techniques.

In another study, Bensic et al. (2005) tried to characterize the main features for small business credit scoring and compared the performance using logistic regression (LR), neural network (NN), and classification and regression trees (CART) on a small dataset. The results showed that the probabilistic NN model achieved the best performance. Furthermore, the findings provided new knowledge about credit scoring modeling in a transitional country. Additionally, West (2000) examined the potential of five neural network architectures in credit scoring accuracy and benchmarked the results with traditional statistical methods: linear discriminant analysis and logistic regression, and other non-parametric methods: decision trees, kernel density estimation, and nearest neighbor. The results showed that neural networks credit models were able to improve credit scoring accuracy from 0.5 to 3%. Moreover, Koh et al. (2006) asserted that the best performing credit scoring models are obtained using logistic regression, neural network, and decision tree.

4. Data collection and variable definition

Drawn from the existing literature, we employ a pooled data of both accepted and rejected applications from different Jordanian commercial banks for the 2006–2011 period.[1] The number of observations from each bank was concealed in order to protect the confidentiality of the banks. The data content is composed of 492 cases. In the provided sample, 292 (59.3%) applications were credit worthy while 200 (40.3%) applications were not. The data collection resulted in the total of 13 variables: seven variables were scale while six variables were categorical. The definition, coding, type, and descriptive of each these variables were shown in Table 1.

[1] An alternative approach of building a credit scoring model is to use a binary classification of good vs. bad firms, where good firms are those that have survived and bad are those that went bankrupt.

Table 1
Proposed variables for building the credit scoring model.

Variable	Type	Variable definition
Age, A	Scale	Applicant's age
Gender, G	Binary	Male or female
Total income, TI	Scale	Total monthly income, and used log for transformed it.
Company's type, CT	Binary	Applicant works in a credible company or not
Guarantor, GU	Binary	Existence of alternative source of repayment if required
Loan amount, LA	Scale	Loan amount, and used log for transformed it.
Loan purpose, LP	Nominal	For different purposes: car, housing, and personal commitments, i.e. Education, marriage, etc.
Period with current employer, PE	Scale	Job experience with current employer
Duration of credit, D	Scale	Loan duration in months
Nationality, N	Binary	If the applicant is Jordanian or foreigner?
Interest rate, IR	Scale	Real interest rate
Debt payment ratio, DPR	Scale	Total debt divided by total income (i.e. DPR measures the applicant's repaying ability: high DPR ratio points to high credit risk, whereas low DPR ratio points to a good credit application).
Credit decision, CD	Binary	1 for accepted/good application & 0 for rejected/bad application.

The dependent variable was the credit decision (CD), a binary variable with two values 1 for accepted application or 0 for rejected application. The tendency to work with only two values "accepted" or "rejected" applications is noticed by many studies (Hand and Henley, 1997; Thomas et al., 2002; Lee et al., 2002; Bensic et al., 2005; West et al., 2005; Huang et al., 2006; Abdou et al., 2007; Tsai and Wu, 2008; Khashman, 2010; Akkoc, 2012; Yap et al., 2011; Blanco et al., 2013). Categorical variables were converted into numerical values in order to be utilized by neural network model. All scale variables were standardized using the rescaling of covariates option in SPSS to improve the network training. SPSS software (version 20) was used to perform the analysis for the current study.

5. Research methodology

5.1. Logistic regression model

Logistic regression (LR) is a predictive model widely used in classification. According to Thomas (2000), LR is a linear regression in which the target variable is a non-linear function of the probability of being good. In addition, he stressed that the classification results of LR model are sensitive to correlations between the independent variables. Therefore, variables used in developing the model should not be strongly correlated. Lahsasna et al. (2008) asserted that the non-linearity of the credit data decreases LR accuracy. Furthermore, Yap et al. (2011) stressed that LR credit scoring model aims to determine the conditional probability of each application belonging to one class, i.e. good or bad given the values of the explanatory variables of the credit applicant. Lee and Chen (2005) supported this view by stressing that each application will be assigned only to one class of the dependent variable.

However, the logistic regression model limits generation of the predicted values of the dependent (response) variable to lie in the interval between zero and one. Logistic regression is a common modeling technique that classifies between two groups using a set of predictor variables (Akkoc, 2012). The LR model is represented as in Eq. (1).

$$\begin{aligned}\ln(p_i/1-p_i) = {} & \beta_0+\beta_1\mathrm{A}+\beta_2\mathrm{G}+\beta_3\mathrm{TI}+\beta_4\mathrm{DPR}\\ & +\beta_5\mathrm{LA}+\beta_6\mathrm{IR}+\beta_7\mathrm{LP}+\beta_8\mathrm{PE}+\beta_9\mathrm{DM}\\ & +\beta_{10}\mathrm{GU}+\beta_{11}\mathrm{N}+\beta_{12}\mathrm{CT}+\varepsilon \qquad (1)\end{aligned}$$

p_i is the probability of being good for a particular customer, i, which is also a function of the predictor variables, X_i (age, gender, total income, DPR, loan amount, interest rate, loan purpose, period in months with current employer, duration in month, guarantor, nationality, and company's type) that represent the applicant's characteristics. β_0 is the intercept, $\beta_j = 1, \ldots, 12)$ are the coefficients associated with the corresponding predictor x_i ($i = 1, \ldots, 12$); $(\ln(p_i/1-p_i))$ represents the credit decision (CD), and ε is errors' terms. Multicollinearity is unfavorable feature of logistic regression, but it is not critical issue because the credit scoring is developed for prediction. Multicollinearity can be diagnosed using VIF (VIF = $1/(1-R^2)$). Furthermore, Gujarati (2003, p. 359) stated that correlations above 0.80 are crucial. Nevertheless, in the current research correlation between predictor variables were considered and the highest correlation was <0.6 (correlation matrix was provided in Appendix).

5.2. Radial basis function scoring model

Artificial neural networks (ANNs) are fruitful non-linear modeling tools. ANNs have a biologically inspired capability that mimics processing capabilities of the human brain (Cao and Parry, 2009). ANN is capable to learn from examples. A neural network model is composed of a number of processing units called neurons cooperating across different layers (Akkoc, 2012) and connected through several connections or weights.

The feed forward is a popular neural network architecture used in many applications. The feed forward architecture implies that neurons are organized in layers in a layered network. Simultaneously, information flows from source to destination through the net. RBF is a popular architecture of multilayer feed forward neural network which produces a predictive model for a target variable based on a set of predictor variables. The benefit of the RBF network is that it uses local approximations to find the input to output map and it needs fewer training cases as well (Huang et al., 2004b). A typical RBF network comprises of three layers: input layer, hidden layer, and output layer. The input layer feeds network inputs; the hidden layer remaps inputs to make them linearly separable; then the output layer performs linear separation (Xie et al., 2011). Bruzzone and Fernandez (1998) pointed that training RBF neural network is carried out in two steps. The connection weights between input and hidden layers are determined first. And then the connection weights between hidden and output layers are found using a supervised algorithm while trying to minimize the sum of squares error function.

Therefore the basic computation in the RBF networks is performed as follows.

After input layer feeds, the predictor variables are the hidden layer. Each hidden neuron receives a p-dimensional input vector, X, of all inputs in the input layer, then it computes the Euclidian distance between a weight vector, W, (centers of hidden neurons) and input vector, X (i.e. $d = |X - W|$). Subsequently, each hidden neuron produces $\varnothing(d)$ to the output layer where $\varnothing(\cdot)$ is a transfer (activation) function such as soft max, Gaussian, etc. (Bruzzone and Fernandez, 1998; Memarian and Balasundram, 2012). The output layer has a linear transfer which is responsible of producing the predicted value, Y. Therefore, Y is the weighted linear combination m radial basis functions and is expressed as in Eq. (2).

$$Y = \sum_{j=1}^{m} W_j \phi(d) + W_0 \tag{2}$$

where m refers to the number of hidden neurons, W_j is the connection weight between the jth hidden neuron and the output layer, and W_0 is the bias term (Xie et al., 2011) for each neuron in the output layer (Memarian and Balasundram, 2012). Finally, the output layer produces an outcome, Y. West (2000) stated that the output, Y, is calculated only from the radial basis function whose weights (centers) are close to the input vector, $\mathbf{X}$, i.e. when the distance, d, is small or close to zero. The activation function is a symmetrical function with a maximum value equals 1. Therefore, the output increases as the distance, d, decreases. Information transmitted outside the net is statistically referred to as dependent variable (Akkoc, 2012). The outcome of the output neuron is the solution of the problem. Moreover, the neural network learn the desired relationship between the independent and dependent variables using a representative training set of (input, target) pairs. The net compares the actual output, Y, with the desired output, T, and if it is not satisfied then it adjusts the connection weights during training in an iterative process until a desirable result is reached.

In the current study, the input layer has 18 neurons, which equal the number of covariates (seven in this case) plus the total number of factor levels; a separate neuron was created for each category of gender, loan purpose, nationality, *company type, and guarantor*. Likewise, the output layer has two neurons with identity activation function; a separate neuron was created for each category of the credit decision. The automatic architecture selection in the SPSS chose six neurons for the hidden layer with softmax activation function. The number of hidden neurons was determined as the best number of hidden units that minimizes the sum of squares error in the testing set.

Table 2
Change in −2 log likelihood ratio.

Variable	Change in −2 log likelihood	df	P value
Loan purpose, LP	11.81	2	0.003
Company type, CT	17.77	1	0.000
Guarantor, GU	11.22	1	0.003
Debt payment ratio, DPR	142.02	1	0.000
Duration in months, DM	5.93	1	0.005
Interest rate, IR	7.99	1	0.005
Total income, TI	5.54	1	0.019

6. Results analysis

6.1. Logistic regression credit scoring analysis

In the current study, 440 cases were used to build the logistic regression scoring model and 52 cases were used to evaluate the developed model. In addition, the forward stepwise method was used in order to extract the most influential variables for model building. Hence, variables were added to the model on sequential steps. At each step, a variable with the largest score statistic and whose significance value is less than 0.05 was added to the model. Additionally, each variable added to the model should have change in the −2 log likelihood ratio of probability less than 0.05.

Table 2 shows that only seven predictor variables out of 12 were significant and most influential to the credit decision (P value < 0.05). These variables are: loan purpose, company type, guarantor, DPR, duration in months, interest rate, and total income.

Table 3 presents the chi-square result that tests the significance of the LR model. It provides statistical evidence that there exists relationship between the selected variables and the dependent variable (credit decision, CD). It shows that the probability of the chi-square (288.280) is less than 0.05. Therefore, the null hypothesis ($\beta_i = 0$) that there is no relationship between the predicted variables (as listed in Table 2) and the dependent variable (CD) is rejected. Hence, it can be confirmed that there exists a relationship between the selected variables and the credit decision.

Table 4 reveals the Wald Statistic results that provide a statistical evidence of the presence of relationship between the CD and each predictor variable entered into the model (as shown in Table 2). The Wald Statistic test was used to examine the hypotheses [null hypothesis (H_0: $\beta_i = 0$) and the alternative hypothesis (H_1: $\beta_i \neq 0$)] for each particular variable. The findings show that all variables have a statistically significant

Table 3
Omnibus tests of model coefficients.

	Chi-square	df	Sig.
Step 7			
Step	5.538	1	0.019
Block	288.280	8	0.000
Model	288.280	8	0.000

Table 4
Variables in the model.

Variables	β	S.E.	Wald	df	Sig.	Exp. (B)
Loan purpose			9.893	2	0.007	
Loan purpose 1, LP_1	−1.89	0.759	6.178	1	0.013	0.15
Loan purpose 2, LP_2	−1.21	0.544	4.943	1	0.028	0.30
Company type, CT	−3.54	1.143	9.588	1	0.002	0.03
Guarantor, GU	−2.58	0.881	8.544	1	0.003	0.08
Debt payment ratio, DPR	−.9.68	1.028	88.732	1	0.000	0.00
Duration in months, DM	−0.01	0.004	6.082	1	0.014	0.99
Interest rate, IR	−33.66	11.907	7.99	1	0.005	0.00
Total income, TI	−1.28	0.546	5.503	1	0.019	0.28
Constant	14.47	2.563	31.872	1	0.000	1,930,182.42

relationship and impact on the credit decision (Sig. < 0.05). Additionally, the coefficients ($\beta_{i'}s$) can be used in Eq. (1) to calculate the probability of accepting or rejecting an application. However, the sign of each β_i determines the direction of the relationship between each variable and the credit decision. In fact, those variables with positive β_i increase the likelihood of a yes answer (accept) to an application, while variables with negative β_i will decrease the probability of accepting an application. This suggests that an increase in the values of DPR, DM, and IR will decrease the probability of accepting an application.

Additionally, Table 5 summarizes the classification capability of the LR model. The correct predictions are presented in diagonal cells while the off diagonal cells have the wrong predictions. It can be observed that 79.6% of the rejected applications were classified correctly, 88.4% of the accepted applications were classified correctly and overall, the correct classification rate of the LR model was 84.8% with a 0.5 cut-off point.

Furthermore, the developed model was tested using a testing subset of 52 cases (19 rejected applications and 33 accepted applications) that was not used to create the model. The overall classification rate for the testing sample was 90.4%. In fact, the LR credit scoring model performed better when classifying accepted applications (97%) than classifying bad applications (78.9%).

6.2. *RBF credit scoring function*

The RBF credit scoring model was built using the same dataset used in developing the LR model and the 12 independent variables. Table 6 shows that the dataset were partitioned into two subsets: 89.4% of the cases were used for training and 10.6% for model testing.

The network comprises of three layers: input layer with 18 neurons, hidden layer of six neurons, and output layer of two neurons. Table 7 displays the model summary in which the percentage of incorrect prediction in the training and testing sets was 19.1% and 13.5%, respectively. The sum of squares error function was used as the stopping criterion.

Table 8 presents the classification results for the RBF model. The overall classification rate in the training and the testing

Table 5
Logistic regression classification results.

Observed	Predicted					
	Model building cases			Testing Cases		
	CD		% Correct	Target		% Correct
	Rejected	Accepted		Rejected	Accepted	
CD						
Rejected	144	37	79.6	15	4	78.9
Accepted	30	229	88.4	1	32	97
Overall %			84.8			90.4

Table 6
Case processing summary.

	Frequency	%
Sample		
Training	440	89.4
Testing	52	10.6
Valid	492	100.0
Excluded	0	
Total	492	

Table 7
Model summary.

Training	
Sum of squares error	57.775
Percent incorrect predictions	19.1%
Training time	00:00:02.236
Testing	
Sum of squares error	6.621
Percent incorrect predictions	13.5%

Table 8
RBF Classification.

Sample	Observed	Predicted		
		Rejected	Accepted	% Correct
Training	Rejected	159	22	87.8
	Accepted	62	197	76.1
	Overall %	50.2	49.8	80.9
Testing	Rejected	16	3	84.2
	Accepted	4	29	87.9
	Overall %	38.5	61.5	86.5

Table 9
Independent variables importance.

Variable	Importance	Normalized importance (%)
Gender, G	4.3	28.3
Loan purpose, LP	4.5	29.6
Company's type, CT	5.2	33.7
Guarantor, GU	4.2	27.6
Nationality, N	4.6	30.0
Age, A	11.1	72.8
Debt payment ratio, DPR	15.3	100.0
Loan amount, LA	11.8	77.3
Total income, TI	10.6	69.1
Period in months with current employer, PE	9.3	61.0
Duration in months, DM	10.9	71.1
Interest rate, IR	8.3	54.3

sample was 80.9% and 86.5%, respectively. However, the RBF model correctly classified: 87.8% of rejected applications and 76.1% of the accepted applications of the training set. Also in the testing set, 84.2% of rejected applications and 87.9% of the accepted applications were classified correctly.

Table 9 shows the importance and the normalized importance of all variables in the RBF model. The importance of an independent variable measures how much the network's predicted value varies for different values of the independent variable. The results show that DPR scored the highest importance, followed by loan amount, age, duration in months, and total income. However, DPR scored 15.3%, which indicates that DPR strongly influence the predicted value of the model (credit decision). On the other hand, guarantor has the lowest importance level of 4.2% which suggests that guarantor has no influence on the RBF predicted value of the credit decision. Normalized importance in column 3 is the importance values divided by the DPR's importance value (highest importance) and displayed as a percentage.

Fig. 1 reveals the descending ranking of the importance and normalized importance values. It suggests that the variables DPR, loan amount, age, duration in months, and total income have high effect on how the network classifies credit applications, whereas guarantor has the least influence on credit decision. However, the way in which the independent variables are correlated to the predicted value of the credit decision is not obvious. Based on commonsense, one could guess that a larger amount of DPR points to a higher probability to reject the credit application being rejected.

6.3. Comparing performance of different credit scoring models

The classification accuracy rate, as well as Type I and Type II errors for the two models are reported in Table 10. In general, classification accuracy rate is the most common quantitative measure used in evaluating the predictive accuracy of classification models. In addition, it represents the percentage of applications that are classified correctly (Abdou et al., 2007). It is evident from Table 10 that the overall classification accuracy rate for the LR model is higher than the overall classification

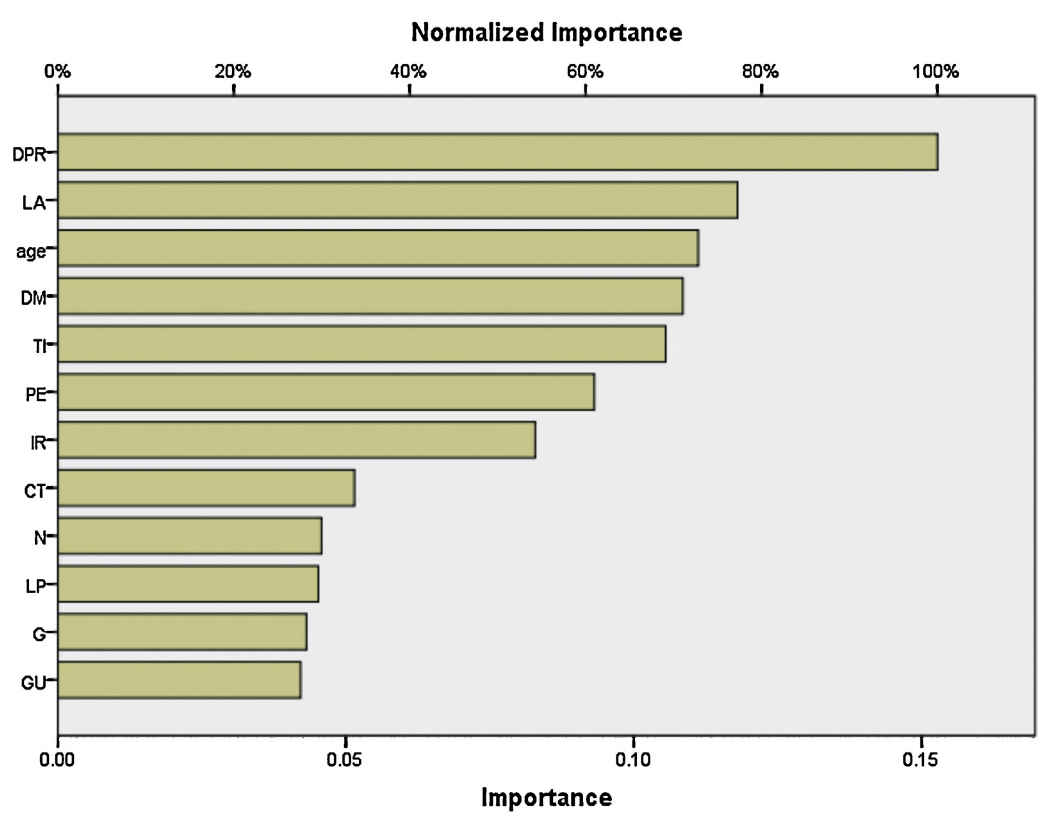

Fig. 1. Normalized importance for the independent variables.

Table 10
Classification results.

Model	Sample	Classification rate (%)	Overall %	Type I error (%)	Type II error (%)
LR	Training	84.8	85.4	11.6	20.4
	Testing	90.4		3.0	21.1
RBF	Training	80.9	81.5	23.9	12.2
	Testing	86.5		12.1	15.8

accuracy rate for the RBF (85.4% against 81.5%) with an accuracy improvement of 3.9%. However, it is interesting to note that the individual group accuracy rate of the two models vary greatly (see Tables 5 and 8). For the accepted applications, the LR is more accurate than the RBF (88.4% vs. 76.1%) for the training set and (97% vs. 87.9%) for the testing set. In contrast, when predicting rejected applications the RBF outperformed the LR (84.2% vs. 78.9%) for the training set and (87.8% vs. 79.6%) for the testing set. This indicates that the RBF is better and more powerful than the LR in screening bad applications.

Furthermore, Type I error results occur when a good application is misclassified as a bad application. In contrast, Type II error results when a bad application is incorrectly classified as good application (credit risk). Additionally, it is believed that type II error is more costly and has higher impact (Kurum et al., 2012; West, 2000; Yap et al., 2011). As reported in Table 10, the LR has lower Type I error than the RBF (11.6% vs. 23.9%) for the training set and (3% vs. 12.1%) for the testing set. Conversely, RBF Type II error is lower than the LR (12.2% vs. 20.4%) and (15.8% vs. 21.1%) for the training and testing samples, respectively.

Furthermore, the non-parametric nature of the RBF model is a key deficiency in explaining the strength and the direction of the relationship between the input variables and the credit decision, and as a consequence makes it hard to justify the credit decision.

7. Conclusions and further research

Both models (LR and RBF) have shown promising results and it can be concluded that there is no an overall best model for credit application evaluation. The LR performed better than the RBF with regard to the overall classification rate. On the other hand, the RBF model outperformed the LR model in screening rejected applications, identifying potential defaulters and hence minimizing Type II error. As mentioned earlier, accepting the bad applications is more costly to the financial institution. This means extending credit to those customers who have a high probability of default which could lead to financial distress and business failure.

The current paper provides insights into the potential and limitation of using two quantitative models: the RBF and the LR for credit scoring applications in the Jordanian commercial banks. The results suggest that the LR is more accurate and interpretive than the RBF model although the RBF showed encouraging results for screening bad applications. However, the decision on the best model is up to the bank's management. Furthermore, this study proposes a further study that compares between different types of ANNs with other statistical techniques such as linear discriminant analysis and decision trees.

Appendix A. Correlation matrix for the variables

	G	LP	CT	GU	N	A	DPR	LA	TI	PE	D	IR	CD
G	1												
LP	0.05	1											
CT	0.02	0.14	1										
GU	0.05	−0.05	−0.11	1									
N	−0.04	0.07	−0.02	0.02	1								
A	0.01	−0.04	−0.09	−0.01	0.03	1							
DPR	0.09	−0.07	−0.01	−0.02	−0.01	0.06	1						
LA	0.09	0.02	−0.02	−0.09	−0.06	0.07	0.43	1					
TI	0.06	0.19	0.03	−0.12	0.03	0.08	0.27	0.52	1				
PE	−0.03	−0.03	−0.03	−0.05	0.04	0.57	−0.07	−0.01	−0.03	1			
D	0.13	−0.26	−0.23	0.04	−0.17	−0.01	0.23	0.29	−0.04	0.04	1		
IR	−0.09	−0.37	−0.20	0.16	−0.04	0.01	−0.50	−0.54	−0.46	0.11	−0.09	1	
CD	−0.06	0.34	0.25	0.07	0.10	−0.12	−0.58	−0.17	−0.06	0.03	−0.30	0.07	1

References

Abdou, H., Pointon, J., El-Masry, A., 2007. On the applicability of credit scoring models in Egyptian banks. Banks Bank Syst. 2 (1), 4–19.

Akkoc, S., 2012. An empirical comparison of conventional techniques, neural networks and the three stage hybrid adaptive neuro fuzzy inference system (ANFIS) model for credit scoring analysis: the case of Turkish credit card data. Eur. J. Operat. Res. 222, 168–178.

Angelini, E., Tollo, G., Roli, A., 2008. A neural network approach for credit risk evaluation. Q. Rev. Econ. Finance 48, 733–755.

Association of Banks in Jordan, 2010. Annual Report, 32. Amman, Jordan.

Bekhet, H., Eletter, S., 2012. Credit risk management for the Jordanian commercial banks: a business intelligence approach. Aust. J. Basic Appl. Sci. 6 (9), 188–195.

Bensic, M., Sarlija, N., Zekic-Susac, M., 2005. Modeling small-business credit scoring by using logistic regression, neural networks and decision trees. Intellect. Syst. Account. Fin. Manage. 13 (3), 133–150.

Blanco, A., Mejias, R., Lara, J., Rayo, S., 2013. Credit scoring models for the microfinance industry using neural networks: evidence from Peru. Exp. Syst. Appl. 40 (1), 356–364.

Boguslauskas, V., Mileris, R., 2009. Estimation of credit risk by artificial neural networks models. Eng. Econ. 4, 7–14.

Bruzzone, L., Fernandez, P., 1998. Supervised training techniques for radial basis function neural network. Electron. Lett. 34 (11), 1115–1116.

Cao, Q., Parry, M., 2009. Neural network earning per share forecasting models: a comparison of backward propagation and genetic algorithm. Decision Support Syst. 47, 32–41.

Central Bank of Jordan (CBJ, 1988, 1989, 1990). Annual reports, Amman, Jordan.

Central Bank of Jordan (CBJ, 2002, 2006, 2010). Statistical database, money and banking, available at statisticaldb.cbj.gov.jo/

Central Bank of Jordan, 2011. Annual Repot, 48. Amman, Jordan.

Chen, M., Huang, S., 2003. Credit scoring and rejected instances reassigning through evolutionary computation techniques. Exp. Syst. Appl. 24, 433–441.

Eletter, S., 2012. Using data mining for an intelligent marketing campaign. Glob. Bus. Econ. Anthol. 2, 276–282.

Emel, A., Oral, M., Reisman, A., Yolalan, R., 2003. A credit scoring approach for the commercial banking sector. Socioecon. Plann. Sci. 37, 103–123.

Gao, L., Zhou, C., Gao, H.B., Shi, Y.R., 2006. Credit scoring model based on neural network with particle swarm optimization. Adv. Nat. Comput. 14, 76–79.

Gouvea, M., 2007. Credit risk analysis applying logistic regression, neural networks and genetic algorithms models. In: POMS 18th Annual Conference, Dallas, Texas, USA, 4–7 May, 2007.

Gujarati, D., 2003. Basic Econometrics. The McGraw-Hill Companies, Inc., New York.

Hand, D.J., Henley, W.E., 1997. Statistical classification methods in consumer credit scoring: a review. J. R. Stat. Soc. 160 (3), 523–541.

Heiat, A., 2012. Comparing performance of data mining models for computer credit scoring. J. Int. Fin. Econ. 12 (1), 78–83.

Huang, W., Lai, K., Nakamori, Y., Wang, S., 2004a. Forecasting foreign exchange rates with artificial neural networks: a review. Inter. J. Inf. Technol. Decis. Making 3 (1), 145–165.

Huang, Z., Chen, H., Hsu, C., Chen, W., Wu, S., 2004b. Credit rating analysis with support vector machines and neural networks: a market comparative study. Decis. Support Syst. 37 (4), 543–558.

Huang, J., Tzeng, G., Ong, C., 2006. Two-stage genetic programming (2SGP) for the credit scoring model. Appl. Math. Comput. 174 (2), 1039–1053.

Jagric, V., Kracun, D., Jagric, T., 2011. Does non-linearity matter in retail credit risk modeling? Finance a uver-Czech J. Econ. Fin. 61 (4), 384–402.

Khandani, A., Kim, A., Lo, A., 2010. Consumer credit-risk model via machine learning algorithms. J. Bank. Finance 34, 2767–2787.

Khashman, A., 2010. Neural network for credit risk evaluation: investigation of different neural models and learning schemes. Exp. Syst. Appl. 37 (9), 6233–6239.

Khrawish, H., 2011. Determinants of commercial banks performance: evidence from Jordan. Int. Res. J. Fin. Econ. 81, 147–159.

Koh, H., Tan, W., Goh, C., 2006. A two-step method to construct credit scoring models with data mining techniques. Int. J. Bus. Inform. 1 (1), 96–118.

Kurum, E., Yildirak, k., Weber, G., 2012. A classification problem of credit risk rating investigated and solved by optimization of ROC curve. Centr. Eur. J. Oper. Res. 20 (3), 529–557.

Lahsasna, A., Ainon, R., Wah, T., 2010. Credit scoring models using soft computing methods: a survey. Int. Arab J. Inform. Technol. 7 (2), 115–123.

Lahsasna, A., Ainon, R., Wah, T., 2008. Intelligent credit scoring model using soft computing approach. In: Paper presented at the International Conference on Computer and Communication Engineering, Kuala Lumpur, Malaysia, 13–15 May, 2008, pp. 396–402.

Lee, T., Chen, I., 2005. A two-stage hyprid credit scoring model using artificial neural networks and multivariate adaptive regression splines. Expert Syst. Appl. 28 (4), 743–752.

Lee, T., Chiu, C., Lu, C., Chen, I., 2002. Credit scoring using the hybrid neural discriminant technique. Exp. Syst. Appl. 23 (3), 245–254.

Malhorta, R., Malhorta, D.K., 2003. Evaluating consumer loans using neural networks. Omega 31 (2), 83–96.

Martens, D., Baesens, B., Van Gestel, T., Vanthienen, J., 2007. Comprehensible credit scoring models using rule extraction from support vector machines. Eur. J. Oper. Res. 183 (3), 1466–1476.

Memarian, H., Balasundram, S., 2012. Comparison between multi-layer perceptron and radial basis function networks for sediment load estimation in a tropical watershed. J. Water Resour. Prot. 4, 870–876.

Paliwal, M., Kumar, U., 2009. Neural networks and statistical techniques: a review of applications. Exp. Syst. Appl. 36 (1), 2–17.

Thomas, L., 2000. A survey of credit and behavioral scoring: forecasting financial risk of lending to consumers. Int. J. Forecast. 16, 149–172.

Thomas, L., Edelman, D., Crook, J., 2002. Credit Scoring and its Applications. Society for Industrial and Application Mathematics, Philadelphia PA, USA.

Tsai, C.F., Wu, J.W., 2008. Using neural network ensembles for bankruptcy prediction and credit scoring. Exp. Syst. Appl. 34 (4), 2639–2649.

West, D., 2000. Neural network credit scoring models. Comp. Operat. Res. 27 (11), 1131–1152.

West, D., Dellana, S., Qian, J., 2005. Neural networks ensemble strategies for financial decision applications. Comp. Operat. Res. 32, 2543–2559.

Witkowska, D., Kaminski, W., Kompa, K., Staniec, I., 2004. Neural networks as a supporting tool in credit granting procedure. e-journal: Inform. Technol. Econ. Manage. 2 (1), ISSN: 1643-8949.

Wu, C., Guo, Y., Zhang, X., Xia, H., 2010. Study of personal credit risk assessment based on support vector machine ensemble. Int. J. Innovative 6 (5), 2353–2360.

Xie, T., Yu, H., Wilamowski, B., 2011. Comparison between traditional neural networks and radial basis function networks. IEEE, 1194–1199.

Yap, P., Ong, S., Husain, N., 2011. Using data mining to improve assessment of credit worthiness via credit scoring models. Exp. Syst. Appl. 38 (10), 1374–1383.