PERGAMON

Expert Systems with Applications 23 (2002) 245–254

Expert Systems with Applications

www.elsevier.com/locate/eswa

Credit scoring using the hybrid neural discriminant technique

Tian-Shyug Lee[a,*], Chih-Chou Chiu[b], Chi-Jie Lu[c], I-Fei Chen[d]

[a]*Department of Business Administration, Fu-Jen Catholic University, Hsin-Chuang, Taipei 24205, Taiwan, ROC*
[b]*Institute of Commerce Automation and Management, National Taipei University of Technology, Taipei, Taiwan, ROC*
[c]*Institute of Applied Statistics, Fu-Jen Catholic University, Hsin-Chuang, Taipei, Taiwan, ROC*
[d]*Graduate Program of Management, Fu-Jen Catholic University, Hsin-Chuang, Taipei, Taiwan, ROC*

Abstract

Credit scoring has become a very important task as the credit industry has been experiencing double-digit growth rate during the past few decades. The artificial neural network is becoming a very popular alternative in credit scoring models due to its associated memory characteristic and generalization capability. However, the decision of network's topology, importance of potential input variables and the long training process has often long been criticized and hence limited its application in handling credit scoring problems. The objective of the proposed study is to explore the performance of credit scoring by integrating the backpropagation neural networks with traditional discriminant analysis approach. To demonstrate the inclusion of the credit scoring result from discriminant analysis would simplify the network structure and improve the credit scoring accuracy of the designed neural network model, credit scoring tasks are performed on one bank credit card data set. As the results reveal, the proposed hybrid approach converges much faster than the conventional neural networks model. Moreover, the credit scoring accuracies increase in terms of the proposed methodology and outperform traditional discriminant analysis and logistic regression approaches.

Keywords: Credit scoring; Discriminant analysis; Neural networks; Model basis

1. Introduction

Starting in the late 1960s, to decide whether to grant credit to customers has gained more and more attention for credit industry due to the industry has been experiencing double-digit growth rate during the past few decades. Credit scoring has become a very important task as the credit industry can benefit from improving cash flow, insuring credit collections, reducing possible risks and implementing better managerial decisions. More and more attention has been paid to credit scoring, and resulting in many different useful techniques, known as the credit scoring models, have been developed by the banks and researchers in order to solve the problems involved during the evaluation process. The objective of credit scoring models is to assign credit applicants to either a 'good credit' group that is likely to repay financial obligation or a 'bad credit' group whose application will be denied because of its high possibility of defaulting on the financial obligation. Therefore credit scoring problems are basically in the scope of the more general and widely discussed discrimination and classification problems (Anderson, 1984; Dillion & Goldstein, 1984; Hand, 1981; Johnson & Wichern, 1998; Morrison, 1990).

In the first beginning, financial institutions always utilized the rules or principles built by the analysts to decide whom to give credit. But it is impossible both in economic and manpower terms to conduct all works with the tremendous increase in the number of applicants. Therefore, there is a need to automate the credit approval decision process. Usually, credit scoring is applied to rank credit information and to target collection activities including the applicant's application form details and the information held by a credit reference agency on the applicant. Besides, the evaluation performance can be improved by using credit scoring with streamlining the process and allowing the credit professional to focus only on unusual accounts. Moreover, the credit scoring can give the credit professional an exposure perspective, mitigate the risk flexibility, and reduce the cost of credit analysis. As a result, accounts with high probability of default can be monitored and necessary actions can be taken in order to prevent the account from being default. In response, the statistical methods, nonparametric statistical methods, and artificial intelligence approaches have been proposed to support the credit decision (Thomas, 2000).

* Corresponding author. Tel.: +886-2-2903-1111x2905; fax: +886-2-2908-9219.

E-mail address: badm1004@mails.fju.edu.tw (T.-S. Lee).

PII: S0957-4174(02)00044-1

Generally, two essential linear statistical tools, discriminante analysis and logistic regression, were most commonly applied to construct credit scoring models. In fact, discriminante analysis is the first tool to be used in building credit scoring models. However, the utilization of linear discriminant analysis (LDA) has often been criticized because of its assumption of the categorical nature of the credit data and the fact that the covariance matrices of the good and bad credit classes are unlikely to be equal (Reichert, Cho, & Wagner, 1983). In addition to the LDA approach, logistic regression is an alternative to conduct credit scoring. Basically, the logistic regression model was emerged as the technique of choice in predicting dichotomous outcomes. For predicting dichotomous outcomes, logistic regression has been concluded as one of the most appropriate techniques (Lee, Jo, & Han, 1997). As a matter of fact, Harrell and Lee (1985) found that logistic regression is as efficient as the LDA approach. A number of explorations of logistic regression model for credit scoring applications have been reported in literature. In addition to these typical methodologies, credit scoring has also lends itself to a recent development of neural networks approach. Neural networks provide a new alternative to LDA and logistic regression, particularly in situations where the dependent and independent variables exhibit complex nonlinear relationships. Even though neural networks have shown to have better credit scoring capability than LDA and logistic regression (Desai, Conway, & Overstreet, 1997; Desai, Crook, & Overstreet, 1996; Jensen, 1992; Piramuthu, 1999; West, 2000). It is, however, also being criticized for its long training process in designing the optimal network's topology and hence has limited its applicability in handling credit scoring problems (Chung & Gray, 1999; Craven & Shavlik, 1997).

Aiming at improving the above-mentioned drawbacks of neural networks and increasing the credit scoring accuracies of the existing approaches, the objective of the proposed study is to explore the performance of credit scoring using a two-stage hybrid modeling procedure in integrating the LDA approach with neural networks technique. The rationale underlying the analyses is firstly to use LDA in modeling the credit scoring problems. Then the significant predictor variables are served as the input variables of the designed neural networks model. Besides, the credit scoring result of discriminant analysis is also included in the input layer as extra information trying to give a better initial solution and increasing the credit scoring accuracy. Please note that it is valuable to use discriminant analysis as a supporting tool for designing the topology of neural networks as we can learn more about the inner workings. Besides, as there is no theoretical method in determining the best input variables of a neural network model, the discriminant analysis procedure can be implemented as a generally accepted method for determining a good subset of input variables when many potential variables are considered and thus giving statistical support in deciding the input vector of the designed neural network model. To demonstrate the feasibility and effectiveness that the inclusion of the obtained predictor variables and the credit scoring results from discriminant analysis would improve the credit scoring accuracy of the neural network model, credit scoring tasks are performed on one bank credit card dataset. As to the structure of the designed neural network model, sensitivity analysis is firstly employed to solve the issue of finding the appropriate setup of the network's topology. Analytic results demonstrated that the proposed hybrid model provides a better initial solution and hence converges much faster than the conventional neural networks model. Besides, in comparison with the traditional neural network approach, the credit scoring accuracy increases in terms of the proposed hybrid methodology. Moreover, the superior credit scoring capability of the proposed technique can be observed by comparing the credit scoring results with those using linear discrimintant analysis and logistic regression approaches.

The rest of the paper is organized as follows. We will briefly review the literature of credit scoring and give a brief outline of discriminant analysis, logistic regression and neural networks in Section 2. The developments as well as the analytic results of credit scoring models using discriminant analysis, logistic regression, neural networks, and the hybrid neural discriminant approach are presented in Section 3. Finally Section 4 addresses the conclusion and discusses the possible future research areas.

2. Research methodology and literature review

2.1. Discriminant analysis

Discriminant analysis was first proposed by Fisher (1936) in the 1930s as a discrimination and classification tools. Nowadays, discriminant analysis has been reported as the most commonly discussed and used statistical technique in modeling classification tasks (Lee, Sung, & Chang, 1999). According to some attributes of the predictor variables, discriminant analysis tends to look for the best linear combination of the predictor variables to classify the studying objects into two or more populations at the optimum accuracy (Cooper & Emory, 1995; Dillion & Goldstein, 1984; Johnson & Wichern, 1998).

As to the statistical assumptions in implementing discriminant analysis, Johnson and Wichern (1998) explicated that discriminant analysis requires the data to be independent and normally distributed while the covariance matrix is also required to comply with the variation homogeneity assumption. If the covariance matrices of the given populations are not equal, then the separation surface of the discriminant function is quadratic and hence in this case the quadratic discrininant analysis (QDA) needs to be used. Despite the fact that LDA is only a special case of QDA with stronger assumptions which should restrict its applications, in fact LDA has been reported to be a more robust method when the theoretical presumptions are violated (Dillon & Goldstein, 1984; Sanchez & Sarabia, 1995; Sharma, 1996). And hence the LDA approach will be used in building the credit scoring model in this paper. Given that the covariance matrices conforming to the prior assumptions of variation homogeneity, Fisher's linear discriminant function is allowed to be used in this paper (Johnson & Wichern, 1998). The LDA can be expressed as

$$D = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n$$

Where D represents the discriminant score, β_0 is the intercept term, and β_i $(i = 1, \ldots, n)$ represents the β coefficient associated with the corresponding explanatory variable X_i $(i = 1, \ldots, n)$.

As we expect, discriminant analysis has been widely devoted to a considerably wide range of application areas, such as medicine, business, education, marketing research, finance, chemistry, biology, engineering and archaeology. Lee et al. (1999) used discriminant analysis to conduct bankruptcy prediction and indicated that discriminant analysis is the most commonly used technique applied for bankruptcy prediction. Kim, Kim, Kim, Ye, and Lee (2000) endeavored to implement a classification analysis on the real estate markets in Korea and to forecast the consumer behaviors. Trevino and Daniels (1995) adopted discriminant analysis in judging the resultant performance from investments and of whether direct investments in American market do have substantial impacts on cooperate investors. Bardos (1998), Desai et al. (1996), Eisenbeis (1978), Falbo (1991), Grablowsky and Talley (1981), Martell and Fitts (1981), Overstreet and Bradley (1994), Overstreet, Bradley, and Kemp (1992), Reichart et al. (1983), and Titterington (1992) as well employed the same analytic tool to build credit scoring models for general bank and credit card sectors.

2.2. Logistic regression

Logistic regression is a widely used statistical modeling technique in which the probability of a dichotomous outcome is related to a set of potential predictor variables in the form:

$$\log[p/(1-p)] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_i x_i$$

where p is the probability of the outcome of interest, β_0 is the intercept term, and β_i $(i = 1, \ldots, n)$ represents the β coefficient associated with the corresponding explanatory variable x_i $(i = 1, \ldots, n)$ (Cox, 1970; Cox & Snell, 1989; Hosmer & Lemeshow, 1989; Pampel, 2000). The dependent variable is the logarithm of the odds, $\{\log[p/(1-p)]\}$, which is the logarithm of the ratio of two probabilities of the outcome of interest. These variables are usually selected for inclusion by using some form of backward or forward stepwise regression technique (Neter, Kutner, Nachtsheim, & Wasserman, 1996; Pampel, 2000) even thought these selection techniques may be prone to problems. And the maximization of the likelihood function is usually applied as the convergent criterion to estimate the coefficients of corresponding parameters when the logistic regression models are utilized.

The logistic regression model does not necessarily require the assumptions of LDA. However, Harrell and Lee (1985) found that logistic regression is as efficient and accurate as LDA even though the assumptions of LDA are satisfied. One advantage of LDA is that ordinary least-square estimation procedure can be implemented to estimate the coefficients of the linear discriminant function, whereas maximum likelihood methods are required for the estimation of logistic regression models. Another advantage of LDA over logistic regression is that prior probabilities and misclassification costs can easily be incorporated into the LDA approach (Desai et al., 1996).

Logistic regression models have been widely used in social research, medical research, biological research, food science, design, control, bankruptcy prediction, market segmentation, and customer behaviors (Kay, Warde, & Martens, 2000; Laitinen & Laitinen, 2000; Suh, Noh, & Suh, 1999; Vellido, Lisboa, & Vaughan, 1999; Wong, Bodnovich, & Selvi, 1997). Logistic regression has also been explored by Henley (1995), Joanes (1993), Laitinen (1999), Westgaard and van der Wijst (2001), Wiginton (1980) in building credit scoring models for personal loan, business loan, and credit card applications.

2.3. Artificial neural networks

A neural network is a computer-intensive, algorithmic procedure for transforming inputs into desired outputs using highly inter-connected networks of relatively simple processing elements (often termed neurons, units or nodes—we will use nodes thereafter). Neural networks are modeled following the neural activity in the human brain. The essential features of a neural network are the nodes, the network architecture describing the connections between the nodes, and the training algorithm used to find values of the network parameters (weights) for a particular network. The nodes are connected to one another in the sense that the output from one node can be served as the inputs to other nodes. Each node transforms an input to an output using some specified function that is typically monotone, but otherwise arbitrary. This function depends on parameters whose values must be determined with a training set of inputs and outputs. Network architecture is the organization of nodes and the types of connections permitted. The nodes are arranged in a series of layers with connections between nodes in different layers, but not between nodes in the same layer. The layer receiving the inputs is called the input layer. The final layer providing the target output signal is the output layer. Any layers between the input and output layers are hidden layers. A simple representation of a neural network with one hidden layer can be shown in Fig. 1 (Rumelhart, Hinton, & Williams, 1986).

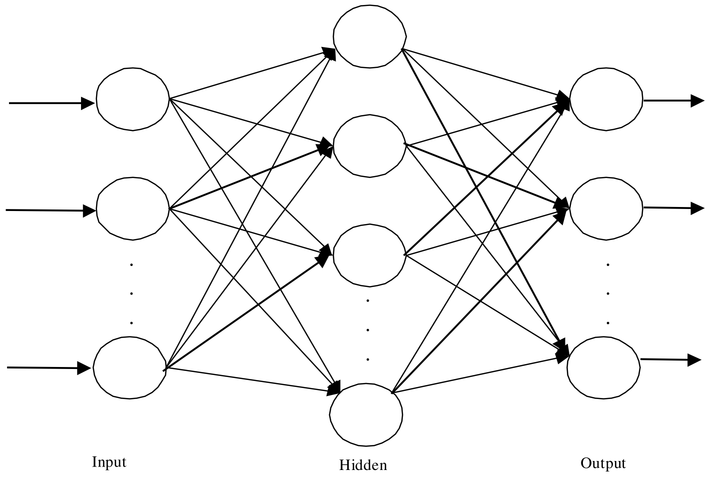

Fig. 1. A three-layer backpropagation neural networks.

Neural networks can be classified into two different categories, feedforward and feedback networks. The feedback networks contain nodes that can be connected to each other, enabling a node to influence other nodes as well as itself. Kohonen self-organizing network and the Hopfield network are examples of this type of network. The nodes in feedforward networks can take inputs only from the previous layer and send outputs to the next layer. The ADALINE and backpropagation neural networks (BPN) are two typical examples of this kind of network. BPN is a network essentially using a gradient steepest descent training algorithm and has been the most often utilized paradigm to date. For the gradient descent training algorithm, the step size, called the learning rate, must be specified first. The learning rate is crucial for BPN since smaller learning rates tend to slow down the learning process before convergence while larger ones may cause network oscillation and unable to converge.

Neural networks are increasingly found to be useful in modeling non-stationary processes due to its associated memory characteristics and generalization capability (Stern, 1996). More and more computer scientists and statisticians have interests in the computational potentials of neural network algorithms. Haykin (1994) wrote a comprehensive reference on artificial neural networks. Anderson and Rosenfeld (1988) edited a collection of papers that chronicled the major developments in neural network modeling. Cheng and Titterington (1994), Repley (1994), and Stern (1996) provided surveys describing the relevance of neural networks to the statistics community. As to the issue of determining the appropriate network topology: the number of layers, and the number of neurons in each layer, and the appropriate learning rates (Cybenko, 1989; Davies, 1994; Hecht-Nielsen, 1990; Hornik, Stinchcombe, & White, 1989; Kang, 1991; Lippmann, 1987; Tang & Fishwick, 1993; Wong, 1991).

Neural networks have been widely used in engineering, science, education, social research, business, forecasting and related fields (Cheng & Titterington, 1994; Repley, 1994; Stern, 1996; Vellido et al., 1999; Wong et al., 1997; Zhang, Patuwo, & Hu, 1998). Neural networks have also been explored by Arminger, Enache, and Bonne (1997), Barney, Graves, and Johnson (1999), Borowsky (1995), Deng (1993), Desai et al. (1997), Desai et al. (1996), Glorfeld (1996), Glorfeld and Hardgrave (1996), Hand and Henley (1997), Jagielska and Jaworski (1996), Jensen (1992), Piramuthu (1999), Piramuthu, Shaw, and Gentry (1994), Richeson, Zimmermann, and Barnett (1994), Robins (1993), Torsun (1996), West (2000), and Williamson (1995) in handling credit scoring problems. The majority of the above references have reported that the credit scoring accuracies of neural networks are better than those using discriminant analysis and logistic regression techniques.

3. Empirical study

In order to verify the feasibility and effectiveness of the proposed two-stage hybrid modeling procedure, one credit card dataset provided by a local bank in Taipei, Taiwan is used in this study. Each bank customer in the dataset contains nine predictor variables, namely, gender, age, marriage status, educational level, occupation, job position, annual income, residential status and credit limits. And the response variable is the credit status of the customer-good or bad credit. Six thousand datasets with respect to the ratio of good and bad credits were randomly selected and then used to build the credit scoring models. Among them, 4000 datasets will be used as the model building set (training sample) while the remaining 2000 will be retained as the validation set (testing sample).

The neural network simulator Qnet97 (1998), developed by Vesta Services Inc., was utilized to develop the neural networks as well as the hybrid credit scoring models. It is a C based simulator that provides a system for developing various neural network configurations using the generalized delta learning algorithm. And the discriminant analysis and logistic regression credit scoring models will be

Table 1
Classification results using discriminant analysis

Actual class	Classified class	
	1 (good credit)	2 (bad credit)
1 (good credit)	739 (74.57%)	252 (25.43%)
2 (bad credit)	320 (31.71%)	689 (68.29%)
Average correct classification rate: 71.40%		

Table 2
Classification results using logistic regression

Actual class	Classified class	
	1 (good credit)	2 (bad credit)
1 (good credit)	755 (76.19%)	236 (23.81%)
2 (bad credit)	295 (29.24%)	714 (70.76%)
Average correct classification rate: 73.45%		

implemented using the popular SPSS software (SPSS, 1998). All the modeling tasks are implemented on an IBM PC with Intel Pentium II 300 MHz CPU processor. The detailed credit scoring results using the above-mentioned four modeling techniques can be summarized as follows.

3.1. Discriminant analysis

Among the variable selection procedures which can be used in this study, the stepwise discriminant approach (Johnson & Wichern, 1998; Neter et al., 1996) is adopted in building the discriminant analysis credit scoring model. Six significant predictor variables are selected in the final discriminant function, namely gender, age, occupation, credit limits, annual income and residential status. The credit scoring results of the training sample using the obtained discriminant function are summarized in Table 1. From the results revealed in Table 1, we can observe that the average correct classification rate is 71.40% with 252 (320) class 1 (2) customers misclassified as class 2 (1) customers (here a class 1 customer is a customer whose credit status is good while a class 2 customer is the one whose credit is bad).

3.2. Logistics regression

The stepwise logistic regression procedure (Neter et al., 1996) is used in building the credit scoring model. Four significant variables, gender, age, credit limits, and annual income were included in the final regression model with the credit scoring results summarized in Table 2. From the result in Table 2, it is observed that the average correct classification rate is 73.45% with 236 (295) class 1 (2) customers misclassified as class 2 (1) customers.

3.3. Neural networks model

Since Vellido et al. (1999) pointed out that more than 75% of business applications using neural networks will use the BPN training algorithms, this study will also use the popular BPN in building the neural network credit scoring model. As recommended by Cybenko (1989), and Hornik et al. (1989) that one-hidden-layer network is sufficient to model any complex system with any desired accuracy, the designed network model will have only one hidden layer. And since there are nine input nodes in the input layer, the initial number of hidden nodes to be tested was chosen to be 17, 18, 19, 20, and 21. And the network has only one output node, the credit status of the customer. As Rumelhart et al. (1986) concluded that lower learning rates tended to give the best network results and the networks were unable to converge when the learning rate is greater then 0.006, learning rates 0.002, 0.004, and 0.006 are tested during the training process. The convergence criteria used for training are a root mean squared error (RMSE) less than or equal to 0.0001 or a maximum of 3000 iterations. The network topology with the minimum testing RMSE is considered as the optimal network topology.

The prediction results of the neural network with combinations of different hidden nodes and learning rates are summarized in Table 3. From Table 3, the {9–19–1} topology with a learning rate of 0.004 gives the best result (minimum testing RMSE). Here 9–19–1 stands for nine neurons in the input layer, nineteen neurons in the hidden layer, and only one neuron in the output layer. To examine the convergence characteristics of the proposed neural networks model, the RMSE during the training process for the {9–19–1} network with the learning rate of 0.004 are depicted in Fig. 2.

The credit scoring results using the designed BPN model can be summarized in Table 4. From the result in Table 4, we can observe that the average correct classification rate is 73.70% with 146 (380) class 1 (2) customers misclassified as class 2 (1) customers. By comparing the results of Tables 2–4, it can be observed that BPN has the highest average correct classification rate in comparison with discriminant analysis and logistic regression approaches.

3.4. Hybrid neural discriminant model

The single-layer BPN model will again be used in building the hybrid credit scoring model. The input layer of the hybrid model contains seven input nodes, as the hybrid model will use the significant predictor variables and the credit scoring results of the obtained LDA scoring model as the input nodes. As there are seven input nodes in the input layer, the initial number of hidden nodes to be tested was chosen to be 13–17. And the network has only one output node, the credit status of the customer. As the networks were still unable to converge when the learning rate is greater then 0.006, learning rates 0.002, 0.004, and 0.006 are tested during the training process. The convergence criteria are the same with the BPN model in Section 3.3. The network

Table 3
BPN model prediction results

Number of nodes in the hidden layer	Learning rate	Training RMSE	Testing RMSE
17	0.002	0.299264	0.302165
	0.004	0.296890	0.299634
	0.006	0.294587	0.297553
18	0.002	0.300309	0.303042
	0.004	0.295744	0.298386
	0.006	0.294342	0.297304
19	0.002	0.296243	0.298976
	0.004	*0.294417*	*0.297184*
	0.006	0.298858	0.301749
20	0.002	0.300296	0.303552
	0.004	0.297616	0.300309
	0.006	0.297047	0.29993
21	0.002	0.299529	0.302198
	0.004	0.295362	0.298191
	0.006	0.299561	0.302636

topology with the minimum testing RMSE is considered as the optimal network topology.

The prediction results of the hybrid model with combinations of different hidden nodes and learning rates are summarized in Table 5. From Table 5, the {7–16–1} topology with a learning rate of 0.006 gives the best result. The RMSE during the training process for the {7–16–1} network are depicted in Fig. 3. In comparison with the credit scoring model only using BPN showing in Fig. 2, the better convergence characteristics of RMSE of the proposed hybrid model can easily be observed.

The credit scoring results using the hybrid model are summarized in Table 6. Table 6 reveals that the average correct classification rate is 77.00% with 183 (277) class 1 (2) customers misclassified as class 2 (1) customers.

Finally, in order to evaluate the credit scoring capabilities of the above four constructed credit scoring models, the summarized results can be shown in Table 7. From the results revealed in Table 7, we can conclude that the hybrid neural discriminant model has the best credit scoring capability in terms of the average classification rate in comparison with LDA, logistic regression, and pure neural network models. Consequently, we can conclude that the credit scoring results of the hybrid neural discriminant model practically support the theoretical presumptions.

3.5. Type I and Type II errors of the constructed models

It is well known that, in order to evaluate the overall credit scoring capability of the designed credit scoring models, the misclassification costs also have to be taken into account. It is apparent that the costs associated with Type I error (a customer with good credit is misclassified as a customer with bad credit) and Type II

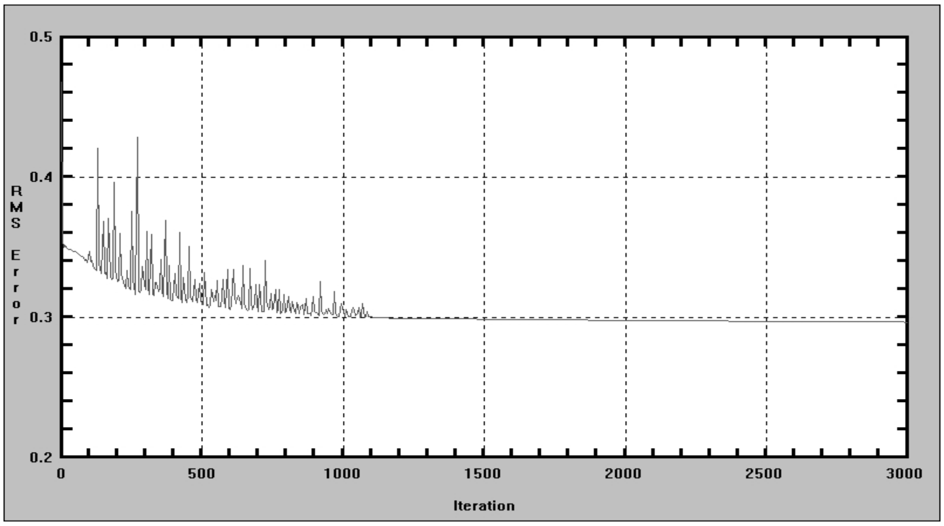

Fig. 2. The RMSE history in the training process for the proposed network.

Table 4
Classification results using BPN

Actual class	Classified class	
	1 (good credit)	2 (bad credit)
1 (good credit)	845 (85.27%)	146 (14.73%)
2 (bad credit)	380 (37.66%)	629 (62.34%)
Average correct classification rate: 73.70%		

error (a customer with bad credit is misclassified as a customer with good credit) are significantly different. In general, the misclassification costs associated with Type II errors are much higher than those associated with Type I errors. As recommended by Dr Hofmann who compiled the German credit data reported that the relative ratio of misclassification costs associated with Type I and Type II errors is 1–5 (West, 2000), and hence special attention should pay to Type II errors of the four constructed models in order to evaluate the overall credit scoring capability. Table 8 summarizes the Type I and Type II errors of the four models being discussed. As the results revealed in Table 8, the hybrid model has the lowest Type II error in comparison with the other three approaches. And hence we can conclude that the hybrid model not only has the best average classification rate, but also has the lowest Type II error and hence can successfully reduce the possible risks of extra losses due to high misclassification costs associated with Type II errors.

Table 5
Integrated hybrid model prediction results

Number of nodes in the hidden layer	Learning rate	Training RMSE	Testing RMSE
13	0.002	0.293617	0.308529
	0.004	0.291904	0.305402
	0.006	0.293828	0.304303
14	0.002	0.292537	0.306499
	0.004	0.292339	0.304224
	0.006	0.294474	0.306317
15	0.002	0.292636	0.307391
	0.004	0.292386	0.304938
	0.006	0.292793	0.302487
16	0.002	0.292969	0.307134
	0.004	0.293267	0.307103
	0.006	*0.296229*	*0.300828*
17	0.002	0.294684	0.313635
	0.004	0.293358	0.304002
	0.006	0.295560	0.303104

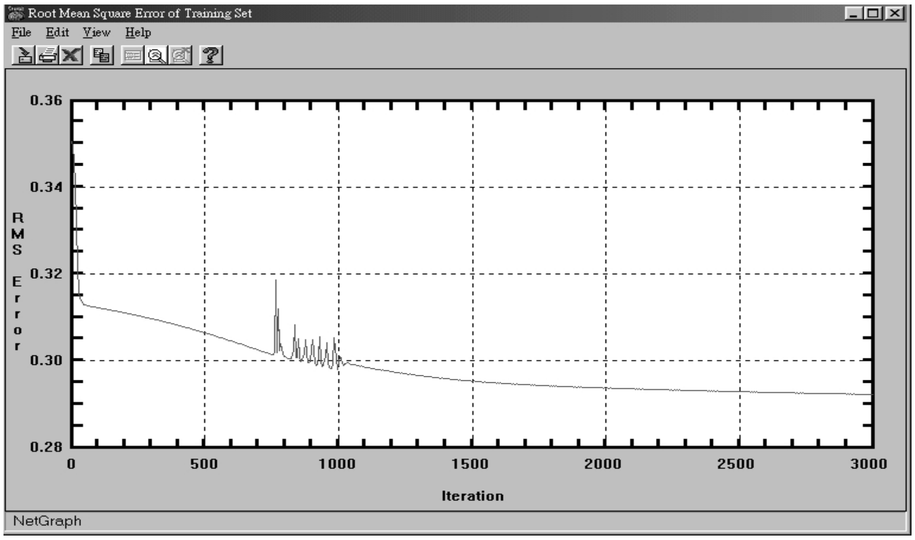

Fig. 3. The RMSE history in the training process for the hybrid model.

Table 6
Classification results using the hybrid model

Actual class	Classified class	
	1 (good credit)	2 (bad credit)
1 (good credit)	808 (81.53%)	183 (18.47%)
2 (bad credit)	277 (27.45%)	732 (72.55%)
Average correct classification rate: 77.00%		

4. Conclusions and areas of future research

Credit scoring has gained more and more attention as the competition between financial institutions has come to a totally conflicting stage. More and more companies are seeking better strategies through the help of credit scoring models and hence credit scoring techniques have been widely used in different credit application areas. And hence credit scoring problems are one of the applications that have gained serious attention over the past decades. Modeling techniques like traditional statistical analyses and artificial intelligence techniques have been developed in order to successfully attack the credit scoring tasks. Discriminant analysis is the most commonly used statistical credit scoring techniques, but often being criticized due to its strong model assumptions. On the other hand, the artificial neural networks is becoming a very popular alternative in the credit scoring tasks due to its associated memory characteristic, generalization capability and outstanding credit scoring capability. However, it is also being criticized for its long training process. The purpose of this study is to explore the performance of credit scoring using a two-stage hybrid modeling procedure in integrating the discriminant analysis approach with artificial neural networks technique. Discriminant analysis is firstly used in modeling the credit scoring problems with the significant predictor variables as the input variables of the designed neural networks model. And the credit scoring result of discriminant analysis is also included in the input layer as extra information trying to give a better initial solution and increase the credit scoring accuracy. Using discriminant analysis as the first-stage tool cannot only provide statistical support in deciding the input variables, but also give a better initial solution for the designed neural network model in getting better credit scoring accuracies.

Table 8
Type I and Type II errors of the four models

Credit scoring models	Performance assessment (%)	
	Type I error	Type II error
Discriminant analysis	25.43	31.71
Logistic regression	23.81	29.24
Back-propagation neural networks	14.73	37.66
Hybrid neural discriminant model	18.47	27.45

For verifying the feasibility on this proposed integrated approach, the credit scoring task is performed on one bank credit card dataset. Analytic results demonstrate that neural network model has the highest average correct classification rate in comparison with discriminant analysis and logistic regression approaches and justify the presumptions that neural networks having better capability of capturing nonlinear relationship among variables. Besides, the designed hybrid model not only has better credit scoring accuracies, but also has the lowest Type II error associated with high misclassification costs. Finally, the hybrid model did get help from the analysis of discriminant analysis in providing statistical support, reducing the number of input variables, and providing a better initial solution. The research findings did support the hypothesis that the two-stage hybrid credit scoring approach proposed in this study will have better credit scoring accuracies and better convergence characteristics for the designed neural networks model.

Future researches may aim at collecting more important variables that will increase the credit scoring accuracies. Using other newly developed classification methodologies, like classification and regression tree (CART) and multivariate adaptive regression splines (MARS), in evaluating their credit scoring capabilities are also recommended. Integrating other artificial intelligence techniques, like fuzzy discriminant analysis, genetic algorithms and gray theory, with neural networks in further refining the network structure and improving the credit scoring accuracies may also being discussed. Other related topics about credit cards like customer retention, market basket analysis, and customer profit analysis models may also being studied in future studies.

Table 7
Credit scoring results of the four constructed models

Credit scoring models	Credit scoring results (%)		
	{1–1}	{2–2}	Average correct classification rate
Discriminant analysis	74.57	68.29	71.40
Logistic regression	76.19	70.76	73.45
Backpropagation neural networks	85.27	62.34	73.70
Hybrid neural discriminant model	81.53	72.55	77.00

Acknowledgments

This research was partially supported by Fu-Jen Catholic University and the National Science Council of the Republic of China under Grant Number NSC 89-2213-E-030-016. Special thanks also goes to the anonymous bank in providing the credit card data that makes this research possible.

References

Anderson, T. W. (1984). *An introduction to multivariate statistical analysis*. New York, NY: Wiley.

Anderson, J. A., & Rosenfeld, E. (1988). *Neurocomputing: Foundations of research*. Cambridge: MIT Press.

Arminger, G., Enache, D., & Bonne, T. (1997). Analyzing credit risk data: A comparison of logistic discriminant classification tree analysis and feedforward networks. *Computational Statistics*, *12*, 293–310.

Bardos, M. (1998). Detecting the risk of company failure at the Banque de France. *Journal of Banking and Finance*, *22*, 1405–1419.

Barney, D. K., Graves, O. F., & Johnson, J. D. (1999). The farmers home administration and farm debt failure prediction. *Journal of Accounting and Public Policy*, *18*, 99–139.

Borowsky, M. (1995). Scoring puts up higher numbers. *US Banker*, *105*(1), 63.

Cheng, B., & Titterington, D. M. (1994). Neural network: A review from a statistical perspective (with discussion). *Statistical Science*, *9*, 2–54.

Chung, H. M., & Gray, P. (1999). Special section: Data mining. *Journal of Management Information Systems*, *16*, 11–16.

Cooper, D. R., & Emory, C. W. (1995). *Business research method*. Orlando, FL: Dryden.

Cox, D. R. (1970). *Analysis of binary data*. London: Methuen.

Cox, D. R., & Snell, E. J. (1989). *Analysis of binary data*. London: Chapman & Hall.

Craven, M. W., & Shavlik, J. W. (1997). Using neural networks for data mining. *Future Generation Computer Systems*, *13*, 221–229.

Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematical Control Signal Systems*, *2*, 303–314.

Davies, P. C. (1994). Design issues in neural network development. *Neurovest*, *5*, 21–25.

Deng, P.-S. (1993). Automatic knowledge acquisition and refinement for decision support: A connectionist inductive inference model. *Decision Sciences*, *24*(2), 371–393.

Desai, V. S., Conway, J. N., & Overstreet, G. A., Jr. (1997). Credit scoring models in the credit union environment using neural networks and genetic algorithms. *IMA Journal of Mathematics Applied in Business and Industry*, *8*, 324–346.

Desai, V. S., Crook, J. N., & Overstreet, G. A., Jr. (1996). A comparison of neural networks and linear scoring models in the credit union environment. *European Journal of Operational Research*, *95*, 24–37.

Dillon, W. R., & Goldstein, M. (1984). *Multivariate analysis methods and applications*. New York: Wiley.

Eisenbeis, R. A. (1978). Problems in applying discriminant analysis in credit scoring models. *Journal of Banking and Finance*, *2*, 205–219.

Falbo, P. (1991). Credit scoring by enlarged discriminant analysis. *OMEGA*, *19*(4), 275–289.

Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, *7*, 179–188.

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

Glorfeld, L. W., & Hardgrave, B. C. (1996). An improved method for developing neural networks—The case of evaluating commercial loan creditworthiness. *Computers and Operations Research*, *23*(10), 933–944.

Grablowsky, B. J., & Talley, W. K. (1981). Probit and discriminant factors for classifying credit applicants: A comparison. *Journal of Economics and Business*, *33*, 254–261.

Hand, D. J. (1981). *Discrimination and classification*. New York, NY: Wiley.

Hand, D. J., & Henley, W. E. (1997). Statistical classification methods in consumer credit scoring: A review. *Journal of the Royal Statistical Society, Series A*, *160*(3), 523–541.

Harrell, F. E., & Lee, K. L. (1985). A comparison of the discrimination of discriminant analysis and logistic regression. In P. K. Se (Ed.), *Biostatistics: Statistics in biomedical, public health, and environmental sciences*. Amsterdam: North-Holland.

Haykin, S. S. (1994). *Neural networks: A comprehensive foundation*. New York, NY: Macmillan.

Hecht-Nielsen, R. (1990). *Neurocomputing*. Menlo Park, CA: Addison-Wesley.

Henley, W. E. (1995). *Statistical aspects of credit scoring*. Doctoral Dissertation. Milton Keynes, UK: The Open University.

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximations. *Neural Networks*, *2*, 336–359.

Hosmer, D. W., & Lemeshow, S. (1989). *Applied logistic regression*. New York: Wiley.

Jagielska, I., & Jaworski, J. (1996). Neural network for predicting the performance of credit card accounts. *Computational Economics*, *9*(1), 77–82.

Jensen, H. L. (1992). Using neural networks for credit scoring. *Managerial Finance*, *18*, 15–26.

Joanes, D. N. (1993). Rejecting inference applied to logistic regression for credit scoring. *IMA Journal of Mathematics Applied in Business and Industry*, *5*, 35–43.

Johnson, R. A., & Wichern, D. W. (1998). *Applied multivariate statistical analysis* (Fourth Edition). Upper Saddle River, NJ: Prentice-Hall.

Kang, S. (1991). *An investigation of the use of feedforward neural networks for forecasting*. PhD Thesis. Kent State University.

Kay, O. W., Warde, A., & Martens, L. (2000). Social differentiation and the market for eating out in the UK. *International Journal of Hospitality Management*, *19*(2), 173–190.

Kim, J. C., Kim, D. H., Kim, J. J., Ye, J. S., & Lee, H. S. (2000). Segmenting the Korean housing market using multiple discriminant analysis. *Construction Management and Economics*, *18*, 45–54.

Laitinen, E. K. (1999). Predicting a corporate credit analyst's risk estimate by logistic and linear models. *International Review of Financial Analysis*, *8*(2), 97–121.

Laitinen, E. K., & Laitinen, T. (2000). Bankruptcy prediction: Application of the Taylor's expansion in logistic regression. *International Review of Financial Analysis*, *9*(4), 327–349.

Lee, H., Jo, H., & Han, I. (1997). Bankruptcy prediction using case-based reasoning, neural networks, and discriminant analysis. *Expert Systems With Applications*, *13*, 97–108.

Lee, G., Sung, T. K., & Chang, N. (1999). Dynamics of modeling in data mining: interpretive approach to bankruptcy prediction. *Journal of Management Information Systems*, *16*, 63–85.

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

Martell, T. F., & Fitts, R. L. (1981). A quadratic discriminant analysis of bank credit card user characteristics. *Journal of Economics and Business*, *33*, 153–159.

Morrison, D. F. (1990). *Multivariate statistical methods*. New York, NY: McGraw-Hill.

Neter, J., Kutner, M. H., Nachtsheim, C. J., & Wasserman, W. (1996). *Applied linear statistical models*. Chicago, IL: Irwin.

Overstreet, G. A., Jr., Bradley, E. L., Jr. (1994). *Applicability of generic linear scoring models in the USA credit union environment: Further analysis*. Working Paper, University of Virginia.

Overstreet, G. A., Jr., Bradley, E. L., Jr., & Kemp, R. S. (1992). The flat-maximum effect and generic linear scoring model: A test. *IMA Journal of Mathematics Applied in Business and Industry*, *4*, 97–109.

Pampel, F. C. (2000). *Logistic regression—A premier*. Thousand Oaks, CA: Sage.

Piramuthu, S. (1999). Financial credit-risk evaluation with neural and neurofuzzy systems. *European Journal of Operational Research*, *112*, 310–321.

Piramuthu, S., Shaw, M. J., & Gentry, J. A. (1994). A classification approach using multi-layered neural networks. *Decision Support Systems*, *11*(5), 509–525.

Qnet 97 (1998). *Neural network modeling for Windows 95/98/NT*, Vesta Services, Winnetka, IL.

Reichert, A. K., Cho, C. C., & Wagner, G. M. (1983). An examination of the conceptual issues involved in developing credit-scoring models. *Journal of Business and Economic Statistics*, *1*, 101–114.

Repley, B. (1994). Neural networks and related methods for classification (with discussion). *Journal of the Royal Statistical Society, Series B*, *56*, 409–456.

Richeson, L., Zimmermann, R. A., & Barnett, K. G. (1994). Predicting consumer credit performance: Can neural networks outperform traditional statistical methods? *International Journal of Applied Expert Systems*, *2*(2), 116–130.

Robins, G. (1993). Credit scoring: Can retailers benefit from neural networks? *Stores*, 34–35.

Rumelhart, D. E., Hinton, D. E., & Williams, R. J. (1986) (*Vol. 1*). *Learning internal representations by error propagation in parallel distributed processing*, Cambridge, MA: MIT Press, pp. 318–362.

Sanchez, M. S., & Sarabia, L. A. (1995). Efficiency of multi-layered feedforward neural networks on classification in relation to linear discriminant analysis, quadratic discriminant analysis and regularized discriminant analysis. *Chemometrics and Intelligent Laboratory Systems*, *28*, 287–303.

Sharma, S. (1996). *Applied multivariate techniques*. New York, NY: Wiley.

SPSS (1998). *Statistic modeling for Windows 95/98/NT*. SPSS Inc.

Stern, H. S. (1996). Neural networks in applied statistics. *Technometrics*, *38*(3), 205–216.

Suh, E. H., Noh, K. C., & Suh, C. K. (1999). Customer list segmentation using the combined response model. *Expert Systems with Applications*, *17*(2), 89–97.

Tang, Z., & Fishwick, P. A. (1993). Feedforward neural nets as models for time series forecasting. *ORSA Journal on Computing*, *5*, 374–385.

Thomas, L. C. (2000). A survey of credit and behavioral scoring: Forecasting financial risks of lending to customers. *International Journal of Forecasting*, *16*, 149–172.

Titterington, D. M. (1992). Discriminant analysis and related topics. In J. N. Crook, & D. B. Edelman (Eds.), *Credit scoring and credit control* (pp. 53–73). Oxford: Oxford University Press.

Torsun, I. S. (1996). A neural network for a loan application scoring system. *The New Review of Applied Expert Systems*, *2*, 47–62.

Trevino, L. J., & Daniels, J. D. (1995). FDI theory and foreign direct investment in the United States: A comparison of investors and non-investors. *International Business Review*, *4*, 177–194.

Vellido, A., Lisboa, P. J. G., & Vaughan, J. (1999). Neural networks in business: A survey of applications (1992–1998). *Expert Systems With Applications*, *17*, 51–70.

West, D. (2000). Neural network credit scoring models. *Computers and Operations Research*, *27*, 1131–1152.

Westgaard, S., & van der Wijst, N. (2001). Default probabilities in a corporate bank Portfolio: A logistic model approach. *European Journal of Operational Research*, *135*(2), 338–349.

Wiginton, J. C. (1980). A note on the comparison of logit and discriminant models of consumer credit behavior. *Journal of Financial and Quantitative Analysis*, *15*, 757–770.

Williamson, A. G. (1995). Refining a neural network credit application system with a genetic algorithm. *Journal of Microcomputer Applications*, *18*, 261–277.

Wong, F. S. (1991). Time series forecasting using backpropagation neural networks. *Neurocomputing*, *2*, 147–159.

Wong, B. K., Bodnovich, A. T., & Selvi, Y. (1997). Neural network applications in business: a review and analysis of the literature (1988–1995). *Decision Support Systems*, *19*, 301–320.

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