

*Article*

## A Soft Intelligent Risk Evaluation Model for Credit Scoring Classification

Mehdi Khashei \* and Akram Mirahmadi

Department of Industrial Engineering, Isfahan University of Technology (IUT), Isfahan 84156-83111, Iran; E-Mail: isa.mirahmadi@in.iut.ac.ir

\* Author to whom correspondence should be addressed; E-Mail: khashei@cc.iut.ac.ir;  
Tel.: +98-313-391-1482.

Academic Editor: Nicholas Apergis

*Received: 27 June 2015 / Accepted: 25 August 2015 / Published: 8 September 2015*

---

**Abstract:** Risk management is one of the most important branches of business and finance. Classification models are the most popular and widely used analytical group of data mining approaches that can greatly help financial decision makers and managers to tackle credit risk problems. However, the literature clearly indicates that, despite proposing numerous classification models, credit scoring is often a difficult task. On the other hand, there is no universal credit-scoring model in the literature that can be accurately and explanatorily used in all circumstances. Therefore, the research for improving the efficiency of credit-scoring models has never stopped. In this paper, a hybrid soft intelligent classification model is proposed for credit-scoring problems. In the proposed model, the unique advantages of the soft computing techniques are used in order to modify the performance of the traditional artificial neural networks in credit scoring. Empirical results of Australian credit card data classifications indicate that the proposed hybrid model outperforms its components, and also other classification models presented for credit scoring. Therefore, the proposed model can be considered as an appropriate alternative tool for binary decision making in business and finance, especially in high uncertainty conditions.

**Keywords:** risk management; classification; credit scoring; soft computing techniques; artificial intelligent; Multi-Layer Perceptrons (MLPs)

**JEL Classification:** C38; C45; D81

---

## 1. Introduction

Credit risk analysis is an important topic in financial risk management. Several different applications are presented in the finance and banking literature for credit risk analysis, such as bank loans, credit cards, mortgages, hire purchase, *etc.* [1]. Credit scoring is generally concerned with evaluating the potential risks corresponding to granting scores. Credit scoring models help lenders to decide who will get credit, how much credit they should get, and what additional strategies will enhance the profitability of the borrowers to lenders. Some of the most important obtained advantages of a reliable credit-scoring model include [2,3]:

- Improving cash flows;
- Insuring proper credit collections;
- Reducing possible credit losses;
- Reducing cost of credit analysis, enabling credit decision almost instantaneously;
- Allowing to offer credit products geared to different risk levels;
- Analyzing the purchasing behavior of existing customers.

Durand [4] posited that the procedure includes collecting, analyzing, and classifying different credit elements and variables, in order to make credit-granting decisions. He noted that, to classify a firm's customers, the objective of the credit evaluation process is to reduce current and expected risk of a customer being "bad" for credit. Thus, credit scoring is an important technology for banks and other financial institutions as they seek to minimize risk [5]. Credit scoring can be considered as a quantitative method of measuring the risk attached to a potential customer, by analyzing their data to determine the likelihood that the prospective borrower will default on a loan [6]. Credit scoring can also be described as a data mining technique, employed to convert data into rules that can be used to guide credit-granting decisions.

The data mining methods for credit scoring were invented in the 1950s. Thus far, different data mining methods have been proposed and developed for credit scoring [7,8]. These methods can be generally divided into two main categories: hard and soft models. Soft data mining models, in contrast to the hard models, use the fuzzy data and/or fuzzy functions for credit scoring, instead of using crisp. In such that models, uncertainties in data and relationships are modeled by fuzzy logic and fuzzy operators. The hard data mining models can be also categorized in two major categories: statistical and intelligent models.

In the literature, a variety of statistical techniques have been proposed and developed for credit scoring. Some of the most well-known and widely-used of statistical credit scoring models include logistic regression [9,10], linear discriminant analysis [11–13], quadratic discriminant analysis, probit regression [14], nearest neighbor analysis [15], and Bayesian network [16,17]. Although statistical models have some valuable advantages in the explanatory purpose of credit scoring, the performances of these models can not satisfy financial managers and decision makers. Therefore, researchers are trying to find and develop more accurate scoring models. This leads to the investigation of new methods in artificial intelligence (AI), such as artificial neural network (ANNs). Several researchers have argued that using artificial neural networks can be an effective way of improving the accuracy in credit scoring.

Paliwal and Kumar [18] reported that such advanced techniques can be superior alternatives to traditional statistical models in practical tasks.

Although statistical and intelligent techniques are precise and accurate classification models, both of them are crisp models that use classical logic in their modeling procedures. Therefore, they cannot effectively model the uncertainties existing in the data and relationships [19]. Therefore, in order to overcome these shortcomings, a second category of credit scoring models has been proposed to the financial industry, soft models. Fuzzy forecasting and classification models, such as fuzzy time series and fuzzy linear regression, are suitable under uncertainty and complexity; however, their performance is not always satisfactory. Therefore, it is not wise to apply them blindly to any type of data. In the literature, several researchers have tried to overcome some of the disadvantages of the traditional fuzzy models for approximating real-world systems [20].

It is an area in which even a small performance improvement can mean a tremendous increase in profit to the lender because of the volume of lending made by using scores. Thus, these efforts lead to the proposal of hybrid methods, which use the advantages of various models. In the literature, using hybrid models, or combining several models, has become a common practice in order to overcome the limitations of single models and improve classification accuracy. Many researchers have argued that performance improves in hybridization. In combined models, the main aim is to reduce the risk of using an inappropriate model and failure by combining several models and obtain results that are more accurate. On the other hand, the motivation for using hybrid models comes from the assumption that neither can completely identify the data generating process. Hence, a single model may not be sufficient in order to identify all the characteristics of the data [21]. A great deal of effort has been devoted in the literature in order to develop and improve hybrid classification models.

In their pioneering work, Bates and Granger showed that a linear combination of forecasts would give a smaller error variance than any of the individual models. Since then, the studies on this topic have expanded dramatically. The basic idea of model combination in classification is to use the unique features of each classification model in order to capture different patterns in the data. Both theoretical and empirical findings suggest that combining different models can be an effective and efficient way to improve classification accuracy [22]. In recent years, more hybrid classification models have also been proposed, integrating different models together in order to improve accuracy. Lee *et al.* [23] have presented credit scoring by integrating back propagation neural networks with linear discriminate analysis. Hsieh [24] has presented a hybrid mining approach in the design of an effective credit-scoring model, integrating clustering algorithms and artificial neural networks. Luo *et al.* [25] have employed support vector machines (SVMs) and clustering-launched classification models for credit scoring.

Hung *et al.* [26] have also adopted three strategies to build hybrid SVM-based credit scoring models to evaluate an applicant's credit score from their features. Li *et al.* [27] have introduced a linear combination of kernel functions to enhance the interpretability of credit classification models, and proposed an alternative model to optimize the parameters based on the evolution strategy. Celikyilmaz and Turksen [20] have introduced a type II fuzzy function system for uncertainty modeling using evolutionary algorithms (ET2FF), and used it in real-life applications, such as financial market modeling. In this model, an improved fuzzy clustering is initially used to find the hidden structures, and then the genetic algorithm is applied to optimize the interval type II fuzzy sets. Chen and Li [28] have proposed a class of hybrid-SVM models. In these models, a support vector machine classifier is,

respectively, combined with conventional LDA, decision tree, F-score, and Rough sets, as pre-processing steps to optimize feature space by removing both redundant and irrelevant features.

Bijak and Thomas [29] have proposed a two-step and simultaneous approach, in which both segmentation and scorecards are optimized at the same time by using Logistic Trees. Chi and Hsu [30] have combined a bank's internal behavioral scoring model with the external credit bureau scoring model to construct the dual scoring model for credit risk management of mortgage accounts. Ping and Yongheng [31] have proposed a hybrid model by using the neighborhood rough set to select input features, and grid search to optimize RBF kernel parameters. Finally, they used hybrid optimal input features and model parameters to solve the credit-scoring problem. Kim and Han [32] have presented a hybrid method of Self-Organizing Map (SOM) and case-based reasoning (CBR) for the prediction of corporate bond rating. Park and Han [33] have attempted to integrate analytic hierarchy process with case-based reasoning as a tool of feature weighting to improve predictive performance of CBR in business failure predictions.

Ahn and Kim [34] have proposed a novel hybrid approach using genetic algorithm (GA) for case-based reasoning in corporate bankruptcies. Capotorti and Barbanera [35] have proposed a hybrid methodology for classification, based on the methodologies of fuzzy sets, partial conditional probability assessments, and rough sets. Akkoc [36] proposed a three-stage hybrid adaptive neuro-fuzzy inference system (ANFIS) for credit scoring, which is based on statistical techniques, artificial neural networks, and fuzzy logic. Laha [37] has proposed a hybrid soft credit-scoring model using fuzzy rule based classifiers. In this model, the rule base is first learned from the training data using a Self-Organizing Map (SOM), and then the fuzzy K-nearest neighbor rule is incorporated to design a contextual classifier that integrates the context information from the training set. Yao [38] proposed a hybrid fuzzy support vector machine (FSVM) for credit scoring using three strategies: (1) using classification and regression trees (CART) to select input features; (2) using multivariate adaptive regression splines (MARS) to select input features; (3) using GA to optimize model parameters.

In this paper, a soft intelligent hybrid classification model of traditional multi-layer perceptrons (MLPs) is proposed in order to yield more accurate results in credit scoring. In the proposed model, a multi-layer perceptron is first used to preprocess the raw data and provide a necessary background in order to apply a fuzzy regression model. Then, the obtained parameters of the first stage are considered in the form of fuzzy numbers and the optimum values of the parameters are calculated using the basic concept of fuzzy regression. In order to show the effectiveness and appropriateness of the proposed model, its performance in the Australian credit data set is compared with those of some fuzzy and nonfuzzy, linear and nonlinear, and intelligent classification models. Empirical results indicate that the proposed model is an effective method to improve classification accuracy.

The rest of the paper is organized as follows: In the next section, the formulation of the proposed hybrid model for classification tasks is introduced. In Section 3, the Australian credit data set is illustrated. In Section 4, the proposed model is applied to the Australian credit data set. In Section 5, the performance of the proposed model is compared to some other classification models, presented in the literature for Australian credit scoring. Finally, the conclusions are discussed.

## 2. Formulation of the Proposed Hybrid Model

Multi-layer perceptrons (MLPs) are flexible computing frameworks and universal approximators that can be applied to a wide range of classification problems with a high degree of accuracy. Several distinguishing features of multi-layer perceptrons make them valuable and attractive for classification tasks. The most important of these is that MLPs, as opposed to the traditional model-based techniques, are data-driven self-adaptive methods in that there are a priori assumptions about the models for problems under study [39]. The parameters of multi-layer perceptrons (MLPs) (weights and biases) are crisp ( $w_{i,j}$  ( $i = 0, 1, 2, \dots, p$   $j = 1, 2, \dots, q$ ),  $w_j$  ( $j = 0, 1, 2, \dots, q$ )). In the proposed model, instead of using crisp, fuzzy parameters, triangular fuzzy numbers are used ( $\tilde{w}_{i,j}$  ( $i = 0, 1, 2, \dots, p$   $j = 1, 2, \dots, q$ ),  $\tilde{w}_j$  ( $j = 0, 1, 2, \dots, q$ )). The model is described using a fuzzy function with a fuzzy parameter [40]:

$$\tilde{y}_t = f(\tilde{w}_0 + \sum_{j=1}^q \tilde{w}_j \cdot g(\tilde{w}_{0,j} + \sum_{i=1}^p \tilde{w}_{i,j} \cdot y_{t-i})), \quad (1)$$

where,  $y_t$  are observations,  $\tilde{w}_{i,j}$  ( $i = 0, 1, 2, \dots, p$   $j = 1, 2, \dots, q$ ),  $\tilde{w}_j$  ( $j = 0, 1, 2, \dots, q$ ) are fuzzy numbers. Equation (1) is modified as follows:

$$\tilde{y}_t = f(\tilde{w}_0 + \sum_{j=1}^q \tilde{w}_j \cdot \tilde{X}_{t,j}) = f(\sum_{j=0}^q \tilde{w}_j \cdot \tilde{X}_{t,j}), \quad (2)$$

where,  $\tilde{X}_{t,j} = g(\tilde{w}_{0,j} + \sum_{i=1}^p \tilde{w}_{i,j} \cdot y_{t-i})$ . Fuzzy parameters in the form of triangular fuzzy numbers  $\tilde{w}_{i,j} = (a_{i,j}, b_{i,j}, c_{i,j})$  are used:

$$\mu_{\tilde{w}_{i,j}}(w_{i,j}) = \begin{cases} \frac{1}{b_{i,j} - a_{i,j}}(w_{i,j} - a_{i,j}) & a_{i,j} \leq w_{i,j} \leq b_{i,j}, \\ \frac{1}{b_{i,j} - c_{i,j}}(w_{i,j} - c_{i,j}) & b_{i,j} \leq w_{i,j} \leq c_{i,j}, \\ 0 & \text{otherwise,} \end{cases} \quad (3)$$

where,  $\mu_{\tilde{w}}(w_{i,j})$  is the membership function of the fuzzy set that represents parameter  $w_{i,j}$ . By applying the extension principle, the membership of  $\tilde{X}_{t,j} = g(\tilde{w}_{0,j} + \sum_{i=1}^p \tilde{w}_{i,j} \cdot y_{t-i})$  in Equation (2) is given as shown in Equation (4) [40]:

$$\mu_{\tilde{X}_{t,j}}(X_{t,j}) = \begin{cases} \frac{(X_{t,j} - g(\sum_{i=0}^p a_{i,j} \cdot y_{t-i}))}{g(\sum_{i=0}^p b_{i,j} \cdot y_{t-i}) - g(\sum_{i=0}^p a_{i,j} \cdot y_{t-i})} & \text{if } g(\sum_{i=0}^p a_{i,j} \cdot y_{t-i}) \leq X_{t,j} \leq g(\sum_{i=0}^p b_{i,j} \cdot y_{t-i}), \\ \frac{(X_{t,j} - g(\sum_{i=0}^p c_{i,j} \cdot y_{t-i}))}{g(\sum_{i=0}^p b_{i,j} \cdot y_{t-i}) - g(\sum_{i=0}^p c_{i,j} \cdot y_{t-i})} & \text{if } g(\sum_{i=0}^p b_{i,j} \cdot y_{t-i}) \leq X_{t,j} \leq g(\sum_{i=0}^p c_{i,j} \cdot y_{t-i}), \\ 0 & \text{otherwise,} \end{cases} \quad (4)$$

where,  $y_{t,i} = 1$  ( $t = 1, 2, \dots, k$   $i = 0$ ),  $y_{t,i} = y_{t-i}$  ( $t = 1, 2, \dots, k$   $i = 1, 2, \dots, p$ ). Considering triangular fuzzy numbers,  $\tilde{X}_{t,j}$  with membership function in Equation (3) and triangular fuzzy parameters  $\tilde{w}_j$  the membership function of  $\tilde{y}_t = f(\tilde{w}_0 + \sum_{j=1}^q \tilde{w}_j \cdot \tilde{X}_{t,j}) = f(\sum_{j=0}^q \tilde{w}_j \cdot \tilde{X}_{t,j})$  is given as:

$$\mu_{\tilde{y}}(y_t) \equiv \begin{cases} \frac{-B_1}{2A_1} + \left[ \left( \frac{B_1}{2A_1} \right)^2 - \frac{C_1 - f^{-1}(y_t)}{A_1} \right]^{1/2} & \text{if } C_1 \leq f^{-1}(y_t) \leq C_3, \\ \frac{B_2}{2A_2} + \left[ \left( \frac{B_2}{2A_2} \right)^2 - \frac{C_2 - f^{-1}(y_t)}{A_2} \right]^{1/2} & \text{if } C_3 \leq f^{-1}(y_t) \leq C_2, \\ 0 & \text{otherwise,} \end{cases} \quad (5)$$

where,

$$\begin{aligned} A_1 &= \sum_{j=0}^q (e_j - d_j) \cdot \left( \frac{g\left(\sum_{i=0}^p b_{i,j} \cdot y_{t,i}\right)}{-g\left(\sum_{i=0}^p a_{i,j} \cdot y_{t,i}\right)} \right), \\ B_1 &= \sum_{j=0}^q (d_j \cdot \left( \frac{g\left(\sum_{i=0}^p b_{i,j} \cdot y_{t,i}\right)}{-g\left(\sum_{i=0}^p a_{i,j} \cdot y_{t,i}\right)} \right) + g\left(\sum_{i=0}^p a_{i,j} \cdot y_{t,i}\right) \cdot (e_j - d_j)), \\ A_2 &= \sum_{j=0}^q (f_j - e_j) \cdot \left( \frac{g\left(\sum_{i=0}^p c_{i,j} \cdot y_{t,i}\right)}{-g\left(\sum_{i=0}^p b_{i,j} \cdot y_{t,i}\right)} \right), \\ B_2 &= \sum_{j=0}^q (f_j \cdot \left( \frac{g\left(\sum_{i=0}^p c_{i,j} \cdot y_{t,i}\right)}{-g\left(\sum_{i=0}^p b_{i,j} \cdot y_{t,i}\right)} \right) + g\left(\sum_{i=0}^p c_{i,j} \cdot y_{t,i}\right) \cdot (f_j - e_j)), \\ C_1 &= \sum_{j=0}^q (d_j \cdot g\left(\sum_{i=0}^p a_{i,j} \cdot y_{t,i}\right)), \\ C_2 &= \sum_{j=0}^q (f_j \cdot g\left(\sum_{i=0}^p c_{i,j} \cdot y_{t,i}\right)), \\ C_3 &= \sum_{j=0}^q (e_j \cdot g\left(\sum_{i=0}^p b_{i,j} \cdot y_{t,i}\right)). \end{aligned}$$

The final point is that the output of the proposed model is fuzzy and continuous, while our classification problem differs in that its output is discrete and nonfuzzy. Therefore, in order to apply the proposed model to classification, certain modifications to the model needed to be made. For this purpose, each class is first assigned a numeric value, and then the membership probability of the output in each class is calculated as follows:

$$P_A = 1 - P_B = \frac{\int_{-\infty}^m f(x) dx}{\int_{-\infty}^{+\infty} f(x) dx} = 1 - \frac{\int_m^{+\infty} f(x) dx}{\int_{-\infty}^{+\infty} f(x) dx} \quad (6)$$

where  $P_A$  and  $P_B$  are the membership probability of class  $A$  and  $B$ , respectively, and  $m$  is the mean of the class values. Finally, the sample is put in the class with which its output has the largest probability. In the proposed model, due to the fact that output is fuzzy, it may be better to apply large class values. The larger class values expand small differences in the output, helping the model to become more sensitive to variations in the input.

### 3. The Australian Credit Data Sets

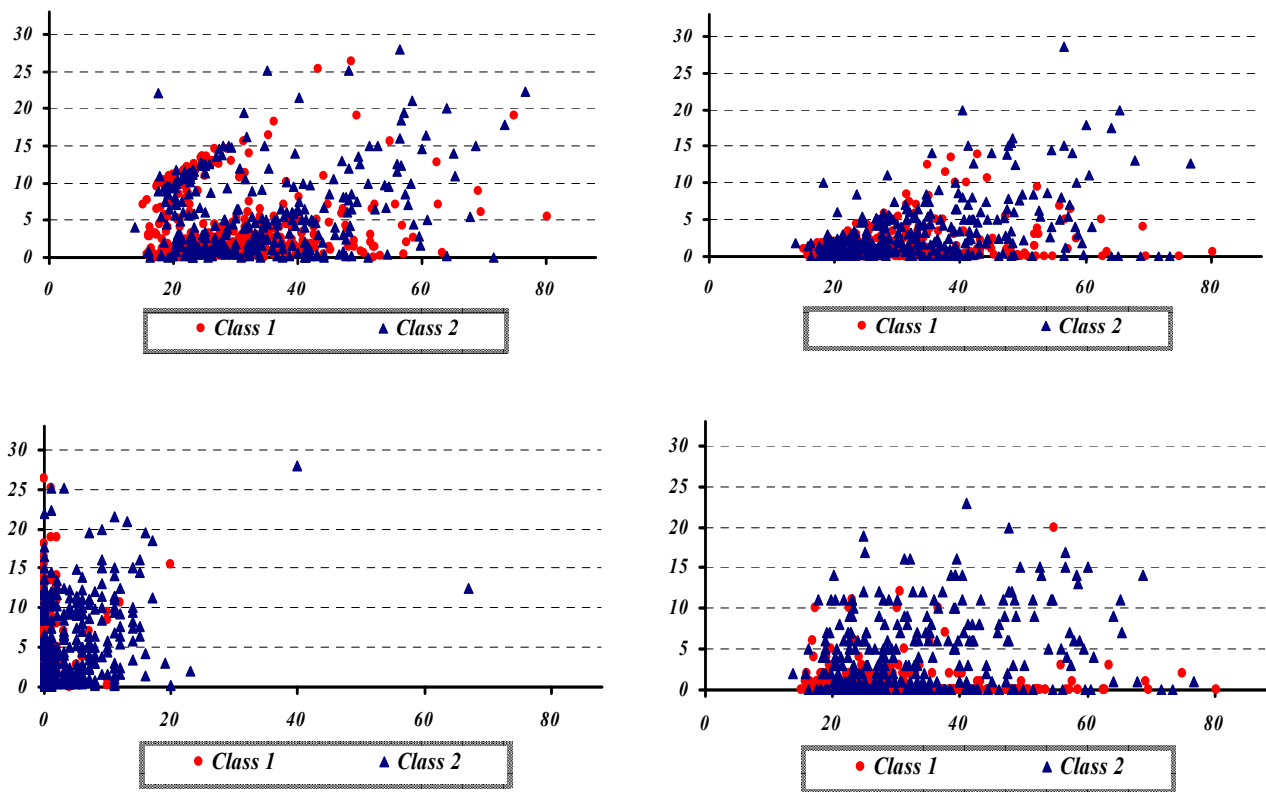
In this section, in order to show the appropriateness and effectiveness of the proposed model for two-class real-life financial data classification, the Australian credit data set has been used. This data set is very interesting for financial researchers because there is a good mix of attributes, including continuous, nominal with small numbers of values, and nominal with larger numbers of values. A brief description of the data set is presented below.

#### *Australian Credit Data Set*

The Australian credit data set contains 690 observations with fourteen attributes in total. Eight of these attributes are discrete, with two to fourteen values, and six of them are continuous attributes. There are 307 positive instances (approximately 44.5%) and 383 negative instances (approximately 55.5%) in this data set. All attribute names and values have been changed to meaningless symbols in order to protect the confidentiality of the data. The data set is summarized in Table 1. In this paper, this data set is randomly divided into training and test data, in which 50% were training and 50% were test examples. The two-dimensional distribution of these two classes against the (A2, A3), (A2, A7), (A3, A10), and (A2, A10), as an example, is shown in Figure 1.

**Table 1.** Australian credit data set.

Attributes	Type	Values	Values (Formerly)
Attribute 1	Discrete	0,1	a,b
Attribute 2	Continuous	13.75–80.25	13.75–80.25
Attribute 3	Continuous	0–28	0–28
Attribute 4	Discrete	1,2,3	p,g,gg
Attribute 5	Discrete	1,2,3,...,14	ff,d,i,k,j,aa,m,c,w,e,q,r,cc,x
Attribute 6	Discrete	1,2,3,...,9	ff,dd,j,bb,v,n,o,h,z
Attribute 7	Continuous	0–28.5	0–28.5
Attribute 8	Discrete	0,1	t,f
Attribute 9	Discrete	0,1	t,f
Attribute 10	Continuous	0–67	0–67
Attribute 11	Discrete	0,1	t,f
Attribute 12	Discrete	1,2,3	s,g,p
Attribute 13	Continuous	0–2000	0–2000
Attribute 14	Continuous	0–100,000	0–100,000
Class	Discrete	0,1	–,+



**Figure 1.** The two-dimensional distribution of the Australian credit cards classes.

#### 4. Application the Proposed Hybrid Model to Australian Credit Scoring

In this section, the procedure of the hybrid proposed model for the Australian credit data set is illustrated. Therefore, in the first stage, in order to obtain the optimum network architecture, based on the concepts of multi-layer perceptrons design [41] and using pruning algorithms in the MATLAB 7 package software, different network architectures are evaluated to compare performance. The best fitted network is selected, and, therefore, the architecture that presented the best accuracy with the test data, is composed of fourteen inputs, fourteen hidden, and one output neuron (in abbreviated form,  $N^{(14-14-1)}$ ). Then, in the second stage, the minimal fuzziness of the fuzzy parameters is determined using Equation (5) with  $h = 0$ . As mentioned previously, the  $h$ -level value influences the width of the fuzzy parameters. In this case, we consider  $h = 0$  in order to yield parameters with a minimum width. The misclassification rate of each model and improvement percentages of the proposed model, in comparison with the models in the test data are summarized in Tables 2 and 3, respectively. The improvement percentage of the proposed model in comparison with the other classification models, is calculated as follows:

$$Improvement(\%) = \frac{(Misclassification\ Rate\ of\ desired\ model - Misclassification\ Rate\ of\ proposed\ model)}{Misclassification\ Rate\ of\ desired\ model} \times 100\% \quad (7)$$



**Table 2.** Misclassification rate of classification models for Australian credit data set.

Model	Classification Error (%)
	Test Data
Linear Discriminant Analysis (LDA)	14.0
Quadratic Discriminant Analysis (QDA)	19.9
K-Nearest Neighbor (KNN)	14.2
Support Vector Machines (SVM)	22.5
Artificial Neural Networks (ANN)	12.3
Proposed Hybrid Model	10.9

**Table 3.** Improvement of the proposed model in comparison with those of other classification models.

Model	Improvement (%)
	Test Data
Linear Discriminant Analysis (LDA)	22.14
Quadratic Discriminant Analysis (QDA)	45.23
K-Nearest Neighbor (KNN)	23.24
Support Vector Machines (SVM)	51.56
Artificial Neural Networks (ANN)	11.38

### Comparison with Other Classifiers

According to the obtained results (Tables 2 and 3), our proposed model has the lowest error on the test portion of the data set in comparison with the other models used for the Australian credit data set, with a misclassification rate of 10.9%. Several different architectures of artificial neural network are designed and examined. The best performing architecture for a traditional multi-layer perceptron (MLP) produces a 12.3% error rate, which, in the proposed model, improves by 11.38%. The linear discriminant analysis (LDA) has the second best performance. The linear discriminant analysis model produces an error rate of 14.0%. The proposed model improves upon this by 22.14% for linear discriminant analysis. As K-nearest neighbor (KNN) scores can be sensitive to the relative magnitude of different attributes, all attributes are scaled by their z-scores before using the K-nearest neighbor model. The best K-nearest neighbor model has an error rate of 14.2%, which is 23.24% higher than the proposed model error. Quadratic discriminant analysis (QDA) has an error rate of 19.9%, which is 45.23% higher than the proposed model. Support vector machine (SVM) with  $C = 0$  misclassifies 22.5% of the test samples, which is 51.56% worse than the proposed model.

## 5. Conclusions

In recent years, credit risk analysis has been an active research area. Risk management and credit scoring is one of the key analytical techniques in credit risk evaluation. Therefore, generating both accurate, as well as explanatory, classification models is becoming increasingly important. In this paper, a soft version of traditional multi-layer perceptrons is proposed as an alternative classification model, using the unique soft computing advantages of fuzzy logic. In the proposed model, instead of crisp weights and biases, fuzzy numbers are used in multi-layer perceptrons for better modeling of the

uncertainties in financial markets. The proposed model, in contrast to linear and quadratic discriminant analyses, does not assume the shape of the partition or the relation between dependent and independent variables. This model does not require storage of training data, unlike the K-nearest neighborhood classifier. Support vector machines require setting a penalty parameter and selecting the kernel function, whereas this classifier, without parameter setting, leans on the training process in order to identify the final classifier. Finally, it does not need a large amount of data to yield accurate results, as with traditional multi-layer perceptrons. In order to indicate the appropriateness and effectiveness of the proposed model, five well-known statistical and intelligent classification models are used for credit scoring classifications. The obtained results show that the proposed model is superior to all alternative models.

### Author Contributions

Both authors contributed equally to all aspects of this paper.

### Conflicts of Interest

The authors declare no conflict of interest.

### References

1. Lewis, E.M. *An Introduction to Credit Scoring*; Athena Press: San Rafael, CA, USA, 1992.
2. West, D. Neural network credit scoring models. *Comput. Oper. Res.* **2000**, *27*, 1131–1152.
3. Lee, K.; Booth, D.; Alam, P. A comparison of supervised and unsupervised neural networks in predicting bankruptcy of Korean firms. *Expert Syst. Appl.* **2005**, *29*, 1–16.
4. Durand, D. *Risk Elements in Consumer Instalment Financing*; National Bureau of Economic Research: New York, NY, USA, 1941.
5. Harris, T. Credit scoring using the clustered support vector machine. *Expert Syst. Appl.* **2015**, *42*, 741–750.
6. Abdou, H.A.; Pointon, J. Credit Scoring, Statistical Techniques and Evaluation Criteria: A Review of the Literature. *Intell. Syst. Account. Financ. Manag.* **2011**, *18*, 59–88.
7. Lando, D. *Credit risk modeling: Theory and applications*, Princeton Series in Finance; Princeton University Press: Princeton, NJ, USA, 2004.
8. Thomas, L.C.; Edelman, D.B.; Crook, J.N. *Credit Scoring and its Applications*, SIAM Monographs on Mathematical Modeling and Computation; SIAM: Philadelphia, PA, USA, 2002.
9. Bolton, C. Logistic Regression and its Application in Credit Scoring. MSc Dissertation, University of Pretoria, Pretoria, South Africa, 2009.
10. Wiginton, J.C. A note on the comparison of logit and discriminant models of consumer credit behaviour. *J. Financ. Quant. Anal.* **1980**, *15*, 757–770.
11. Lee, G.; Sung, T.K.; Chang, N. Dynamics of modeling in data mining: Interpretive approach to bankruptcy prediction. *J. Manag. Inform. Syst.* **1999**, *16*, 63–85.
12. Lee, T.S.; Chen, I.F. A two-stage hybrid credit scoring model using artificial neural networks and multivariate adaptive regression splines. *Expert Syst. Appl.* **2005**, *28*, 743–752.

13. Bellotti, T.; Crook, J. Support vector machines for credit scoring and discovery of significant features. *Expert Syst. Appl.* **2008**, *36*, 3302–3308.
14. Grablowsky, B.J.; Talley, W.K. Probit and discriminant functions for classifying credit applicants: A comparison. *J. Econ. Bus.* **1981**, *33*, 254–261.
15. Henley, W.E.; Hand, D.J. Nearest neighbor analysis in credit scoring. *Statistician* **1996**, *45*, 77–95.
16. Baesens, B.; Egmont-Petersen, M.; Castelo, R.; Vanthienen, J. Learning Bayesian network classifiers for credit scoring using Markov chain Monte Carlo search. In Proceedings of the 16th International Conference on Pattern Recognition (ICPR'02), Quebec, Canada, 2002; Volume 3, p. 30049.
17. Pavlenko, T.; Chernyak, O. Credit risk modeling using Bayesian networks. *Int. J. Intell. Syst.* **2010**, *25*, 326–344.
18. Paliwal, M.; Kumar, U.A. Neural networks and statistical techniques: A review of applications. *Expert Syst. Appl.* **2009**, *36*, 2–17.
19. Khashei, M.; Bijari, M.; Raissi, G.A. Improvement of auto-regressive integrated moving average models using fuzzy logic and artificial neural networks (ANNs). *Neurocomputing* **2009**, *72*, 956–967.
20. Celikyilmaz, A.; Turksen, I.B. Uncertainty modeling of improved fuzzy functions with evolutionary systems. *IEEE Trans. Syst. Man Cybern.* **2008**, *38*, 1098–1110.
21. Khashei, M.; Bijari, M.; Hejazi, S.R. Combining seasonal ARIMA models with computational intelligence techniques for time series forecasting. *Soft Comput.* **2012**, *16*, 1091–1105.
22. Khashei, M.; Hamadani, A.Z.; Bijari, M. A novel hybrid classification model of artificial neural networks and multiple linear regression models. *Expert Syst. Appl.* **2012**, *39*, 2606–2620.
23. Lee, T.S.; Chiu, C.C.; Lu, C.J.; Chen, I.F. Credit scoring using the hybrid neural discriminant technique. *Expert Syst. Appl.* **2002**, *23*, 245–254.
24. Hsieh, N.-C. Hybrid mining approach in the design of credit scoring models. *Expert Syst. Appl.* **2005**, *28*, 655–665.
25. Luo, S.-T.; Cheng, B.-W.; Hsieh, C.-H. Prediction model building with clustering-launched classification and support vector machines in credit scoring. *Expert Syst. Appl.* **2009**, *36*, 7562–7566.
26. Hung, C.L.; Chen, M.C.; Wang, C.J. Credit scoring with a data mining approach based on support vector machines. *Expert Syst. Appl.* **2007**, *33*, 847–856.
27. Li, J.; Wei, L.; Li, G.; Xu, W. An evolution strategy-based multiple kernels multi-criteria programming approach: The case of credit decision making. *Decis. Support Syst.* **2011**, *51*, 292–298.
28. Chen, F.-L.; Li, F.-C. Combination of feature selection approaches with SVM in credit scoring. *Expert Syst. Appl.* **2010**, *37*, 4902–4909.
29. Bijak, K.; Thomas, L.C. Does segmentation always improve model performance in credit scoring. *Expert Syst. Appl.* **2012**, *39*, 2433–2442.
30. Chi, B.-W.; Hsu, C.-C. A hybrid approach to integrate genetic algorithm into dual scoring model in enhancing the performance of credit scoring model. *Expert Syst. Appl.* **2012**, *39*, 2650–2661.
31. Ping, Y.; Yongheng, L. Neighborhood rough set and SVM based hybrid credit scoring classifier. *Expert Syst. Appl.* **2011**, *38*, 11300–11304.
32. Kim, K.S.; Han, I. The cluster-indexing method for case based reasoning using self-organizing maps and learning vector quantization for bond rating cases. *Expert Syst. Appl.* **2001**, *21*, 147–156.

33. Park, C.; Han, I. A case-based reasoning with the feature weights derived by analytic hierarchy process for bankruptcy prediction. *Expert Syst. Appl.* **2002**, *23*, 255–264.
34. Ahn, H.; Kim, K.-J. Bankruptcy prediction modeling with hybrid case-based reasoning and genetic algorithms approach. *Appl. Soft Comput.* **2009**, *9*, 599–607.
35. Capotorti, A.; Barbanera, E. Credit scoring analysis using a fuzzy probabilistic rough set model. *Comput. Stat. Data Anal.* **2012**, *56*, 981–994.
36. Akkoc, S. 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. Oper. Res.* **2012**, *222*, 168–178.
37. Laha, A. Building contextual classifiers by integrating fuzzy rule based classification technique and k-NN method for credit scoring. *Adv. Eng. Inform.* **2007**, *21*, 281–291.
38. Yao, P. Hybrid fuzzy SVM model using CART and MARS for credit scoring. In Proceedings of the International Conference on Intelligent Human-Machine Systems and Cybernetics, IHMSC'09, Hangzhou, China, 26–27 August 2009; pp. 392–395.
39. Khashei, M.; Bijari, M. An artificial neural network (p, d, q) model for time series forecasting. *Expert Syst. Appl.* **2010**, *37*, 479–489.
40. Khashei, M.; Hejazi, S.R.; Bijari, M. A new hybrid artificial neural networks and fuzzy regression model for time series forecasting. *Fuzzy Sets Syst.* **2008**, *159*, 769–786.
41. Khashei, M. Forecasting the Isfahan Steel Company Production Price in Tehran Metals Exchange Using Artificial Neural Networks (ANNs). Master of Science Thesis, Isfahan University of Technology, Isfahan, Iran, 2005.