Credit Scoring Using Data Mining Techniques with Particular Reference to Sudanese Banks

Eiman kambal, Izzeldin Osman
Sudan University for Science and Technology
Khartoum, Sudan
eimankambal@uofk.edu, izzeldin@acm.org

Methag Taha, Noon Mohammed, Sara Mohammed
University of Khartoum
Khartoum, Sudan
methag@live.com, zozo_noon@hotmail.com,
Sara10301@hotmail.com

Abstract— **One of the key success factors of lending organizations in general and banks in particular is the assessment of borrower credit worthiness in advance during the credit evaluation process. Credit scoring models have been applied by many researchers to improve the process of assessing credit worthiness by differentiating between prospective loans on the basis of the likelihood of repayment. Thus, credit scoring is a very typical Data Mining (DM) classification problem. Many traditional statistical and modern computational intelligence techniques have been presented in the literature to tackle this problem. The main objective of this paper is to describe an experiment of building suitable Credit Scoring Models (CSMs) for the Sudanese banks. Two commonly discussed data mining classification techniques are chosen in this paper namely: Decision Tree (DT) and Artificial Neural Networks (ANN). In addition Genetic Algorithms (GA) and Principal Component Analysis (PCA) are also applied as feature selection techniques. In addition to a Sudanese credit dataset, German credit dataset is also used to evaluate these techniques. The results reveal that ANN models outperform DT models in most cases. Using GA as a feature selection is more effective than PCA technique. The highest accuracy of German data set (80.67%) and Sudanese credit scoring models (69.74%) are achieved by a hybrid GA-ANN model. Although DT and its hybrid models (PCA-DT, GA-DT) are outperformed by ANN and its hybrid models (PCA-ANN, GA-ANN) in most cases, they produced interpretable loan granting decisions.**

Index Terms—**Data mining, credit scoring, artificial neural network, decision tree, principal component analysis, genetic algorithms**.

I. INTRODUCTION

Banks and other lending organizations are ''profit-seeking'' to make money for their shareholders [1]. These organizations provide financial products and services to clients while managing a diversity of risks [1]. Although it is risky, loan granting is one of the main services which still constitutes the core of the income to commercial banks and other lending organizations [2]. Credit risk appears when wrong decisions associated with the approval of the loan application are taken. The wrong credit risk assessment leads to increase in the number of defaulters and as a consequence could derive banks towards bankruptcy. An evidence of the potential impact of credit risk decisions socially and economically both locally and globally is the U.S. Sub-prime mortgage crisis (2007-2008) [3]. The consequence of this crisis is the incidence of the global financial crisis [3].

Credit risk evaluation decisions are complex and unstructured problems which cannot be readily solved formally. So credit risk is one of challenges to which financial institutions are exposed. In the past credit risk was evaluated through judgmental and subjective decisions which were issued by loan officers using their experience and analysis of data. Hence, loan granting decisions were inaccurate, more subjective and time consuming [4].

The productive direction is to automate credit risk evaluation process through formal evaluation methods which are called credit scoring. These methods have been developed by banks and researchers to classify customers according to their different risk levels based on the available credit history information. Therefore, credit scoring can be modeled as a data mining (DM) classification problem. A wide range of DM classification techniques such as Linear Discriminant Analysis(LDA), Logistic Regression(LR), Decision Trees (DT), Artificial Neural Network(ANN), Support Vector Machine(SVM)…. etc. have been used to develop credit scoring models[4-6].

This paper describes an experiment of building suitable credit scoring models (CSMs) for the Sudanese banks. This paper firstly briefly surveys the DM techniques that are most widely used for credit scoring. Two commonly discussed DM classification techniques are chosen in this experiment namely: DT and ANN. In addition Genetic Algorithms (GA) and Principal Component Analysis (PCA) are also applied as feature selection techniques. Six CSMs are developed in this paper: ANN, PCA-ANN, GA-ANN, DT, PCA-DT, and GA-DT. These models are compared in terms of accuracy, precision, Type I and Type Π errors, computational time, and transparency. The proposed models were implemented using Agricultural Bank of Sudan credit data set. The German credit data set from UCI Machine Learning Repository was used as a benchmarking credit data set [7].

The structure of the paper is organized as follows: in Section 2 a brief survey of DM techniques which are used for credit scoring is presented. Section 3 describes the proposed credit scoring models; showing the details about experiments' data sets, preprocessing of input data, and the modeling process. In Section 4 the results of training and testing credit scoring models are presented; and comparisons between the evaluation of the results are provided. Finally, Section 5 concludes this work and suggests future directions.

II. Data mining techniques in credit scoring

In this paper these techniques are categorized as Statistical, Artificial intelligence (AI), and Hybrid.

A. *Statistical Approach*

This subsection surveys two statistical parametric techniques and one non-parametric, namely: LDA, LR, and DT.

1) Linear Discriminant Analysis

LDA is a simple statistical model which was utilized by Durand [8] to develop the first credit scoring model. In addition, Baesens et al., and West [5, 9] proposed LDA in building a credit scoring model. In their studies LDA performed well in many cases when it was compared with other techniques such as LR, K-nearest neighbor(KNN), DT, SVM, and ANN. LDA is still one of the most commonly used techniques in developing credit scoring models[4]. However, LDA sometimes suffers from lack of accuracy due to the presumptions of linear relationship between response and independent variables, and the equality of covariance matrices of the good and bad credit classes[9].

2) Logistic Regression

In contrast to LDA, LR model does not require the assumptions of LDA [10]. LR models have been widely used for developing credit scoring models [5, 9, 11, 12]. In these studies the LR credit scoring models achieved better in terms of accuracy when LR was compared with other models such as LDA, ANN, KNN, and DTs. LR was identified as a good alternative to ANN. However, it has weakness due to the model assumption, that independent variables must be linearly related to the logit of the dependent variable [10].

3) Decision Tree

Construction of DT is very easy and does not require any domain knowledge or parameter setting [13]. Moreover, DTs do not require the assumption about probability distribution of response variable and are also applicable whatever the nature of response and explanatory variables [13]. For these advantages DTs have been applied successfully for credit scoring applications in a number of studies [11, 14-18]. In spite of the greater flexibility of DTs, they have the disadvantages of: greater demand for computational resources and structure of DT depends on the observed data; a small change alters the structure of tree [13]; and the accuracy of DTs is not stable and is easily affected by noisy data and the redundancy of the data attributes [19]. Therefore, some researchers addressed these problems in [19, 20] studies, where hybrid credit scoring models were developed by combining DT with rough set (RS), GA and K-means to reduce the number of features and remove noise from data. These hybrid models outperform the single DT models.

B. *Artificial Intelligence Approach*

A wide range of AI techniques provide the best alternative for conventional statistical techniques. Techniques, such as ANN, SVM, and Evolutionary computational techniques are all widely used in building credit scoring models.

1) Artificial Neural Network

Credit scoring is the most famous financial application in which ANN has been used [5, 12, 15, 17, 21-25]. In most studies, researchers compare ANN with traditional statistical methods such as DA, LR, Probit regression, Naive Bayes (NB), Classification and Regression Tree (CART), and KNN. ANN achieved better performance than these techniques, so it is considered to be the proper alternative to these conventional techniques in credit scoring [4, 26]. Despite the high classification rate of ANN, it has been criticized in the development of CSM, for its: poor performance in case of irrelevant and large number of attributes [27]. Oreski et al. [28] addressed this problem by applying GA as the feature selection technique and compared it with other feature selection techniques such as forward selection, information gain, ….etc. These experiments concluded that GA when applied to ANN is significantly better than other techniques. The only shortcoming of GA-ANN is that it takes long runtime. ANNs have been also criticized for their poor explanation capability [6]. Enhancement of the transparency of neural networks acts as one of the success factors for ANN in developing CSMs. Therefore, Akkoç [29] developed interpretable CSM by hybridization of ANN with fuzzy logic. Furthermore Baesens et al. [30] provided another treatment of the transparency problem by using neural network rule extraction techniques.

2) Support Vector Machine

As a promising recent competitor SVMs have been successfully applied to credit scoring problem. Baesens et al. and others [5, 24, 31, 32] applied SVM to construct credit scoring model and compare it with other classifications techniques in terms of accuracy, misclassification, computational cost, and other evaluation criteria. SVM classifier yields good results in most studies in terms of specific criteria. Despite the successful results achieved by this technique, it still suffers from two major drawbacks: sensitivity toward outliers and noisy data[33], and its black–box nature[34]. Fuzzy SVM model (FSVM) was proposed by Tang et al. [33] to deal with outliers and noisy data problem. On the other hand Martens et al. [34] attempted to open the black box of SVM in the application of credit scoring by using SVM rule extraction techniques and suitable ANN rule extraction techniques.

3) Evolutionary Computational Techniques

Two evolutionary computational techniques namely GA and Genetic Programming (GP) have been used in credit scoring. GA credit scoring outperformed LDA, LR and a variety of ANN in terms of classification accuracy [35]. In addition, GA can be combined with other classification techniques to enhance their accuracy. For example Desai and

Overstreet [36] combined GA with ANN as a parameter setting technique. Lahsasna et al. [37] combined GA with fuzzy logic to extract optimal fuzzy rules. One drawback of GA is the considerable computational cost and the lack of comprehensibility [6].

GP credit scoring achieved better results in terms of accuracy when it was compared with other methods such as ANN, LR, probit analysis, etc. [27, 38, 39].

However, the success of evolutionary computation techniques in credit scoring systems have often been criticized because they are considered as black-box techniques whose resulting decisions are not easily interpretable for the financial and business analysts[40].

C. Hybrid approach in credit scoring models

Both statistical techniques and AI techniques have been explored for credit scoring, but there are no reliable conclusions indicating which techniques are better. Recently, there has been a growing interest that existing applications of single AI technique can be further improved by two approaches of hybridization [41], these are:

1. Tightly coupled systems (hybrid systems): A hybrid system would consist of a combination of two or more simple techniques. The system can overcome the limitations of one technique and gain the advantages of the others. In a hybrid credit scoring model, each simple technique has a specific role such as selection of features, classification, optimization of parameter setting, detection of outliers and noisy data, enhancement of transparency, ….. etc. Neuro-fuzzy [29], Genetic fuzzy [37], GA-ANN [28], GA-SVM [42], DT-RS [20] are examples of these hybrid systems. They were suggested by many researchers to enhance the performance of credit scoring systems.

2. Loosely coupled (Ensemble methods): these methods use multiple learners to solve the same problem where each learner can be identified as a separate unit. Many studies have shown that such ensemble methods performed better than single AI techniques for credit scoring [43-45]. In spite of superiority of ensemble credit scoring models in terms of accuracy when compared with single classifier models, they suffer from many drawbacks [46] such as: increased storage; increased computation to classify a new object; and most importantly the lack of transparency which may lead to limiting the usage of ensemble learning methods in credit scoring. Therefore, Wall et al. [47] employed rule extraction to enhance interpretability of ensemble learner.

III. Proposed Credit Scoring Models

This section explains the process of constructing credit scoring models.

A. Software Package

The main software package used in our experiments is RapidMiner 5.3.007. It is an open-source Java-based data mining software. This software can be free downloaded and installed from Rapid home page http://rapid-i.com.

B. Data sets

In this research two credit datasets were used to evaluate the performance of the different classification techniques. The first one was German credit dataset from the UCI Repository of Machine Learning Databases and the second was Sudanese credit dataset which was provided by Agricultural Bank of Sudan.

The German credit dataset contains 1000 instances of loan applications. The original data has a mix of 20 categorical and numerical attributes; recording various financial and demographic information about the applicants. A numeric version of this dataset is also available where the categorical attributes are transformed into numerical ones and a few indicator variables are added, which increases the dimension to 24 input numerical values. The data instances are labeled as classes 1 (good borrower, 700 instances) and 2 (bad borrower, 300 instances) [7].

The Sudanese credit dataset contains 1300 cases, where 720 are classified as non-defaulters and 580 as defaulters. It consists of 17 attributes, which are grouped into three groups: applicant demographic indicators, financial indicators and loan application information .See TABLE I.

TABLE I. ATTRIBUTES OF THE SUDANESE CREDIT DATASET

Attribute#	Attribute Description	Attribute Type
1	Gender	categorical
2	Age	numerical
3	ID Type	categorical
4	Occupation	categorical
5	Have Phone or not	categorical
6	Marital Status	categorical
7	# children	numerical
8	# spouses	numerical
9	Monthly salary	numerical
10	Monthly Expenditures	numerical
11	Finance size	categorical
12	Finance duration	categorical
13	Payment method	numerical
14	Finance form	categorical
15	Loan type	categorical
16	Insurance description	categorical
17	Operational type	categorical

C. Data Preparation

Since a Sudanese credit dataset was not available, constructing of it is the first step in this research. The process of building Sudanese credit dataset consists of the following tasks:

1. Identification and collection of the relevant data (under direction of expertise loan officers).

2. Integration of data (data is scattered into more than one department).
3. Imputation of missing values.
4. Normalization of the numerical attributes using Min-max normalization method.
5. Removing the outliers using RapidMiner outlier detection distance method.
6. Transformation of the categorical attributes to numerical (to gain numerical version of a data set).
7. Storing Data into Excel format and export it to RapidMiner data repository

D. Data Modeling

In this research two experiments to build credit scoring models were conducted. In these experiments two classifiers were employed. The first one is the multi-layer perceptron (MLP) neural network (with one hidden layer consisting of three neurons). The second classifier chosen is the widely recognized decision tree algorithm CART [14], which accepts both categorical and numerical attributes in constructing the decision tree. The criteria used for CART model to select the optimal attribute for splitting was Gini-index [48]. Since feature selection plays an important role in classification, the two feature selection techniques namely PCA[49], and GA[49] were used in these experiments.

Holdout and 10-fold cross-validation were used in experiment1 and experment2 respectively. These techniques are commonly used for assessing classifier accuracy, based on randomly sampled partitions of the given data set [48].

IV. Experimental Results

A. Evaluation Criteria

Five measures were used in this research to evaluate the performance of credit scoring models in the two experiments; accuracy, precision, Type I and Type Π errors, computational time, and transparency.

Accuracy of the classification algorithms evaluates how accurately an algorithm will classify future data, that is, data on which the algorithm has not been trained. Precision (of the class defaulter) is the percentage of tuples classified as "defaulter" that are actually defaulter tuples. Precision (of the class non-defaulter) is the percentage of tuples classified as "non-defaulter" that are actually non-defaulter tuples. TypeⅠerror (the rate of classifying customers as "non-defaulters" when they are defaulters). TypeⅡ error (the rate of classifying customers as "defaulters" when they are non-defaulters). These measures are expressed as [48] :

Accuracy = *Tpos+Tneg/ (pos + neg) ;*

Type I error rate= *Fneg /(pos + neg) ;*

Type II error rate=*Fpos / (pos + neg);*

Precision (*pos*) = *Tpos/ (Tpos+ Fpos);*

Precision (*neg*) = *Tneg/ (Tneg+ Fneg).*

Where:

pos**:** The number of positive ("*defaulter*") samples,

Tpos**:** The number of true positives ("*defaulter*" customers that were correctly classified as such)

Fpos: The number of false positives ("*non-defaulters"* were incorrectly labeled as ("*defaulters*")

neg: The number of negative ("*non-defaulter*") samples,

Tneg : The number of true negatives ("*non-defaulter* " customers that were correctly classified as such)

Fneg: The number of false negatives ("*defaulter*" customers that were incorrectly labeled as "*non-defaulter*").

Computational time: Run time taken for training and testing the model. Transparency (Interpretability): A subjective criterion which refers to the level of understanding and insight that is provided by the classifier.

B. Results and Discussion

1) Experiment 1

Two tables (TABLE Ⅱ and TABLE Ⅲ) show the results of experiment 1.

Results of Experiment 1 reveal that:

For the German credit dataset (TABLE Ⅱ) and Sudanese credit dataset (TABLE Ⅲ):

- GA-ANN model (with training to testing ratio 70-30) outperforms all others in terms of accuracy and precision (Non-defaulter). But it takes the longer computational time to be trained and tested (511, 413 seconds for German and Sudanese data sets respectively).
- PCA-ANN model (with training to testing ratio 70-30) achieved lower Type Π error rates (2% and 0% for the German and Sudanese data sets respectively)

For the German dataset (TABLE Ⅱ):

- GA-ANN (with training to testing ratio 70-30) yields the lower Type I error rate.
- PCA-ANN (with training to testing ratio 70-30) yields the higher Type I error rate and lower Type Π error rate.

For the Sudanese dataset TABLE Ⅲ:

- PCA-DT (with training to testing ratio 60-40) yields the lower Type I error rate and higher Type Π error rate.

Considering the results of accuracy, precision (Non-defaulter), and Type I error rate:

Experiment1 shows that: The best results obtained by the German dataset (80.67%, 85.83%, 11.67%) are better than the best results obtained by the Sudanese dataset (69.74%, 70.63%, 16.35%).

2) Experiment 2 Results

Two tables (TABLE IV and TABLE V) show the results of this experiment.

Results of experiment 2 reveal that:

For the German credit dataset (TABLE IV) and Sudanese credit dataset (TABLE V):

TABLE II. HOLDOUT RESULTS OF SCORING MODELS FOR THE GERMAN CREDIT DATASET

Technique		ANN		PCA-ANN		GA-ANN		DT		PCA-DT		GA-DT	
		German credit data set											
Training (%)-testing (%)		70-30	60-40	70-30	60-40	70-30	60-40	70-30	60-40	70-30	60-40	70-30	60-40
Accuracy (%)		69.00	71.75	65.33	66.50	**80.67**	77.50	67.00	70.25	63.67	60.75	75.67	74.75
Precision%	Non-defaulter	69.08	72.30	65.61	68.33	**85.83**	80.94	71.17	76.16	69.63	67.30	83.33	78.14
	Defaulter	**68.42**	**68.42**	60.00	50.00	56.60	63.75	55.13	56.30	48.84	36.47	55.95	57.58
Type I error (%)		27.00	23.75	32.67	28.5	**11.67**	15.25	21.33	16.75	21.67	25.75	12.00	18.25
Type II error (%)		4.00	4.5	**2.00**	5.00	7.66	7.25	11.67	13.00	14.66	13.5	12.33	7.00
Computational Time in seconds		10	10	4	3	511	439	6	5	1	1	429	339

TABLE III. HOLDOUT RESULTS OF SCORING MODELS FOR THE SUDANESE CREDIT DATASET

Technique		ANN		PCA-ANN		GA-ANN		DT		PCA-DT		GA-DT	
		Sudanese credit data set											
Training (%)-testing (%)		70-30	60-40	70-30	60-40	70-30	60-40	70-30	60-40	70-30	60-40	70-30	60-40
Accuracy (%)		65.13	62.12	58.46	57.12	**69.74**	68.85	52.31	57.88	52.56	55.58	65.64	66.15
Precision %	Non-defaulter	67.42	64.45	58.46	57.12	**70.63**	68.24	56.22	60.74	56.82	62.72	67.92	67.06
	Defaulter	60.32	57.47	Unknown	Unknown	68.12	**70.50**	46.50	53.09	47.06	50.00	60.80	66.28
Type I error (%)		22.05	23.65	41.54	42.88	18.98	23.27	26.15	24.62	24.36	**16.35**	21.79	21.54
Type II error (%)		12.82	14.23	**0.00**	**0.00**	11.28	7.88	21.54	17.50	23.08	28.07	12.57	12.31
Computational Time in seconds		7	6	2	2	413	407	4	4	1	1	47	57

- GA-ANN model outperforms all others in terms of accuracy (75% and 66.54% for German and Sudanese data sets respectively).
- GA-DT and GA-ANN achieved higher precision percentages of non-defaulter and defaulter correspondingly.
- PCA-DT takes shorter computational time.

For the German dataset (TABLE IV):

- DT and PCA-ANN yield lower Type I and Type Π error rates respectively.

For the Sudanese dataset (TABLE V):

- GA-DT and GA-ANN yield lower Type I and Type Π error rates respectively.

Considering the results of accuracy, precision (Non-defaulter), and Type I and Type Π error rates:

Experiment 2 shows that: The better results obtained by the German dataset (75.00%, 78.88%, 15%, 4.3%) are significantly better than the best result of the Sudanese dataset (66.54%, 65.99%, 19.23%, and 8.77%).

3) General remarks for the two experiments

- Cross-validation does not outperform the holdout technique in all cases.
- PCA enhanced the accuracy of DT for only one case in TABLE V (where the accuracy of DT is 55.62% and the accuracy of PCA-DT is 57.38%); while for all other cases, using PCA as a feature selection technique decreased the accuracy of DT and ANN models.
- GA enhanced the accuracy of DT and ANN models for all models in the two experiments.
- For the Sudanese credit dataset GA in addition to enhancing the accuracy of ANN models, it also increased the precision of defaulter and non- defaulter and decreased the rate of Type I and Type Π error rates in the two experiments
- The highest precision (Defaulter) is 69.52% for the Sudanese dataset which is better than the higher precision (Defaulter) for the German dataset (66.89%).

TABLE IV. CROSS-VALIDATION RESULTS OF CREDIT MODELS FOR THE GERMAN CREDIT DATASET

Technique	German credit data set					Computational Time in seconds
	Accuracy (%)	Precision (%)		Type I Error (%)	Type II Error (%)	
		Non-defaulter	Defaulter			
ANN	72.80	78.76	55.47	15.8	11.4	72
PCA-ANN	70.00	71.88	50.00	25.7	**4.3**	21
GA-ANN	**75.00**	76.41	**66.89**	20.1	4.9	2625
DT	70.40	78.69	50.68	**15**	14.6	30
PCA-DT	60.60	72.17	34.84	19.2	20.2	**8**
GA-DT	73.20	**78.88**	56.35	15.8	11	2416

TABLE V. CROSS-VALIDATION RESULTS OF CREDIT SCORING MODELS FOR THE SUDANESE CREDIT DATASET

Technique	Sudanese credit data set					
	Accuracy (%)	Precision (%)		Type I Error (%)	Type II Error (%)	Computational Time in seconds
		Non-defaulter	Defaulter			
ANN	60.77	62.03	58.24	25.38	13.85	38
PCA-ANN	52.54	55.28	44.61	33.23	14.230	13
GA-ANN	**66.54**	65.33	**69.52**	24.69	**8.77**	5234
DT	55.62	58.99	50.30	25.07	19.30	38
PCA-DT	57.38	61.31	52.30	21.84	20.76	**6**
GA-DT	62.69	**65.99**	58.41	**19.23**	18.07	681

V. CONCLUSIONS AND RESEARCH DIRECTIONS

The main goal of this paper is to develop efficient and suitable credit scoring models for the Sudanese credit data set. Like many developing countries, in Sudan credit agencies and credit bureaus do not exist. Financial organizations have not built credit data sets from the performing and non-performing loans in the past. Hence, obtaining a credit data set was a real challenge. So, this research started by constructing credit data-set from credit data provided by Agricultural Bank of Sudan.

This paper firstly surveys the most widely used data mining techniques in credit scoring and presents the pros and cons of each technique. ANN and DT were chosen as the main classifiers in two experiments. In addition PCA and GA were employed as feature selection techniques. Holdout and 10-fold cross-validation techniques were used in experiment1 and experment2 respectively.

The results of these experiments showed that GA-ANN credit scoring model outperformed all other models in terms of accuracy for the two credit datasets in two experiments. The highest accuracy for the Sudanese and German credit scoring models was 69.74% and 80.67% respectively (achieved in experimnt2 with 70%-30%, training to testing ratio). DT and hybrid of DT models achieved less than ANN and hybrid of ANN in most of the criteria, but they have better transparency capability.

Therefore the paper concludes that the capability of credit scoring may be affected by the structure and quality of data and there is no best technique for credit scoring problems for all situations.

The paper recommends that:

i. This is a first attempt to build credit scoring model for Sudanese credit data. Hence, further studies are needed and the credit dataset has to be extended and collected from more than one banking sector.
ii. An attempt for improving the performance of these proposed models have to be conducted using other classification techniques such as SVM, Case-Base reasoning, and Rough set as well as using other alternatives of the hybrid models and ensemble techniques.
iii. Explanation of loan granting decision is important for bankers and consumers. Hence, transparency is of special importance to credit scoring models. Therefore it has to be addressed in future research.

ACKNOWLEDGMENT

Authors would like to thank the Agricultural Bank of Sudan for kindly providing the masked credit data that made this research possible.

REFERENCES

[1] A. B. Emel, M. Oral, A. Reisman, and R. Yolalan, "A credit scoring approach for the commercial banking sector," Socio-Economic Planning Sciences, 2003. 37(2): p. 103-123.

[2] A. Mukherjee, credit scoring model using data mining techniques-a pragmatic approach. 2010, Germany: VDM Verlag Dr. Müller.

[3] A. B. Russell " What Gave Rise to the Credit Crisis, " in UICIFD E-Book2010: http://blogs.law.uiowa.edu/ebook/sites/default/files/Part_5_1.pdf. p. 1-19.

[4] H. Abdou and J. Pointon, " Credit Scoring , Statistical Techniques and Evaluation Criteria : A Review of the Literature, " Intelligent Systems in Accounting, Finance & Management., 2011. 18(2-3): p. 59-88.

[5] B.Baesens, et al. , "Benchmarking state-of-the-art classification algorithms for credit scoring, " Journal of the Operational Research Society, 2003. 54(6): p. 627-635.

[6] A. Lahsasna, R. Ainon, and T. Wah, "Credit scoring models using soft computing methods: a survey, " The International Arab Journal of Information Technology, 2010. 7(2).

[7] A. Asuncion and D.J. Newman, UCI machine learning repository,2007, http://www.ics.uci.edu/~mlearn/MLRepository.html:Irvine, CA:University of California, School of Information and Computer Science .

[8] D. Durand, " Risk Elements in Consumer Installment Financing, " National Bureau of Economic Research, Editor 1941: USA. New York.

[9] D. West, " Neural network credit scoring models, " Computers & Operations Research 2000. 27 p. 1131-1152.

[10] M.-c. Lee, "Enterprise Credit Risk Evaluation models : A Review of Current Research Trends," International Journal of Computer Applications, 2012. 44(11): p. 37-44.

[11] B.Wah, S.Huat, N. Huselina, and M. Husain, " Using data mining to improve assessment of credit worthiness via credit scoring models, " Expert Systems With Applications, 2011. 38(10): p. 13274-13283.

[12] H.C.Koh and W.C. Tan, "A Two-step Method to Construct Credit Scoring Models with Data Mining Techniques, " International Journal of Business and Information, 2006. 1(1): p. 96-118.

[13] P. GIUDICI and S. FIGINI, Applied Data Mining for Business and Industry. Second ed 2009: John Wiley & Sons Ltd.

[14] T.-S. Lee, C.-C. Chiu, Y.-C. Chou, and C.-J. Lu " Mining the customer credit using classification and regression tree and multivariate adaptive regression splines, " Computational Statistics & Data Analysis, 2006. 50(4): p. 1113-1130.

[15] W. Zhong-yin " Credit valuation modeling and evaluation using data mining for consumer loan market with skewed data," Department of Industrial Management2007, National Taiwan University of Science and Technology.

[16] W. Li and J. Liao, "An Empirical Study on Credit Scoring Model for Credit Card by Using Data Mining Technology, " Seventh International Conference on Computational Intelligence and Security2011, IEEE: Sanya, Hainan, China p. 1279-1282.

[17] A. Arzy Soltan and M. Mehrabioun Mohammadi "A hybrid model using decision tree and neural network for credit scoring problem," Management Science Letters, 2012. 2(5): p. 1683-1688.

[18] Y. B. Wah and I. R. Ibrahim "Using Data Mining Predictive Models to Classify Credit Card Applicants, " in 6th International Conference on Advanced Information Management and Service (IMS)2010, IEEE: Seoul, Korea (South). p. 394-398.

[19] D. Zhang, S. C. H. Leung, and Z. Ye " A Decision Tree Scoring Model Based on Genetic Algorithm and K-Means Algorithm," in Third International Conference on Convergence and Hybrid Information Technology(ICCIT)2008, IEEE: Busan, Korea. p. 1043-1047.

[20] Y. Jiang, X. Zhou, and D. Zhang "A New Approach Based on a Rough Set and a Decision Tree to Bank Customer Credit Evaluation, " International Symposium on IT in Medicine and Education2008, IEEE Xiamen, China. p. 61-65.

[21] H. Abdou, J. Pointon, and A. El-Masry " Neural nets versus conventional techniques in credit scoring in Egyptian banking, " Expert Systems with Applications, 2008. 35: p. 1275-1292.

[22] I.-C. Yeh and C.-h. Lien "The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients, " Expert Systems with Applications, 2009. 36(2): p. 2473-2480.

[23] A.Khashman "Neural networks for credit risk evaluation: Investigation of different neural models and learning schemes, " Expert Systems with Applications, 2010. 37(9): p. 6233-6239.

[24] N. Nwulu, S. Oroja, and M. Ilkan "Credit Scoring Using Soft Computing Schemes : A Comparison between Support Vector Machines and Artificial Neural Networks," in nternational Conference, DEIS . 2011. London, UK: Springer-Verlag Berlin Heidelberg.

[25] A. Blanco, R. Pino-Mejías, J. Lara, and S. Rayo, "Credit scoring models for the microfinance industry using neural networks: Evidence from Peru," Expert Systems with Applications, 2013. 40(1): p. 356-364.

[26] A. Lahsasna, R.N. Ainon, and T.Y. Wah, "Credit Scoring Models Using Soft Computing Methods : A Survey," The International Arab Journal of Information Technology, 2010. 7(2).

[27] C. Ong, J. Huang, and G. Tzeng " Building credit scoring models using genetic programming," Expert Systems with Applications, 2005. 29(1): p. 41-47.

[28] S. Oreski, D. Oreski, and G. Oreski, "Hybrid system with genetic algorithm and artificial neural networks and its application to retail credit risk assessment, " Expert Systems with Applications, 2012. 39: p. 12605-12617.

[29] S. Akkoç, "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, " European Journal of Operational Research, 2012. 222(1): p. 168-178.

[30] B. Baesens, R. Setiono, C. Mues, J. Vanthienen, and K. Ridge, "Using Neural Network Rule Extraction and Decision Tables for Credit-Risk Evaluation, " Management Science, 2003. 49(3): p. 312-329.

[31] L.Zhang and X. Hui, "Application of Support Vector Machines, " 6th International Symposium of Neural Networks (ISNN) 2009, Springer-Verlag: China. p. 283-290.

[32] T. Bellotti and J. Crook, "Support vector machines for credit scoring and discovery of significant features," Expert Systems with Applications, 2009. 36(2): p. 3302-3308.

[33] B. Tang and S. Qiu, "A new Credit Scoring Method Based on Improved Fuzzy Support Vector Machine," Computer Science and Automation Engineering (CSAE)2012, IEEE: China. p. 73-75.

[34] J. H. David Martens, Rudy Setiono, Jan Vanthienen, and Bart Baesens, "Rule Extraction from Support Vector Machines: An Overview of Issues and Application in Credit Scoring," An Overview of Issues and Application in Credit Scoring, Studies in Computational Intelligence (SCI)2008, Springer-Verlag Berlin Heidelberg p. 33–63.

[35] C. J. Desai V., and Overstreet G., "Credit Scoring Models in the Credit Union Environment Using Neural Networks and Genetic Algorithms," Computer Journal of Mathematics Applied in Business and Industry, 1997. 8(4): p. 324-346.

[36] A. C. Bahnsen and A. M. Gonzalez, "Evolutionary algorithms for selecting the architecture of a MLP Neural Network: A Credit Scoring Case," in 11th IEEE International Conference on Data Mining Workshops2011: Vancouver.

[37] A. Lahsasna, R. N. Ainon, and T. Y. Wah, "Enhancement of transparency and accuracy of credit scoring models through genetic fuzzy classifier," Maejo International Journal of Science and Technology, vol. 4, pp. 136-158, 2010.

[38] A I Marqués, V García, and J. S. Sánchez, "A literature review on the application of evolutionary computing to credit scoring,"

Journal of the Operational Research Society advance online publication 2012. 7.

[39] H.T. Elshoush and I.M. Osman, "Alert correlation in collaborative intelligent intrusion detection systems—A survey," Applied Soft Computing, 2011. 11: p. 4349-4365.

[40] W. Xu, S. Zhou, D. Duan, and Y. Chen, "A Support Vector Machine Based Method for Credit Risk Assessment," in 7th International Conference on E-Business Engineering2010, IEEE: Shanghai, China. p. 50-55.

[41] G.Wang, J. Hao, J.Ma, and H. Jiang, "A comparative assessment of ensemble learning for credit scoring," Expert Systems with Applications, 2011. 38: p. 223-230.

[42] B.Twala, "Multiple classifier application to credit risk assessment," Expert Systems with Applications, 2010. 37: p. 3326-3336.

[43] H. Abdou, "Genetic programming for credit scoring: The case of Egyptian public sector banks," Expert Systems with Applications 2009. 36: p. 11402–11417.

[44] R. Singh and R. R. Aggarwal, "Comparative Evaluation of Predictive Modeling Techniques on Credit Card Data," International Journal of Computer Theory and Engineering, 2011. 3(5): p. 1-6.

[45] A. Ghodselahi and A. Amirmadhi, "Application of Artificial Intelligence Techniques for Credit Risk Evaluation," International Journal of Modeling and Optimization, 2011. 1(3):p.243-249.

[46] S. B. Kotsiantis, I. D. Zaharakis, and P. E. Pintelas, "Machine learning: a review of classification and combining techniques," Artificial Intelligence Review, 2007. 26: p. 159-190.

[47] R. Wall, P. Cunningham, P. Walsh, and S. Byrne, "Explaining the output of ensembles in medical decision support on a case by case basis," Artificial Intelligence in Medicine, 2003. 28: p. 191-206.

[48] J. Han and M. Kamber, Data Mining:Concepts and Techniques, 2 ed 2006: Morgan Kaufmann.

[49] M. Šušterši, D. Mramor, and J. Zupan, "Consumer Credit Scoring Models With Limited Data," Expert Systems with Applications, 2009. 36(3): p. 4736-4744.