JAIDM
Journal of AI and Data Mining
Vol. 1, No.2, 2013, 119-129.

Credit scoring in banks and financial institutions via data mining techniques: A literature review

S. M. Sadatrasoul[1*], M.R. Gholamian[1], M. Siami[1], Z. Hajimohammadi[2]

1. Department of Industrial engineering, Iran University of Science and technology, Tehran, Iran
2. Department of Computer Science, Amirkabir University of technology, Tehran, Iran

Received 22 November 2012; accepted 11 March 2013
**Corresponding author:sadatrasoul@iust.ac.ir (S. M. Sadatrasoul)*

Abstract
This paper presents a comprehensive review of the studies conducted in the application of data mining techniques focus on credit scoring from 2000 to 2012. Yet, there isn't adequate literature reviews in the field of data mining applications in credit scoring. Using a novel research approach, this paper investigates academic and systematic literature review and includes all of the journals in the Science direct online journal database. The studies are categorized and classified into enterprise, individual and small and midsized (SME) companies credit scoring. Data mining techniques are also categorized to single classifier, Hybrid methods and Ensembles. Variable selection methods are also investigated separately because there is a major issue in a credit scoring problem. The findings of this literature review reveals that data mining techniques are mostly applied to an individual credit score and there is inadequate research on enterprise and SME credit scoring. Also ensemble methods, support vector machines and neural network methods are the most favorite techniques used recently. Hybrid methods are investigated in four categories and two of the frequently used combinations are "classification and classification" and "clustering and classification". This review of literature analysis provides scope for future research and concludes with some helpful suggestions for further research.

Keywords: *Credit scoring, Banks and financial institutions, Literature review, Data mining.*

1. Introduction
Credit scoring consists of the assessment of risk associated with lending to an organization or a consumer (an individual). There are so many papers used intelligent and statistical techniques since the 1930s. In that decade, numerical score cards were first introduced by mail-order companies [1]. It seems that since then, although statistical techniques are used in some papers especially in hybrid techniques which mainly combine different techniques strengths to overcome their weaknesses, the usage of data mining techniques in the area of research has increased and become the dominant area in the field.
When assessing the credit, according to the context we can roughly summarize the different kind of scoring as follows [2]:

- **Application (credit) scoring:** It refers to the assessment of the credit worthiness for new applicants. It quantifies the default, associated with credit requests, by questions in the application form, e.g., present salary, number of dependents, and time at current address. Usually, a credit score is a number that quantifies the creditworthiness of a person;
- **Behavioral scoring:** It involves principles that are similar to application scoring, with the difference that it refers to existing customers. In fact, the decision about that how the lender has to deal with the borrower is in this area. Behavioral scoring models use customer's historical data, e.g., account activity, account balance, frequency about

past due, and age of account to predict the time to default;

- **Collection scoring:** It is used to divide customers with different levels of insolvency into groups, separating those who require more decisive actions from those who don't need to be attended to immediately. These models are distinguished according to the degree of delinquency (early, middle, late recovery) and allow a better management of delinquent customers, from the first signs of delinquency (30–60 days) to subsequent phases and debt write-off;
- **Fraud detection:** fraud scoring models rank the applicants according to the relative likelihood that an application may be fraudulent.

This paper investigates credit scoring problems used data mining techniques. Over the past few years, a number of review articles have appeared in different publications. Hand and Henely reviewed several statistical classification models in consumer credit scoring [3]. They concluded that there is not a best method for scoring and selecting the best method depends on parameters like data structure, and the variables used other contextual characteristics. They concluded that when the data is not structured, it's better to use flexible intelligent methods like neural networks.

Thomas surveys the statistical and operational research techniques used to support credit and behavioral scoring decisions. He also discusses the need for Profit scoring, in terms of the profit, a consumer will bring to the lending organization. He explained that Profit scoring would allow organizations to have a tool that is more aligned to their objective of profitability than the present tools to measure customer's delinquency. The paper concludes that developing more quality information systems credit and behavioral scoring area are going to have more studies in new areas like profit scoring [4].

Kamleitner and Kirchlerpresent a conceptual process model, and stress the character of credit use, and review credit literature with regard to the three major parts of the consumer credit process, which are processes before, processes at, and processes after credit takes up [5]. They conclude their study with nine findings and two major gaps about credit process.

Abdou and Pointon reviewed articles based on credit scoring applications in various areas especially in finance and banking based on statistical techniques [6]. Their study also include some of data mining techniques, and comparison of different techniques accuracy for different UCI datasets, they conclude that there is no overall best statistical technique in building scoring models.

This paper is an up to date review, which is defined in the new area and has new objectives. First, it is to develop a framework for classifying data mining application in the credit scoring and provides a comprehensive review of new articles in the area based on the framework. Second, it is to provide a guideline for new researchers and practitioners in credit scoring area especially for those who want to use data mining techniques. Third, it is to investigate the pre-process and especially variable selection techniques used in the area.

The rest of the paper is organized as follows: Section 2 presents review methodology, section 3 gives the classified articles based on section 2 methodology, in section 4 the discussions are represented and the important insights of the research is analyzed and bolded. Section 5 concludes the research and future directions in the field are suggested.

2. Methodological framework

As there are many previous works in the area of credit scoring, the literature review was based on the descriptor, “credit scoring". Full text of articles reviewed and the ones that were not actually related to the data mining techniques are excluded. Other selection criteria are as follows:

- Only Science direct online journal database were used;
- Only those articles that were in published journals and used the data mining techniques are included;
- Masters and doctorial theses, conference papers, working papers and internal reports, text books are excluded from the review mainly because academics prefer journals to acquire and disseminate information.

Figure 1 shows the methodological framework of the research.

The primary databases have about 110 articles and with further investigations and refining the results 44 articles were remained and other 66 articles were eliminated because they were not related to the application of data mining techniques in credit scoring. Each of the 44remaining articles was studied and reviewed carefully and classified in 5 tables according to their type of study.

3. Classification method

In this section, a graphical conceptual framework shown in Figure 2 is used for classifying credit scoring and data mining techniques. The conceptual framework is designed by literature review of current researches and books in credit scoring area [1]. As shown in Figure 2, the given framework consists of two levels. The first level includes three types of credit scoring problem comprising Enterprise credit score, individual's credit score and small and midsized credit score.

(i) **Individual (consumer) credit score:** The individual credit score, scores people credit using variables like applicant age, marital status, income and some other variables and can include credit bureau variables.

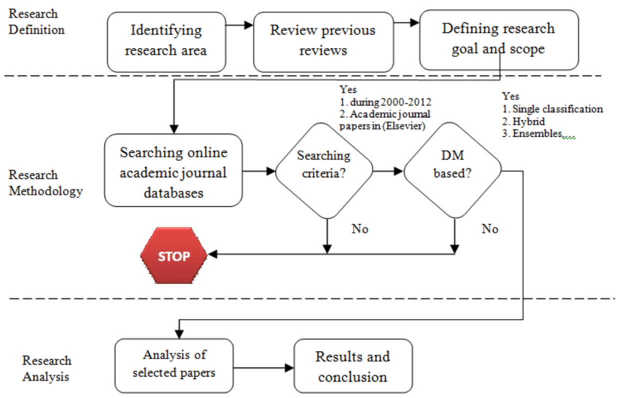

Figure 1. Methodological framework of research

(i) **Enterprise credit score:** using audited financial accounts variables and other internal or external, industrial or credit bureau variables, the enterprise score is extracted.

(ii) **SME credit score:** For SME and especially small companies financial accounts are not reliable and it's up to the owner to withdraw or retain cash, there are also other issues, for example small companies are affected by their partners and their bad/good financial status affects them, so monitoring the SMEs counterparts is another way of scoring them [1]. As a matter of fact, small businesses have a major share of the world economy and their share is growing, so SME scoring is a major issue which is investigated in this paper.

Although some differences can be found for scoring of export guarantees, EXIM banks and other

Figure 2. Classification framework for intelligent techniques in credit scoring

institutions which have not the profit as their main goal, they are excluded because of their low literature [1].

The second layer, comprised from three types of solutions and variable selection, they are presented below.

- **Variable selection:** Selecting appropriate and more predictive variables is fundamental for credit scoring [7]. Variable selection is the process of selecting the best predictive subset of variables from the original set of variables in a dataset [8]. There are many different methods for selecting variables include Stepwise regression, Factor analysis, and partial least square.
- **Single classifier**: Credit scoring is a classification problem and mainly classified applicant to good or bad. There are many data mining techniques for classification including support vector machine, and decision tree.
- **Hybrid approaches:**
  The main idea behind the hybrid approaches is that different methods have different strengths and weaknesses. This notion makes sense when the methods can be combined in some extent. This combination covers the weaknesses of the others. There are four different hybrid methods [9].
  - **Classification + Clustering**
    Clustering is an unsupervised learning technique and it cannot distinguish data accurately like supervised techniques. Therefore, a classifier can be trained first, and its output is used as the input for the cluster to improve the clustering results.

In the case of credit scoring, one can cluster good applicants in different groups.

- **Clustering + Classification**
In this approach, clustering technique is done first in order to detect and filter outlier. Then the remained data, which are not filtered, are used to train the classifier in order to probably improve the classification result.
- **Classification + Classification**
In this approach, the aim of the first classifier is to 'pre-process' the data set for data reduction. That is, the correctly classified data by the first classifier are collected and used to train the second classifier. It is assumed that for a new testing set, the second classifier could provide better classification results than single classifiers trained by the original dataset [9].
- **Clustering + Clustering**
For the combination of two clustering techniques, the first cluster is also used for data reduction. The correctly clustered data by the first cluster are used to train the second cluster. Finally, for a new testing set, it is assumed that the second cluster could provide better results.

- **Ensemble approaches:**

Ensemble methods aggregate the predictions made by multiple classifiers to improve the overall accuracy. They construct a set of classifiers from the training data and predict the classes of test samples by combining the predictions of these classifiers [10]. There are several types of Ensembles include bagging and boosting.

Table 1. Distribution of articles according to the proposed classification model

Credit scoring categories	Data mining application class	Data mining techniques	Prescreening/Variable selection	References
Enterprises	Ensemble	NN cross validation, bagging, and boosting Ensemble strategies compared with multilayer perceptron neural network	-	[11]
		Bagging, Boosting (adaboost), staking ensembles based on Logistic Regression, Decision Tree, Artificial Neural Network and Support Vector Machine compared with each other	-	[12]
		Subagging compared with 5 other methods	Manually based on strong correlation	[13]
Individuals	Single classification	Genetic programming compared with weight of evidence and Probit analysis	-	[14]
		Back-propagation artificial neural network compared with logistic regression	genetic algorithm and principle component analysisfor variable selection	[15]
		Neural networks(multilayer perceptron, mixture-of-experts, radial basis function, learning vector quantization, and fuzzy adaptive resonance) compared to linear discriminate analysis, logistic regression, k nearest neighbor, kernel density estimation, and decision trees	-	[16]
		Probabilistic neural nets and multi-layer feed-forward nets are compared with conventional techniques (discriminant analysis, probit analysis and logistic regression)	-	[17]
		Rule base	-	[18]
		Expert system compared with ٣٧ techniques include intelligent and statistical	-	[19]
		Multi-Layer Perceptrons compared with other 14 methods	principal component analysis (PCA) and different treatment methods of experiences	[20]
		Artificial neural network (RBF) compared with SVM and logistic regression	New feature selection based on rough set and tabu search	[21]
		Two evolutionary rule learners compared with neuro fuzzy classifier, Fisher discriminant analysis, Bayes' classification rule, Artificial neural networks, C4.5 decision trees	-	[22]
		Two staged MARS and NN hybrid compared with discriminant analysis, logistic regression, artificial neural networks and MARS	multivariate adaptive regression splines (MARS)	[23]
		SVM compared with neural networks, genetic programming, and decision tree classifiers	Genetic algorithm	[24]
		SVM compared with Multilayer Perceptrons (MLP)	No variable selection/include data encoding and discritization	[25]
		Genetic programming compared with ANN, decision trees, rough sets, and logistic regression.	No variable selection/ include discritization	[26]
		SVM(RBF Kernel , KGPF Kernel) compare with logistic regression	-	[27]
		SVM compare with logistic regression, discriminant analysis and k-Nearest neighbors	Feature selection by SVM	[28]
		Three link analysis algorithms compared with traditional SVM	SVM prescreening	[29]
		Clustering-launched classification(CLC) compared with SVM , SVM +GA	-	[30]
		SVM grid search compared with CART and MARS	CART and MARS	[31]
		Different feature selection for SVM compare with original SVM	discriminate analysis, decision tree, Roughs set and Fscore	[32]

Credit scoring categories	Data mining application class	Data mining techniques	Prescreening/Variable selection	References
		Random subspace method compared with Bagging, Class Switching, Rotation Forest and stand-alone classifiers	-	[33]
		CART and MARS compared with discriminant analysis, logistic regression, neural networks, and support vector machine	-	[34]
		rule extraction techniques for SVM compared with Trepan, G-REX and three other methods	-	[35]
		Genetic algorithm compared with logistic and linear regression	pre-process categorical and continuous variables to code them as a set of dummy variables	[36]
		using grid search to optimize RBF kernel parameters of SVM compared with linear discriminant analysis, logistic regression and neural networks	neighborhood rough set compared with t_Test, Correlations, Stepwise, CART, MARS, Pawlak's rough set	[37]
		Support vector machine with variable selection compared with genetic programming, neural network, SVM based genetic algorithm	F score	[38]
		Radial bases function with feature selection compared with J48 and logistic regression	based on rough set and scatter search	[39]
		Decision tree chi-square automatic interaction detector (CHAID), compared with logistic regression and weight of evidence and scorecard	Manual data preprocessing and cleaning	[40]
		Random forest and gradient boosting compared with 8 other methods	-	[41]
		Multi layer perceptron and Classification and regression trees compared with discriminant analysis and logistic regression	categorizing the data using dummy variables	[42]
	Classification + Classification	Back propagation neural networks combined with discriminant analysis	Discriminant analysis also works as a variable selection	[43]
		Hybrid neural networks (NNs) and genetic algorithms compared with discriminant analysis and CART	-	[44]
		Two-stage genetic programming compared with other 6 methods	-	[45]
		ANN and case based reasoning(CBR) compared with discriminant analysis, Logistic regression, CART and ANN	MARS	[46]
	Clustering + Classification	Self organizing map and k-means for clustering and neural network for classification	-	[47]
		Self organizing map and fuzzy k-nn rule compare with fuzzy rule base	-	[48]
	Ensemble	Three layer back-propagation neural network single classifier compared with multiple classifier	-	[49]
		Vertical bagging decision trees model(VBDTM) compared with other 10 methods	Rough set	[50]
		Least squares support vector machines (LSSVM) compared with 19 other individual classification models	-	[51]
		Hybrid clustering using Two-step and k-means, Ensembles and association rules	Discretization of continuous values with Optimal associate binning and Rank important features with Pearson chi-square test)	[52]
		Two-bagged and the three-bagged based on decision tree compared with different bagging based on logistic regression	-	[53]
		Random subspace(RS)-Bagging decision tree(DT) and Bagging-RS DT compared with single DT and four other methods	-	[54]
SME	Single classification	Classification and Regression Tree (CART) compared with 5 different variables selected	-	[55]

4. Analysis of credit scoring research based on Classification method

This paper provides a new review of literature on the application of data mining in credit scoring based on Figure 2. The distribution of 44 articles was classified by using the proposed classification method shown in Tables 1-5. The following subsections present analysis of data mining techniques in credit scoring.

4.1. Distribution of articles by data mining application classes

The 44 classified articles and their techniques are analyzed and shown in Table 1. All articles were read carefully and categorized based on the type of credit and type of the main data mining techniques. Also other Techniques which are used as the benchmark are mentioned obviously and separated using "compared with" statement. Any preprocessing or especially variable selection techniques in each article are extracted and determined in Penultimate column. Some articles include both enterprise and individual credit scoring, They are categorized in enterprise level because they have mainly used datasets which haven't seen in previous works, and have more contributions to the knowledge in the field [11]. There are no articles in some categories for example in Enterprise credit scoring using hybrid methods, so no room was specified in this regard.

It can be seen that the most of the publications were in the individual's credit scoring with 40 articles (91%). After that, Enterprise credit scoring had the second step (3 with 7%) and SME credit scoring had the third step (1 with 2%). About 22 articles (50%) used a preprocessing method and 17 articles (39%) use variables selection methods some of them manually and others with known techniques. It is clear that data preprocessing and variable selection is used credit scoring research especially for those who used datasets other than UCI benchmark datasets.

4.2. Articles by their main contribution

Table 2 comprises a complete list of the 44 articles in the review; the "main idea" column of the table shows the main idea and objective of each research.

Table 2. Distribution of articles by their main contribution

Ref #	MAIN IDEA
[11]	Ensembles of NN predictors provide more accurate generalization than a single model.
[12]	Comparative assessment of the performance of three popular ensemble methods (Bagging, Boosting, and Stacking).
[13]	The main objective is to build and validate robust models able to handle missing information, class unbalancedness and non-iid data points.
[14]	Investigate the ability of GP in the analysis of credit scoring models in Egyptian public sector banks.
[15]	Using a new method for variable selection because of high correlation between them and evaluating the results using ANN on the newly introduced data.
[16]	Comparing different neural networks versus traditional commercial techniques.
[17]	To investigate the ability of neural nets and conventional techniques in evaluating credit risk in Egyptian banks.
[18]	Giving a complementary view of redundancy in rule bases based on the contribution of individual rules to the overall system's accuracy.
[19]	Machine learning methods haven't any statistically significant advantage over the expert system's accuracy when problems were treated as a classification.
[20]	Solving the problem of imbalanced class distributionscan lead the algorithms to learn overly complex models and can over fit the data.
[21]	A new feature selection based on rough set and tabu search has been proposed.
[22]	Proposing two evolutionary fuzzy rule learners.
[23]	Introducing a new two-stage hybrid modeling procedure using MSRS and NN.
[24]	Increase SVM accuracy by hybrid method and feature reduction.
[25]	To develop a useful visual decision-support tool Using SVM.
[26]	Proposing genetic programming as a more sophisticated model to significantly improving the accuracy of the credit scoring.
[27]	To present a novel and practical adaptive scoring system based on incremental kernel methods.
[28]	To show that support vector machines are competitive against traditional methods on a large credit card database.
[29]	Three link analysis algorithms based on preprocess of support vector machine proposed to estimate an applicant's credit.
[30]	Using a new classifier named clustering-launched classification (CLC) for credit scoring.
[31]	To show that hybrid SVM has better capability of capturing nonlinear relationship among variables.
[32]	Using different feature selection methods for SVM.
[33]	''Random Subspace" method outperforms the other ensemble methods tested in the paper.
[34]	Explore the performance of credit scoring using two commonly discussed data mining techniques CART and MARS.
[35]	Extracting rules from SVM to overcome its complexity.
[36]	Seeks to determine the impact of in correct problem specification on performance that results from having different objectives for model construction and assessment.
[37]	To constructs a hybrid SVM-based credit scoring models to evaluate the applicant's credit score.
[38]	A new strategy to reduce the computational time for credit scoring using SVM incorporated with F score for feature reduction.
[39]	A novel approach, called RSFS, to feature selection based on rough set and scatter search is proposed.
[40]	Constructions of credit scoring model based on data mining technique and compare it to a scorecard.
[41]	Compare several techniques that can be used in the analysis of imbalanced credit scoring data sets.
[42]	To make a practical contribution in instance sampling to model building on credit scoring datasets.
[43]	Using NN and discriminant analysis Hybrid models to improve the performance.
[44]	Using GA-based inverse classification to conditional acceptance of rejected customers classified sooner with NN.
[45]	An improvement in accuracy might translate into significant savings, so a more sophisticated model based on Two-stage genetic programming is introduced.
[46]	Introduce a reassigning credit scoring model (RCSM) involving two stages to decrease the Type I error.
[47]	Presents a hybrid mining approach in the design of an effective credit scoring model based on clustering and neural network.
[48]	Introduce a ''soft'' classifier to produce a measure of support for the decision that provides the analyst with a greater insight.
[49]	Comparing classifier NN ensembles versus single NN classifiers and best single classifier.
[50]	A novel credit-scoring model called vertical bagging decision trees model (abbreviated to VBDTM) is proposed.
[51]	Several ensemble models based on least squares support vector machines (LSSVM) are used to reduce bias.
[52]	Introducing the concept of class-wise classification as a preprocessing step in order to obtain an efficient ensemble classifier.
[53]	A new bagging-type variant procedure called poly-bagging is proposed.
[54]	Random subspace (RS)-Bagging decision tree (DT) and Bagging-RS DT, to reduce the influences of the noise data and redundant attributes.
[55]	A decision tree-based technology credit scoring introduced for start-ups and SMEs.

4.3. Distribution of articles by data mining techniques

Table 3 shows the distribution of articles by the main data mining techniques used in different credit scoring domains and benchmark techniques used for comparison are excluded. The variable selection techniques are also included in Table3 [32]. Some articles used data mining techniques other than the main issue of classification or clustering in credit scoring, for example [14] Kohnenused map for analysis of the overall sample and tested sub-sample. These techniques used for issues other than classification are excluded because they are not concerned with the main objective of the review. In some articles, different types of techniques are used and discussed all of those different types add a single value to the number of technique used [16,17]. Some articles use meta-heuristics or search algorithms to find or tune data mining algorithms parameters. For example, an article used grid search to optimize model parameters, and these algorithms are also included [31]. Ensembles mainly used one (with different parameter settings) or more classification techniques, and in these situations, the data mining technique is reported only in ensemble raw and techniques behind and the ensembles are not reported and computed [12].

The analysis shown that 23 different techniques are used 79 times and artificial neural networks are mostly used and ranked first (12 with 15.2%). Following techniques are Ensemble methods with 11articles (14%) and support vector machines with 9 articles (11.4%).

Because of robustness, transparency needs and also regulators on the credit scoring in some countries do the auditing process. Banks cannot use many of above mentioned methods [56].By using rule bases, decision trees banks can easily interpret the results and explore the rejecting reasons to the applicant and regulatory auditors. Therefore rule based techniques, and other types of decision tree methods are used in 14 articles (17.7%). This shows that these types of techniques are also one of the favorite techniques in credit scoring problems.17 articles used different variable selection techniques, among them rough sets are the most favorite 5 articles (29.4%) used, and are followed by MARS from which 4 articles (23.5%) used.

A brief description of the three most used techniques are as follows:

Neural networks: Artificial Neural Networks (ANNs) are non-linear techniques that imitate the human brains functionality. They are used broadly in classification, clustering and optimization problems[10]. ANNs are able to recognize the complex and non-linear patterns between input and output variables in credit scoring which then predict the creditworthiness of a new applicant. They can also use for clustering applicants.

Support vector machines: SVM is the state-of-the-art technology based on statistical learning, it is designed for binary classification and aims to develop an optimal hyper plain in way that maximizes the margins of separation between the negative and positive data sets [57]. Because in many cases, the used datasets are linearly non-separable, and a non-linear transformation of the data set into a higher dimensional space in done[10]. In the case of credit scoring, SVM is used to classify the applicants usually based on non-linear input variables.

Ensemble methods:

Ensemble methods combines the predictions of different classifiers [10]. An ensemble method can use a unique classifier with different parameters tuned or different classifiers combined. There are several types of Ensembles include bagging, boosting, random forests. In the case of credit scoring, different classifiers classify an applicant and using a voting mechanism the final decision is kept for an applicant.

4.4. Distribution of articles by journal

Table 5 shows distribution of articles by journal. Articles related to credit scoring publications are from 10 different journals. Most of the publications are dedicated to the “Expert system with applications” journal (32 with 72.7%).European Journal of Operational Research and Computers and Operations Research are followed (6 with 13.5%totally).

5. Conclusion and future directions

Application of data mining techniques is an emerging and growing trend in credit scoring. This paper gathered and analyzed 44 articles, which applied data mining techniques to credit scoring between 2000 and 2012. The aim of this paper is to develop a framework for classifying data mining application in the credit scoring, and provides a guideline for new researchers. Practitioners in credit scoring area especially for those who want to use data mining techniques lastly investigate preprocesses and especially variable selection technique which is used in the area. The findings of the paper are:

- Individuals (consumer) credit scoring has dedicated the most articles from three area of credit scoring research.
- Only one article from Korea focused on SME credit scoring and the reason is that Korean government valued a knowledge-based economy.
- Although there are few literature on SME credit scoring, research on the application of data mining in credit scoring will increase significantly in future in the area of small and midsized companies as they are the companies of future which are more knowledge based.

- The majority of articles especially those who built their models based on real non UCI datasets used variable selection in their model building process.
- Decision trees, rule based classifiers, expert system and any other rule extraction techniques from different data mining techniques are welcomed to the credit scoring and banking industry because of their explicit conditions in accepting/rejecting applicants, and that they are easily understandable by business people compared to other techniques.
- Policy making and evaluating in credit scoring in banks are mainly done with using rules, so the reason is of the importance of new ways through effective rule design and implementation in credit industry.
- “Classification + Clustering” methods are a type of hybrid methods which is not used in reviewed articles but it can identify and extracts potential good and bad applicants groups. Identifying good customer groups helps banks and financial institutes know their customers better and plan their marketing strategies based on different customer clusters.
- With respect to the world financial crises, SMEs are financially weak and easily affected and are bankrupted by fluctuations. Papers focusing on extracting and financially clustering self sufficient silos of business groups are welcomed in the industry to prevent defaults domino effect. This issue applies other data mining techniques in the area of creditworthy business social networks.
- With respect to the research findings, some key papers focused on the area of profit scoring is suggested that profit concept versus default concept developed more financial gains for banks.
- In the field of credit scoring, imbalanced data sets frequently occur as the number of non-worthy applicants is usually much lower than the number of worthy. Some Academics and practitioners reported that non-worthy applicants are usually ten times lower than worthy applicants. So sampling issues on real world credit datasets focused on field of work in the area of credit scoring and there are few researches in the area.

Table 3. Statistics of articles on credit Scoring and data mining techniques.

NO.	Interpretation	Individual credit scoring	Enterprise credit scoring	SME credit scoring	Total
1	Artificial neural networks	12			12
2	Ensembles	8	3		11
3	Support vector machine	9			9
4	Genetic Algorithm	5			5
5	rule based (Fuzzy/non Fuzzy)	5			5
6	Rough set theory	5			5
7	Classification and regression trees	3		1	4
8	multivariate adaptive regression splines	4			4
9	Genetic programming	3			3
10	Grid search	3			3
11	Decision Tree	2			2
12	Discriminant analysis	2			2
13	F score	2			2
14	k-means	2			2
15	Principle component analysis	2			2
16	K nearest neighbor	1			1
17	Expert system	١			1
18	clustering-launched classification	1			1
19	Tabu search	1			1
20	Case-based reasoning	1			1
21	Two-step clustering	1			1
22	Scatter search	1			1
23	Chi-square automatic interaction detector	1			1
Total		75	3	1	79

Table 4. Distribution of articles by journal title.

Journal title	Number	Percentage (%)
Expert Systems with Applications	32	72.7
European Journal of Operational Research	4	9
Computers & Operations Research	2	4.5
Nonlinear Analysis: Real World Applications	1	2.2
Applied Mathematics and Computation	1	2.2
Procedia Computer Science	1	2.2
Advanced Engineering Informatics	1	2.2
Computational Statistics & Data Analysis	1	2.2
Knowledge-Based Systems	1	2.2
International Journal of Forecasting	1	2.2
Total	44	100

- There are so many validation and test methods in the area and accuracy rate, Type I and II errors, Areas under ROC curve are mostly used in the research. These methods are mainly done on "in sample" and "out of sample" records of applicants and "Out of time" and back testing issues are ignored in the reviewed articles. It's another area but it mainly needs the records of applicant's statues at least more than three years.
- The area of collection scoring is rather new in academic publications although there are so much research and software products in the outside market.
- One of the main reasons for limited research in other areas of credit scoring, which includes behavioral scoring, collection scoring, and profit scoring is the lack of appropriate data. So, bridging the gap between academics and Practitioners is of interest. This gap helps practitioners to use data mining techniques better and easier in their works. Establishing benchmark databases like UCI credit databases in other areas of credit research help to develop data mining applications in credit industry research.

This study has some limitations. First, it is limited to the science direct online database and there is a wild variety of online databases. Second, the articles are selected with "credit scoring" keyword and articles that used data mining techniques are selected based on reading articles one by one. Finally, articles which noted above on credit scoring don't use the keywords which are not included.

References

[1] Edelman, D.B. and J.N. Crook. (2002). Credit scoring and its applications. Society for Industrial Mathematics.

[2] Van Gestel, T. and B. Baesens. Credit Risk Management: Oxford University Press.

[3] Hand, D.J. and W.E. Henley. (1997). Statistical classification methods in consumer credit scoring: a review. Journal of the Royal Statistical Society: Series A (Statistics in Society). 160(3), 523-541.

[4] Thomas, L.C. (2000). A survey of credit and behavioural scoring: forecasting financial risk of lending to consumers. International Journal of Forecasting. 16(2), 149-172.

[5] Kamleitner, B. and E. Kirchler. (2007). Consumer credit use: a process model and literature review. Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology.57(4), 267-283.

[6] Abdou, H.A. and J. Pointon. (2011). Credit scoring, statistical techniques and evaluation criteria: a review of the literature. Intelligent Systems in Accounting, Finance and Management.

[7] Leung, K., et al. (2008). A comparison of variable selection techniques for credit scoring.

[8] Cios, K.J., et al. (1998). Data mining methods for knowledge discovery. Kluwer Academic Publishers.

[9] Tsai, C.F. and M.-L. Chen. (2010). Credit rating by hybrid machine learning techniques. Applied Soft Computing. 10(2), 374-380.

[10] Tan, P.N., M. Steinbach, and V. Kumar. (2006). Introduction to data mining. Pearson Addison Wesley Boston.

[11] West, D., S. Dellana, and J. Qian. (2005). Neural network ensemble strategies for financial decision applications. Computers & Operations Research. 32(10), 2543-2559.

[12] Wang, G., et al. (2011). A comparative assessment of ensemble learning for credit scoring. Expert Systems with Applications. 38(1), 223-230.

[13] Paleologo, G., A. Elisseeff, and G. Antonini. (2010). Subagging for credit scoring models. European Journal of Operational Research. 201(2). 490-499.

[14] Hussein A, A. (2009). Genetic programming for credit scoring: The case of Egyptian public sector banks. Expert Systems with Applications. 36(9), 11402-11417.

[15] Šušteršič, M., D. Mramor, and J. Zupan. (2009). Consumer credit scoring models with limited data. Expert Systems with Applications. 36(3, Part 1), 4736-4744.

[16] David, W. (2000). Neural network credit scoring models. Computers & Operations Research. 27(11–12),1131-1152.

[17] Abdou, H., J. Pointon, and A. El-Masry. (2008). Neural nets versus conventional techniques in credit scoring in Egyptian banking. Expert Systems with Applications. 35(3), 1275-1292.

[18] Arie, B.D. (2008) Rule effectiveness in rule-based systems: A credit scoring case study. Expert Systems with Applications. 34(4), 2783-2788.

[19] Ben-David, A. and E. Frank. (2009). Accuracy of machine learning models versus "hand crafted" expert systems – A credit scoring case study. Expert Systems with Applications. 36(3, Part 1), 5264-5271.

[20] Huang, Y.M., C.M. Hung, and H.C. Jiau. (2006). Evaluation of neural networks and data mining methods on a credit assessment task for class imbalance problem. Nonlinear Analysis: Real World Applications. 7(4), 720-747.

[21] Wang, J., K. Guo, and S. Wang. (2010). Rough set and Tabu search based feature selection for credit scoring. Procedia Computer Science. 1(1), 2425-2432.

[22] Hoffmann, F., et al. (2007). Inferring descriptive and approximate fuzzy rules for credit scoring using evolutionary algorithms. European Journal of Operational Research. 177(1), 540-555.

[23] Lee, T.S. and I.F. Chen. (2005). A two-stage hybrid credit scoring model using artificial neural networks and multivariate adaptive regression splines. Expert Systems with Applications. 28(4), 743-752.

[24] Huang, C.L., M.C. Chen, and C.J. Wang. (2007). Credit scoring with a data mining approach based on support vector machines. Expert Systems with Applications. 33(4), 847-856.

[25] Li, S.T., W. Shiue, and M.-H. Huang. (2006). The evaluation of consumer loans using support vector machines. Expert Systems with Applications. 30(4), 772-782.

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

[27] Yingxu, Y. (2007). Adaptive credit scoring with kernel learning methods. European Journal of Operational Research. 183(3), 1521-1536.

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

[29] Xu, X., C. Zhou, and Z. Wang. (2009). Credit scoring algorithm based on link analysis ranking with support vector machine. Expert Systems with Applications. 36(2, Part 2), 2625-2632.

[30] Luo, S.T., B.-W. Cheng, and C.-H. Hsieh. (2009). Prediction model building with clustering-launched classification and support vector machines in credit scoring. Expert Systems with Applications. 36(4), 7562-7566.

[31] Chen, W., C. Ma, and L. Ma. (2009). Mining the customer credit using hybrid support vector machine technique. Expert Systems with Applications. 36(4), 7611-7616.

[32] Chen, F.L. and F.C. Li. (2010). Combination of feature selection approaches with SVM in credit scoring. Expert Systems with Applications. 37(7), 4902-4909.

[33] Nanni, L. and A. Lumini. (2009). An experimental comparison of ensemble of classifiers for bankruptcy prediction and credit scoring. Expert Systems with Applications. 36(2, Part 2), 3028-3033.

[34] Lee, T.S., et al. (2006). Mining the customer credit using classification and regression tree and multivariate adaptive regression splines. Computational Statistics & Data Analysis. 50(4), 1113-1130.

[35] Martens, D., et al. (2007). Comprehensible credit scoring models using rule extraction from support vector machines. European Journal of Operational Research. 183(3), 1466-1476.

[36] Steven, F. (2009). Are we modelling the right thing? The impact of incorrect problem specification in credit scoring. Expert Systems with Applications. 36(5), 9065-9071.

[37] Ping, Y. and L. Yongheng. (2011). Neighborhood rough set and SVM based hybrid credit scoring classifier. Expert Systems with Applications. 38(9), 11300-11304.

[38] Hens, A.B. and M.K. Tiwari. (2012). Computational time reduction for credit scoring: An integrated approach based on support vector machine and stratified sampling method. Expert Systems with Applications.

[39] Wang, J., et al. (2012). Rough set and scatter search metaheuristic based feature selection for credit scoring. Expert Systems with Applications.

[40] Yap, B.W., S.H. Ong, and N.H.M. Husain. (2011). Using data mining to improve assessment of credit worthiness via credit scoring models. Expert Systems with Applications. 38(10), 13274-13283.

[41] Brown, I. and C. Mues. (2012). An experimental comparison of classification algorithms for imbalanced credit scoring data sets. Expert Systems with Applications. 39(3), 3446-3453.

[42] Crone, S.F. and S. Finlay. (2012). Instance sampling in credit scoring: An empirical study of sample size and balancing. International Journal of Forecasting. 28(1), 224-238.

[43] Lee, T.-S., et al. (2002). Credit scoring using the hybrid neural discriminant technique. Expert Systems with Applications. 23(3), 245-254.

[44] Chen, M.-C. and S.-H. Huang. (2003). Credit scoring and rejected instances reassigning through evolutionary computation techniques. Expert Systems with Applications. 24(4), 433-441.

[45] 45. Huang, J.-J., G.-H. Tzeng, and C.-S. Ong. (2006). Two-stage genetic programming (2SGP) for the credit scoring model. Applied Mathematics and Computation. 174(2), 1039-1053.

[46] Chuang, C.-L. and R.-H. Lin. (2009). Constructing a reassigning credit scoring model. Expert Systems with Applications. 36(2, Part 1), 1685-1694.

[47] Nan-Chen, H. (2005). Hybrid mining approach in the design of credit scoring models. Expert Systems with Applications. 28(4), 655-665.

[48] Arijit, L. (2007). Building contextual classifiers by integrating fuzzy rule based classification technique and k-nn method for credit scoring. Advanced Engineering Informatics. 21(3), 281-291.

[49] Tsai, C.-F. and J.-W. Wu. (2008). Using neural network ensembles for bankruptcy prediction and credit scoring. Expert Systems with Applications. 34(4), 2639-2649.

[50] Zhang, D., et al. (2010). Vertical bagging decision trees model for credit scoring. Expert Systems with Applications. 37(12), 7838-7843.

[51] Zhou, L., K.K. Lai, and L. Yu. (2010). Least squares support vector machines ensemble models for credit scoring. Expert Systems with Applications. 37(1), 127-133.

[52] Hsieh, N.-C. and L.-P. Hung. (2010). A data driven ensemble classifier for credit scoring analysis. Expert Systems with Applications. 37(1), 534-545.

[53] Louzada, F., et al. (2011). Poly-bagging predictors for classification modelling for credit scoring. Expert Systems with Applications. 38(10), 12717-12720.

[54] Wang, G., et al. (2012). Two credit scoring models based on dual strategy ensemble trees. Knowledge-Based Systems. 26(0), 61-68.

[55] Sohn, S.Y. and J.W. Kim. (2012). Decision tree-based technology credit scoring for start-up firms: Korean case. Expert Systems with Applications. 39(4), 4007-4012.

[56] Thomas, L.C. (2009). Consumer credit models: pricing, profit, and portfolios. Oxford University Press, USA.

[57] Vapnik, V.N. (2000). The nature of statistical learning theory. Springer Verlag.