

Download full-text PDF

Download citation

Copy link

Article

PDF Available

A Corporate Credit Rating Model Using Support Vector Domain Combined with Fuzzy Clustering Algorithm

July 2012 · [Mathematical Problems in Engineering](#) 2012(1-2)
DOI:[10.1155/2012/302624](#)
License · [CC BY 3.0](#)

Authors:



Xuesong Guo
Xi'an Jiaotong University



Zhengwei Zhu



Jia Shi
Xi'an Jiaotong University

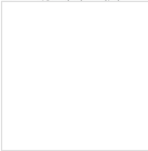
Citations (16)


References (51)


Figures (13)


Abstract and Figures


Corporate credit-rating prediction using statistical and artificial intelligence techniques has received considerable attentions in the literature. Different from the thoughts of various techniques for adopting support vector machines as binary classifiers originally, a new method, based on support vector domain combined with fuzzy clustering algorithm for multiclassification, is proposed in the paper to accomplish corporate credit rating. By data preprocessing using fuzzy clustering algorithm, only the boundary data points are selected as training samples to accomplish support vector domain specification to reduce computational cost and also achieve better performance. To validate the proposed methodology, real-world cases are used for experiments, with results compared with conventional multiclassification support vector machine approaches and other artificial intelligence techniques. The results show that the proposed model improves the performance of corporate credit-rating with less computational consumption.











Figures - uploaded by [Xuesong Guo](#) Author content
Content may be subject to copyright.

Discover the world's research

- 25+ million members
- 160+ millic publi pages
- 2.3+ billion citations

Join for free

 Public Full-text ¹

Content uploaded by [Xuesong Guo](#) Author content
Content may be subject to copyright.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2012, Article ID 302624, 20 pages
doi:[10.1155/2012/302624](#)

Research Article

A Corporate Credit Rating Model Using Support Vector Domain Combined with Fuzzy Clustering Algorithm

Xuesong Guo, Zhengwei Zhu, and Jia Shi

School of Public Policy and Administration, Xi'an Jiaotong University, Xi'an 710049, China

Correspondence should be addressed to Xuesong Guo, guoxues1@163.com

Received 11 February 2012; Revised 19 April 2012; Accepted 9 May 2012

Academic Editor: Wanquan Liu

Copyright © 2012 Xuesong Guo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Corporate credit-rating prediction using statistical and artificial intelligence techniques has received considerable attentions in the literature. Different from the thoughts of various techniques for adopting support vector machines as binary classifiers originally, a new method, based on support vector domain combined with fuzzy clustering algorithm for multiclassification, is proposed in the paper to accomplish corporate credit rating. By data preprocessing using fuzzy clustering algorithm, only the boundary data points are selected as training samples to accomplish support vector domain specification to reduce computational cost and also achieve better performance. To validate the proposed methodology, real-world cases are used for experiments, with results compared with conventional multiclassification support vector machine approaches and other artificial intelligence techniques. The results show that the proposed model improves the performance of corporate credit-rating with less computational consumption.

1. Introduction

Techniques of credit ratings have been applied by bond investors, debt issuers, and governmental officials as one of the most efficient measures of risk management. However, company credit ratings are too costly to obtain, because agencies including Standard and Poor’s (S&P), and Moody’s are required to invest lots of time and human resources to accomplish critical analysis based on various aspects ranging from strategic competitiveness to operational level in detail [1–3]. Moreover, from a technical perspective, credit rating constitutes a typical multiclassification problem, because the agencies generally have much more than two categories of ratings. For example, ratings from S&P range from AAA for the highest-quality bonds to D for the lowest-quality ones.

The final objective of credit rating prediction is to develop the models, by which knowledge of credit risk evaluation can be extracted from experiences of experts and to be applied in much broader scope. Besides prediction, the studies can also help users capture fundamental characteristics of different financial markets by analyzing the information applied by experts.

Although rating agencies take emphasis on experts’ subjective judgment in obtaining ratings, many promising results on credit rating prediction based on different statistical and Artificial Intelligence (AI) methods have been proposed, with a grand assumption that financial variables extracted from general statements, such as financial ratios, contain lots of information about company’s credit risk, embedded in their valuable experiences [4, 5].

Among the technologies based on AI applied in credit rating prediction, the Artificial Neural Networks (ANNs) have been applied in the domain of finance because of the ability to learn from training samples. Moreover, in terms of defects of ANN such as overfitting, Support Vector Machine (SVM) has been regarded as one of the popular alternative solutions to the problems, because of its much better performance than traditional approaches such as ANN [6–11]. That is, an SVM’s solution can be globally optimal because the models seek to minimize the structural risk [12]. Conversely, the solutions found by ANN tend to fall into local optimum because of seeking to minimize the empirical risk.

However, SVM, which was originally developed for binary classification, is not naturally modified for multiclassification of many problems including credit ratings. Thus, researchers have tried to extend original SVM to multiclassification problems [13], with some techniques of multiclassification SVM (MSVM) proposed, which include approaches that construct and combine several binary classifiers as well as the ones that directly consider all the data in a single optimization formulation.

In terms of multiclassification in the domain of credit rating containing lots of data, current approaches applied in MSVM still have some drawbacks in integration of multiple binary classifiers as follows.

- (1) Some unclassifiable regions may exist if a data point belongs to more than one class or to none.
- (2) Training binary classifiers based on two-class SVM multiple times for the same data set often result in a highly intensive time complexity for large-scale problems including credit ratings prediction to improve computational consumption.

To overcome the drawbacks associated with current MSVM in credit rating prediction, a novel model based on support vector domain combined with kernel-based fuzzy clustering is proposed in the paper to accomplish multiclassification involved in credit ratings prediction.

2. Literature Review

2.1. Credit Rating Using Data Mining Techniques

Major researches applying data mining techniques for bond rating prediction can be found in the literature.

Early investigations of credit rating techniques mainly focused on the applicability of statistical techniques including multiple discriminant analysis (MDA) [14, 15] and logistic regression analysis (LRA) [16], and so forth, while typical techniques of AI including ANN

Table 1: Prior bond rating prediction using AI techniques.

Research	Number of categories	AI methods applied	Data source	Samples size
[20]	2	BP	U.S	30/17

Download full-text PDF	Download citation	Copy link			
	[17]	6	BP, RPS	U.S (S&P)	110/60
	[23]	6	BP	U.S (S&P)	N/A
	[24]	6	BP	U.S (Moody's)	299
	[25]	5	BP with OPP	Korea	126
	[26]	6	BP, RBF	U.S (S&P)	60/60
	[27]	5	CBR, GA	Korea	3886
	[28]	5	SVM	U.S (S&P)	N/A
	[29]	5	BP, SVM	Taiwan, U.S	N/A

[17, 18] and case-based reasoning (CBR) [19], and so forth are applied in the second phase of research.

The important researches applying AI techniques in bond-rating prediction are listed in Table 1. In summary, the most prior ones accomplish prediction using ANN with comparison to other statistical methods, with general conclusions that neural networks outperformed conventional statistical methods in the domain of bond rating prediction.

On the other hand, to overcome the limitations such as overfitting of ANN, techniques based on MSVM are applied in credit rating in recent years. Among the models based on MSVM in credit rating, method of Grammar and Singer was early proposed by Huang et al., with experiments based on different parameters so as to find the optimal model [29]. Moreover, methodologies based on One-Against-All, One-Against-One, and DAGSVM are also proposed to accomplish S&P's bond ratings prediction, with kernel function of Gaussian RBF applied and the optimal parameters derived form a grid-search strategy [28]. Another automatic-classification model for credit rating prediction based on One-Against-One approach was also applied [30]. And Lee applied MSVM in corporate credit rating prediction [31], with experiments showing that model based on MSVM outperformed other AI techniques such as ANN, MDA, and CBR.

2.2. Multiclassification by Support Vector Domain Description

Support Vector Domain Description (SVDD), proposed by Tax and Duin in 1999 [32] and extended in 2004 [33], is a method for classification with the aim to accomplish accurate estimation of a set of data points originally. The methods based on SVDD differ from two or multiclass classification in that a single object type is interested rather than to be separated from other classes. The SVDD is a nonparametric method in the sense that it does not assume any particular form of distribution of the data points. The support of unknown distribution of data points is modeled by a boundary function. And the boundary is “soft” in the sense that atypical points are allowed outside it.

The boundary function of SVDD is modeled by a hypersphere rather than a hyperplane applied in standard SVM, which can be made with less constrains by mapping the data points to a high-dimensional space using methodology known as kernel trick, where the classification is performed.

Download full-text PDF

Download citation

Copy link

Download full-text PDF

Download citation

Copy link

Citations (16)

References (51)

... Zhang & Chi, 2018) Multi-objective Programming Artificial Intelligence and Machine Learning 22 (Michael & Constantin, 2011) Multi-criteria Outranking Modeling (Hsu Che et al., 2014) Bagging-DT (K.-j. Kim & Ahn, 2012), (Mileris & Boguslauskas, 2011) Artificial Intelligence (Dwivedi & Tripathi, 2019), (Shi et al., 2016) (R. H. Abiyev, 2014) (Cao & Xiong, 2014),, (Guo et al., 2012) , (Gong, 2017), (Baofeng et al., 2016), , (Estevez & Carballob, 2015), (Al-Najjar & Al-Najjar, 2014), (Guo et al., 2012), (Hongxia et al., 2010), (Shyu, 2008), (Jiang, 2007), (Kao & Wang, 2006), (D. Zhang & Zhang, 2004) Neural Network Random Forest (RF) (Taba et al., 2019) Novel Algorithm Data Mining 5 (Bae & Kim, 2015), (Hristova et al., 2017), (Moradi & rafiei, 2019), (G. ...

... Zhang & Chi, 2018) Multi-objective Programming Artificial Intelligence and Machine Learning 22 (Michael & Constantin, 2011) Multi-criteria Outranking Modeling (Hsu Che et al., 2014) Bagging-DT (K.-j. Kim & Ahn, 2012), (Mileris & Boguslauskas, 2011) Artificial Intelligence (Dwivedi & Tripathi, 2019), (Shi et al., 2016) (R. H. Abiyev, 2014) (Cao & Xiong, 2014),, (Guo et al., 2012), (Gong, 2017), (Baofeng et al., 2016), , (Estevez & Carballob, 2015), (Al-Najjar & Al-Najjar, 2014), (Guo et al., 2012) , (Hongxia et al., 2010), (Shyu, 2008), (Jiang, 2007), (Kao & Wang, 2006), (D. Zhang & Zhang, 2004) Neural Network Random Forest (RF) (Taba et al., 2019) Novel Algorithm Data Mining 5 (Bae & Kim, 2015), (Hristova et al., 2017), (Moradi & rafiei, 2019), (G. ...

Elements of Credit Rating: A Hybrid Review and Future Research Agenda

Article Full-text available

Feb 2021

Prashant Ubarhande · Arti Chandani

View Show abstract

... Since modelling, evaluating and determining consumer credit risk have become more important day by day, banks need to take a lot of precautions which is about monitoring and collection of loans. Investors and banks may be able to benefit by choosing portfolios that are more closely aligned with their preferences [8] [9][10]. Therefore, while they give a loan, they evaluate some basic variables, for instance, customer's income, age, sex, job and past credit record. ...

DETERMINATION OF VARIABLES IN CALCULATION OF CREDIT SCORE IN TURKEY TO DECREASE DEFAULT RISK

Presentation Full-text available

Jul 2021

Hakkı Aydoğdu · Deniz Unal

View Show abstract

... The ratings of the agencies that sort risk are viewed by financial agents as efficient for risk management, although they represent a very costly procedure [3] . However, recent events have tarnished the image of these agencies. ...

Sovereign Credit Risk Assessment with Multiple Criteria Using an Outranking Method

Article Full-text available

Sep 2018 · MATH PROBL ENG

Diogo Ferreira de Lima Silva · Julio Cezar Soares Silva · Lucimário G. O. Silva · Adiel Teixeira de Almeida-Filho

View Show abstract

... The second mapping T connects each unit in A to a particular weight value W. For multi-dimensional input space, the required storage capacity of A will show a geometric growth, it may be very large. In this case, it can be used hash programming techniques to reduce the storage capacity [10] . The output of the CMAC neural network is equal to the sum of the C weights that have been activated. ...

Multiple multidimensional fuzzy reasoning algorithm based on neural network

Article Full-text available

Oct 2018

Zhiwei Zhao · Guiqiang Ni · Yuanyuan Shen · Nasruddin Hassan

View Show abstract

... Since modelling, evaluating and determining consumer credit risk have become more important day by day, banks need to take a lot of precautions which is about monitoring and collection of loans. Investors and banks may be able to benefit by choosing portfolios that are more closely aligned with their preferences [8] [9][10]. Therefore, while they give a loan, they evaluate some basic variables, for instance, customer's income, age, sex, job and past credit record. ...

Determination of Variables in Calculation of Credit Score in Turkey to Decrease Default Risk

Article

Jan 2018


Deniz Unal · Hakkı Aydoğdu

View




Download full-text PDF



Download citation

Copy link

Article
Jan 2022
Ravi Makwana · Dhruvil Bhatt · Kirtan Delwadia ·  Bhaskar Chaudhury
[View](#)

Loan Default Prediction with Machine Learning Techniques
[Conference Paper](#)
Aug 2020
Lili Lai
[View](#)

Sequence-based clustering applied to long-term credit risk assessment
[Article](#)
Sep 2020 · [EXPERT SYST APPL](#)
 Richard Le ·  Hyejin Ku ·  Doobae Jun
[View](#) [Show abstract](#)

Fractional-Order Modeling and Fuzzy Clustering of Improved Artificial Bee Colony Algorithms
[Article](#)
Aug 2019
 Shenghan Zhou · Xingxing Xu · Zhenzhong Xu ·  Yiyong Xiao
[View](#) [Show abstract](#)

Credit Scoring Using Information Fusion Technique
[Conference Paper](#)
Nov 2018
Di Wang · Zuoquan Zhang
[View](#)

[Show more](#)

Download full-text PDF

 Download citation

 Copy link



Recommendations [Discover more](#)

- Article

Full-text available

Learning networks for tornado detection

February 2006 · International Journal of General Systems

Theodore B. Trafalis · Budi Santosa · Michael Richman

In this paper, different types of learning networks, such as artificial neural networks (ANNs), Bayesian neural networks (BNNs), support vector machines (SVMs) and minimax probability machines (MPMs) are applied for tornado detection. The last two approaches utilize kernel methods to address non-linearity of the data in the input space. All methods are applied to detect when tornadoes occur, ... [\[Show full abstract\]](#)

[View full-text](#)
- Article

Application to Classifier Design

January 2008

Sadaaki Miyamoto · Hidetomo Ichihashi · Katsuhiro Honda

This chapter is devoted to a description of the postsupervised classifier design using fuzzy clustering. We will first derive a modified fuzzy c-means clustering algorithm by slightly generalizing the objective function and introducing some simplifications. The k-harmonic means clustering [177, 178, 179, 119] is reviewed from the point of view of fuzzy c-means. In the algorithm derived from the ... [\[Show full abstract\]](#)

[Read more](#)
- Article

Full-text available

Fuzzy Clustering Data Given in the Ordinal Scale

January 2017 · International Journal of Intelligent Systems and Applications

Oleksii Tyshchenko · Yevgeniy Bodyanskiy · ZHENGBING HU · Viktoriia O. Samitova

A fuzzy clustering algorithm for multidimensional data is proposed in this article. The data is described by vectors whose components are linguistic variables defined in an ordinal scale. The obtained results confirm the efficiency of the proposed approach.

[View full-text](#)
- Article

Support vector classifier based on fuzzy c-means and Mahalanobis distance

October 2010 · Journal of Intelligent Information Systems

Yong Zhang · Fuding Xie · Dan Huang · Min Ji

This paper presents a fuzzy support vector classifier by integrating modified fuzzy c-means clustering based on Mahalanobis distance into fuzzy support vector data description. The proposed algorithm can be used to deal with the outlier sensitivity problem in traditional multi-class classification problems. The modified fuzzy c-means clustering algorithm based on Mahalanobis distance takes into ... [\[Show full abstract\]](#)

[Read more](#)
- Conference Paper

Bill classification by using the LVQ method

February 2001

T. Kosaka · S. Omatu · T. Fujinaka

For pattern classification problems the neuro-pattern recognition, which is the pattern recognition based on the neural network approach, has been increasingly popular since it can classify various patterns similar to human beings. In this paper we adopt the learning vector quantization (LVQ) method to classify the various bank notes. The reasons to use LVQ are that it can process the ... [\[Show full abstract\]](#)

[Read more](#)



Company

[About us](#)
[News](#)
[Careers](#)

Support

[Help Center](#)

Business solutions

[Advertising](#)
[Recruiting](#)