STUDY OF PERSONAL CREDIT RISK ASSESSMENT BASED ON SUPPORT VECTOR MACHINE ENSEMBLE

Chong Wu, Yingjian Guo, Xinying Zhang and Han Xia

School of Economy and Management
Harbin Institute of Technology
No.92 West Da-Zhi Street, Harbin, Heilongjiang Province, P. R. China
wuchong@hit.edu.cn

ABSTRACT. *Credit risk assessment, essentially a classification problem, plays an important role in bank credit risk management. Hence, how to assess credit risks in commercial banks effectively becomes an important issue in the financial field. In recent years, a number of studies have provided the discipline within sights into the practice of credit risk management. Credit risk assessment, a classification problem in nature, is a basic task in credit risk management of commercial banks whose purpose is to analyze the credit risk of banks so as to offer a decision-making ground for loan business. Support vector machine (SVM) ensemble has been proposed to improve classification performance recently. However, ensemble strategies currently used can not evaluate the output importance degrees of component SVM classifiers when combining component predictions to the final decision. To overcome this shortcoming, we propose a method of support vector machine (SVM) ensemble based on fuzzy integral to discriminate good creditors from bad ones. Compared the experiment results of SVM ensemble with that of single SVM and the neural network ensemble, the proposed method outperforms the single SVM and the neural network ensemble in terms of classification accuracy. Our experiments show that the proposed method, stable, highly accurate, strong robust and feasible, can provide a sound credit assessment system.*
Keywords: Credit index, Credit assessment, Support vector machine ensemble, Customer to customer, Bagging, Fuzzy integral

1. **Introduction.** Credit risk assessment is the set of decision models and techniques that aid lenders in granting consumer credit by assessing the risk of lending to different consumers. Credit risk assessment, a classification problem in nature, is a basic task in credit risk management of commercial banks whose purpose is to analyze the credit risk of banks so as to offer a decision-making ground for loan business. An accurate assessment of risk could be translated into a more efficient use of resources or a less loan loss for a bank. In recent years, many models like statistical methods and artificial neural networks are usually used to evaluate credit ranks [1-3]. However, there are some disadvantages (1) what the traditional statistical method studies is the Incremental Theory when the samples tend to be infinitive, so it demands a certain amount of sample data; (2) the availability of methods is closely related to many necessary assumptions and conditions when clustering models are established. For example, Multiple Discriminant Analysis requires data to obey the multiple normal distribution and equal covariance, while in reality, large amounts of data do not meet these assumptions. Therefore statistical methods can not meet research needs. Artificial neural network has a local optimal problem, poor generalization ability, weak theory and uncontrolled converge speed.