Chapter 2[1]

Application and Comparison of Classification Techniques in Controlling Credit Risk

Lan Yu[1], Guoqing Chen[2,3], Andy Koronios[4], Shiwu Zhu[2], Xunhua Guo[2]
[1]Department of Computer Science and Technology
[2]School of Economics and Management
[3]Research Center for Contemporary Management,
Tsinghua University, Beijing 100084, China.
[4]School of Computer and Information Science, University of South Australia, Australia.

Abstract: Credit rating is a powerful tool that can help banks improve loan quality and decrease credit risk. This chapter examines major classification techniques, which include traditional statistical models (LDA, QDA and logistic regression), *k*-nearest neighbors, Bayesian networks (Naïve Bayes and TAN), decision trees (C4.5), associative classification (CBA), a neural network and support vector machines (SVM), and applies them to controlling credit risk. The experiments were conducted on 244 rated companies mainly from the Industrial and Commercial Bank of China. The receiver operating characteristic curve and the Delong-Pearson method were adopted to verify and compare their performance. The results reveal that while traditional statistical models produced the poorest outcomes, C4.5 or SVM did not perform satisfactorily, and CBA seemed to be the best choice for credit rating in terms of predictability and interpretability.

***Key Words*:** Classification, credit risk, ROC curve, Delong-Pearson method.

---

[1] Liao, T.W. and E. Triantaphyllou, (Eds.), **Recent Advances in Data Mining of Enterprise Data**, *World Scientific*, Singapore, pp. 111-145, 2007.