Chapter 9
Feature Selection for Bankruptcy Prediction:
A Multi–Objective Optimization Approach

A. Gaspar-Cunha
University of Minho, Portugal

F. Mendes
University of Minho, Portugal

J. Duarte
Instituto Superior de Engenharia do Porto, Portugal

A. Vieira
Instituto Superior de Engenharia do Porto, Portugal

B. Ribeiro
University of Coimbra, Portugal

A. Ribeiro
Technical University of Lisbon, Portugal

J. Neves
Technical University of Lisbon, Portugal

ABSTRACT

In this work a Multi-Objective Evolutionary Algorithm (MOEA) was applied for feature selection in the problem of bankruptcy prediction. This algorithm maximizes the accuracy of the classifier while keeping the number of features low. A two-objective problem, that is minimization of the number of features and accuracy maximization, was fully analyzed using the Logistic Regression (LR) and Support Vector Machines (SVM) classifiers. Simultaneously, the parameters required by both classifiers were also optimized, and the validity of the methodology proposed was tested using a database containing financial statements of 1200 medium sized private French companies. Based on extensive tests, it is shown that MOEA is an efficient feature selection approach. Best results were obtained when both the accuracy and the classifiers parameters are optimized. The proposed method can provide useful information for decision makers in characterizing the financial health of a company.

1. INTRODUCTION

Financial bankruptcy prediction is of high importance for banks, insurance companies, creditors and investors. One of the most important threats for business is the credit risk associated with counterparts. The rate of bankruptcies have increased in recent years and its becoming harder to estimate as companies become more complex and develop sophisticated schemes to hide their real situation. Due to the recent financial crisis and regulatory concerns, credit risk assessment is a very active area both for academic and business community. The ability to discriminate between faithful customers from potential bad ones is thus crucial for commercial banks and retailers (Atiya, 2001).

Different approaches have been used to analyze this problem, like discriminant analysis (Eisenbeis, 1977) and Logit and Probit models (Martin, 1977). However, most of these methods have important limitations. Discriminant analysis is limited due to its linearity, restrictive assumptions, for treating financial ratios as independent variables and can only be used with continuous independent variables. Furthermore, the choice of the regression function creates a bias that restricts the outcome and they are also very sensitive to exceptions, while has an implicit Gaussian distribution on data, which is inappropriate in many cases.

More recently other approaches have been applied for bankruptcy classification, such as Artificial Neural Networks (ANN) (Atiya, 2001; Charitou et al., 2004; Neves & Vieira, 2006), Evolutionary Algorithms (EA) and Support Vector Machines (SVM) (Fan & Palaniswami, 2000). ANN, EA and SVM are used as complementary tools to classify credit risk. Some of the studies performed show that ANN outperforms discriminant analysis in bankruptcy prediction (Neves & Vieira, 2006; Coats & Fant, 1993; Yang, 1999; Tan & Dihardjo, 2001). Huang et al. (2008) conclude that financial ratios are important tools in prediction of business failures and that they are commonly used to develop the models or classifiers. In their work failed firms are targeted aiming to seek out relevant features of their financial ratios. To this end, automatic clustering techniques are employed to automatically divide targeted failed firms into some clusters according to characteristics of financial ratios. In order to simplify the task of analysis, as well as to increase the classification accuracy, feature selection techniques are used to reduce the overall number of financial ratios analyzed. Also, in their paper the authors, particularly emphasizes the importance of both expert knowledge and data mining techniques in feature selection. This means that it is preferable to conduct the analysis task using not only the data mining technique but also the expert knowledge, and to compare their performances of classification accuracies in terms of the feature selection. In this way, more accurate results and practical insights can be obtained. More recently, Wu (2010) proposed a method which directly explores the features of failed firms rather than researching pairs of failed and non-failed firms. To this end, automatic clustering techniques and feature selection techniques are employed for this study. Taking these conclusions into account, it is generally recognized that further research is needed to achieve higher predictive capabilities, which is the avenue of the present research (Vieira et al., 2009).

Banks collect large amounts of data available from companies and other creditors. These data is often inconsistent and redundant and needs considerable manipulation to make it useful for problems like credit risk analysis. First, it is necessary to build a set of ratios that may be appropriated for the problem. Then, is necessary to further restrict the number of these ratios, or attributes with higher information content in order to reduce the complexity of the problem. Finally, these reduced set of attributes, or features, are used to train any classification algorithm designed to predict the company financial health.

Related Content

Algorithms and Computations in BL-Algebras

Celestin Lele (2010). *International Journal of Artificial Life Research (pp. 29-47).*

www.irma-international.org/article/algorithms-computations-algebras/49682

Fast Learning in Neural Networks

Darryl Charles, Colin Fyfe, Daniel Livingstoneand Stephen McGlinchey (2008). *Biologically Inspired Artificial Intelligence for Computer Games (pp. 91-104).*

www.irma-international.org/chapter/fast-learning-neural-networks/5908

Multi-Objective Binary Fish School Search

Mariana Gomes da Motta Macedo, Carmelo J. A. Bastos-Filho, Susana M. Vieiraand João M. C. Sousa (2018). *Critical Developments and Applications of Swarm Intelligence (pp. 53-72).*

www.irma-international.org/chapter/multi-objective-binary-fish-school-search/198921

Materialized View Selection using Marriage in Honey Bees Optimization

Biri Arunand T.V. Vijay Kumar (2015). *International Journal of Natural Computing Research (pp. 1-25).*

www.irma-international.org/article/materialized-view-selection-using-marriage-in-honey-bees-optimization/146882

Evolutionary Algorithms

P. Collet (2007). *Handbook of Research on Nature-Inspired Computing for Economics and Management (pp. 45-58).*

www.irma-international.org/chapter/evolutionary-algorithms/21119