Knowledge-Based Systems 000 (2017) 1–9

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Predicting the listing statuses of Chinese-listed companies using decision trees combined with an improved filter feature selection method

Ligang Zhou [a,*], Yain-Whar Si [b], Hamido Fujita [c]

[a] School of Business, Macau University of Science and Technology, Taipa, Macau
[b] Department of Computer and Information Science, University of Macau, Macau
[c] Faculty of Software and Information Science, Iwate Prefectural University, Iwate, Japan

ARTICLE INFO

Article history:
Received 23 December 2016
Revised 28 April 2017
Accepted 3 May 2017
Available online xxx

Keywords:
Multi-class classification
Listing-status prediction
Decision tree C4.5
Decision tree C5.0

ABSTRACT

Predicting the listing statuses of Chinese-listed companies (PLSCLC) is an important and complex problem for investors in China. There is a large quantity of information related to each company's listing status. We propose an improved filter feature selection method to select effective features for predicting the listing statuses of Chinese-listed companies. Due to the practical concerns of analysts in finance about the performance and interpretability of the prediction models, models based on decision trees C4.5 and C5.0 are employed and are compared with several other widely used models. To evaluate the models' robustness with time, the models are also tested under rolling time windows. The empirical results demonstrate the efficacy of the proposed feature selection method and decision tree C5.0 model.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

To help make investors aware of the risks of the listed companies in Chinese stock markets, the China Security Regulatory Commission (CSRC) issued special listing rules for the listed companies in 1998. According to the rules, a listed company that falls into abnormal operational or financial conditions will be given a special risk warning label, known as special treatment (ST), to indicate its risk.

According to the "Rules Governing the Listing of Stocks (RGLS)" on the Shanghai Stock Exchange (SHSE) [1] and Shenzhen Stock Exchange (SZSE) [2] released in July 2012, there are two types of risk warning: indication of the delisting risk and indication of other risks. The special treatment of a stock with a delisting risk warning or other risk warning is to add the prefix of "*ST" or "ST" to the stock symbol, respectively. In addition, the limit of the price increase/decrease within a trading day is 5% for a stock with a risk warning and 10% for a normal stock. In China's stock market, a stock with a delisting or other risk warning is always called an ST stock and the corresponding company is called an ST company.

The RGLS on SHSE and SZSE lists similar conditions for implementing the delisting risk warning on the listed companies. The conditions of abnormal financial conditions are specified as follows [1,2]:

(1) The audited net profit of the company was negative in the last two consecutive fiscal years.
(2) The audited net worth of the company was negative in the last fiscal year.
(3) The audited operating income of the company was less than 10 million Yuan in the last fiscal year.
(4) The financial statements for the last fiscal year were noted with an adverse opinion or disclaimer opinion from the auditing company.
(5) The company has been commanded by the China Securities Regulatory Commission (CSRC) to correct serious errors and false records, but has failed to do so within the specified time window, and the company's stocks have been suspended from trading for two months.
(6) The company fails to disclose its annual report or semi-annual report within the statutory time window and the company's stock has been suspended from trading for two months.

If any of the following conditions occur, the listed company will be given another risk warning:

(1) The operation of the company has been seriously affected and cannot be recovered in three months.
(2) The main bank account of the company has been blocked.

* Corresponding author.
E-mail addresses: mrlgzhou@gmail.com (L. Zhou), fstasp@umac.mo (Y.-W. Si), fujita.2578@gmail.com (H. Fujita).

http://dx.doi.org/10.1016/j.knosys.2017.05.003
0950-7051/© 2017 Elsevier B.V. All rights reserved.

Please cite this article as: L. Zhou et al., Predicting the listing statuses of Chinese-listed companies using decision trees combined with an improved filter feature selection method, Knowledge-Based Systems (2017), http://dx.doi.org/10.1016/j.knosys.2017.05.003

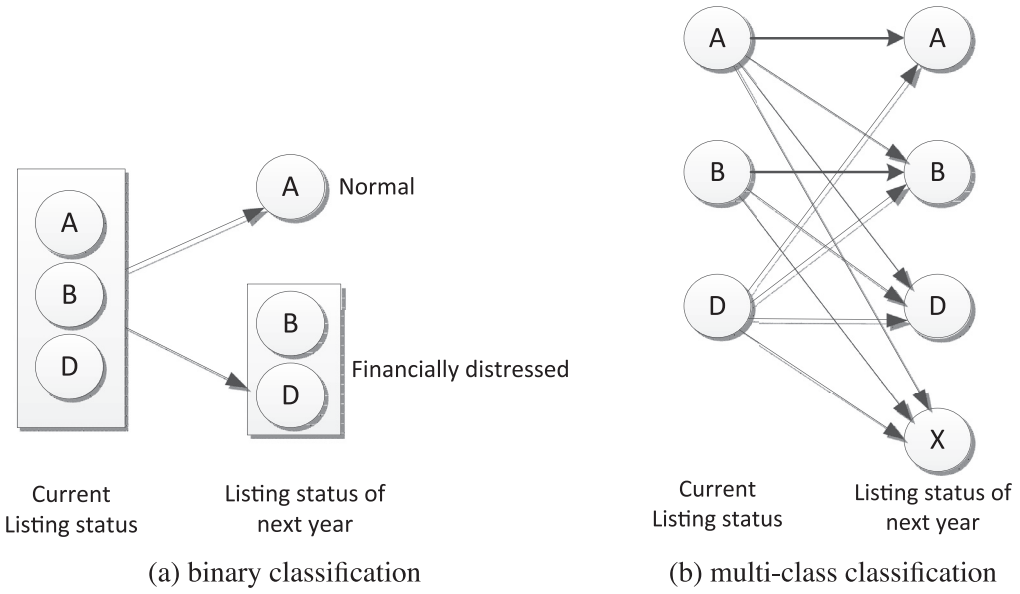

Fig. 1. The different formulations of the listing-status prediction problem.

(3) The board of the directors cannot hold meetings on a regular basis and cannot reach a resolution.
(4) The company funds the controlling shareholders or their stakeholders, or the company violates the regulations on providing guarantees for external obligations.

As indicated above, the main reason for a normal company to receive ST is its poor financial performance. Therefore, financial ratios are important factors in the listing status prediction.

A normal company may be given special treatment; at the same time, an ST company has the opportunity to get the risk warning removed if its financial status has improved and it can meet certain requirements specified in RGLS. If a company with a risk warning fails to meet some requirements, it may be delisted. Since July 1999, the ST firms that exhibit no sign of financial improvement in the following year will be issued a particular transfer warning (hereinafter called PT firms) issued by CSRC [3]. However, on February 25, 2002, the PT warning policy was canceled by CSRC. Therefore, in this study, the PT warning is not considered and the listing statuses of the stocks will be categorized into four groups: (1) a normal company without any risk warning; (2) a company with delisting risk warning; (3) a company with another risk warning; and (4) a delisted company. These four listing-status groups are denoted by A, B, D and X, respectively, in this research.

Due to the different risk levels implied by the listing statuses, it is crucial for investors to forecast the stocks' listing statuses to manage the risk of their portfolios or make stock investment decisions in Chinese stock markets.

Most preliminary research [3–5] forecasts listing status of companies in China by classifying a currently listed company as either a normal company or a financially distressed company, as shown in Fig. 1(a). Since a risk warning given to a company always indicates the company's financial distress, the prediction of the listing status of a company in China is usually formulated as a financial distress prediction problem. In a prediction model constructed for the normal listed companies, a company can have one of two possible financial statuses in the considered forecasting period: normal or financially distressed. Therefore, the financial distress prediction problem is a typical binary classification problem. Altman et al. [3] developed a model called Z_{china} Score to predict financial distress of the listed companies in China. They defined the "ST" and "PT" companies as financially distressed firms and the model was constructed based on a training sample consisting of 30 financially distressed and 30 healthy companies that were announced in 1998 or 1999. Ding, et al. [4] developed a prediction model based on support vector machines (SVM) to predict the financial distress of Chinese high-tech manufacturing companies. Li and Sun [5] proposed a hybrid Gaussian case-based reasoning system for predicting the business failures of Chinese-listed companies.

This study predicts the listing statuses of the listed companies, based on four different listing statuses, as shown in Fig. 1(b). The study demonstrates that the listing status prediction problem is a multi-class classification (MC) problem in practice, which is simplified as a binary classification problem in most existing studies [3–5].

Since the listing status can indicates a company's risk level and affect the liquidity of the company's stock, correct prediction of a listed company's listing status is very important for the investors and stake holders of the company. Moreover, the interpretability of the predictive models can help financial analysts judge the reliability of the models and thus increase the applicability of the models. Therefore, we select decision tree models C4.5 and C5.0 for predicting the listing statuses of Chinese-listed companies. C5.0 is an extension of decision tree C4.5 with improvements in computational efficiency and computational space. Some widely used approaches for multi-class classification are also used for benchmarking, such as linear discriminant analysis (LDA), Adaboost (AD), Neural networks (NN), random forest (RF), and Bayesian networks (BN). These multi-class classification models can be implemented easily and have been successfully applied in different applications [6–10].

The performance in predicting the listing statuses of Chinese-listed companies can be affected not only by the classification approach but also by the selection of features. There are different categories of features that can be used in PLSCLC, such as macroeconomic factors, company characteristics, financial indicators and market information. Most previous studies have demonstrated that financial indicators and market information are most effective in financial distress prediction [3,11–14]. However, there are hundreds of different financial indicators and market information that can be

Please cite this article as: L. Zhou et al., Predicting the listing statuses of Chinese-listed companies using decision trees combined with an improved filter feature selection method, Knowledge-Based Systems (2017), http://dx.doi.org/10.1016/j.knosys.2017.05.003

Download English Version:

https://daneshyari.com/en/article/4946218

Download Persian Version:

https://daneshyari.com/article/4946218

Daneshyari.com