

# Study of Financial Warning Ensemble Model for Listed Companies Based on Unbalanced Classification Perspective

Wei Cong, School of Management Science and Engineering, Nanjing University of Information Science and Technology, Nan Jing, China

## ABSTRACT

Using the ensemble learning method to mine valuable information from a sea of financial data accumulated on the market of financial securities is very important for studying data processing. On the basis of financial data from A-share companies listed on Shanghai Stock Market, this article takes the perspective of unbalanced classification of ST stocks to carry out a study of the construction of a financial warning model for the listed companies. In our experiment, HDRF (HDRandom Forest, Hellinger Distance based Random Forest), ensemble classification models of Bagging, AdaBoost, and Rotation Forest, which take Hellinger distance decision tree (HDDT) as the base classifier, and the ensemble classification model which takes the C4.5 decision tree as the base classifier, are compared in respect of both the area under the ROC curve and the F-measure. As shown in the experimental results, the HDRF and the HDDT based classifier, as an ensemble method, are effective for financial data of listed companies.

## KEYWORDS

Financial Warning, Hellinger Distance Based Random Forest (Hdrrandom Forest, HDRF), Hellinger Distance Decision Tree (HDDT), Unbalanced Classification

## 1. INTRODUCTION

A key direction for studying the integration of analytical processing and ensemble learning of financial data is to accurately mine valuable information from a sea of real-time, incomplete and unbalanced data on China's securities market. The objective of this article is to introduce different ensemble learning methods from the perspective of unbalanced classification to explore the ensemble classification model suitable for the corresponding data structure. There has been no research in using the after mentioned algorithm for financial warning prediction. The contribution of this paper is to build a relatively effective classification model to improve the forecasting ability of ST shares in financial data of listed companies, which has certain theoretical and practical significance for the construction and application of financial warning model under background of financial big data.

DOI: 10.4018/IJIIIT.2020010103

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

### 1.1. Theories Relating to Financial Warning for Listed Companies

Financial warning, as an important content of the management of a company's financial risks, for a listed company, the financial warning means to make statistics of the company's data such as financial statements and operating plans to predict its operational and financial activities as well as to find out operational and financial risks in its operation and management activities, so as to inform shareholders and managers with warning information.

A number of scholars have already made systematic studies on financial warning. Among them, Qin (2012) applied a Logistic regression analysis model to empirically study the selection of financial warning variables for companies listed on Shanghai and Shenzhen A-share markets whose stocks were, for the first time, prefixed with "ST". Xiao and Yang (2018) constructed  $L_{1/2}$ -regularized Logistic Regression model and took ST and non-ST manufacturing companies in Shanghai and Shenzhen stock markets as the research objects for comparative prediction. Song et al. (2019) used Cox regression to build financial crisis warning model for minor enterprises to achieve relatively accurate prediction of healthy enterprises. By using the annual reporting data of two sets of 37 pairs of "A" shares, Wang (2008) built a warning model for listed companies based on a multiple regression model and checked the accuracy as well, and he also gave his opinion and suggestions about the financial warning system. Li et al. (2013) adopt multiple discriminant analysis to modify a Z-Score baseline model, and establish a financial distress early warning model applicable to listed real estate companies in China. The findings indicate that the average accuracy of the financial distress early warning model reaches higher than 90%. Wang et al. (2018) combined PSO-BP and FOA-BP neural network model to select data of high-tech enterprises to effectively predict the financial crisis of enterprises. As indicated in the above studies, in most of the empirical studies about the financial warning for listed companies, a reasonable statistical analysis model is usually built based on financial data to check the validity and make a systematic analysis of the warning model. This paper has studied financial warning ensemble model for listed Companies based on Unbalanced Classification Perspective.

### 1.2. Brief Introduction of ST Shares in Listed Companies

According to the regulations of CSRC and securities exchanges (Jiang & Wang, 2005), when the financial condition or other condition of a listed company is abnormal, which makes it difficult for investors to judge the company prospects and the equity of investors may be damaged, the share exchanges of the company should be specially treated and the share is called ST (Special Treatment) share. The main reason to decide whether a share is ST share is that the audit results of two recent accounting years indicate that the net profits are negative, or in the most recent fiscal year of the audit results show that the shareholders' equity is lower than the registered capital, the securities exchange and the CSRC identified as abnormal financial position.

If the stock of a listed company is rated as a ST (special treatment) share, then it means that there may be serious problems in the company's financial state or the operation and management of the company violates the requirements on a listed company. Then, warning information of financial crisis or of poor operation and management of the company can be acquired from the market (Liu, 2016).

Classification of listed companies' ST shares is an important issue that securities regulators, shareholders and managers of the companies and stock investors have to pay attention to. A key direction to research financial warning system is to make accurate prediction of ST shares, which is also the premise for making good decisions on operation and management (Zheng, 2013).

In the current paper, hence, a financial crisis will be defined if the financial state of a listed company is judged to be abnormal as a result of the stock of the company being rated as a ST share. The ensemble financial warning model for listed companies will also be studies from the perspective of the classification of ST shares by using the data mining and unbalanced classification techniques.

15 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: [www.igi-global.com/article/study-of-financial-warning-ensemble-model-for-listed-companies-based-on-unbalanced-classification-perspective/243369](http://www.igi-global.com/article/study-of-financial-warning-ensemble-model-for-listed-companies-based-on-unbalanced-classification-perspective/243369)

## Related Content

---

### A 2D Positioning Application in PET Using ANNs

Fernando Mateo, Ramón J. Aliaga, Jorge D. Martínez, José Ma Monzóand Rafael Gadea (2009). *Encyclopedia of Artificial Intelligence* (pp. 1576-1582). [www.irma-international.org/chapter/positioning-application-pet-using-anns/10448](http://www.irma-international.org/chapter/positioning-application-pet-using-anns/10448)

### Constraint Processing

Roman Barták (2009). *Encyclopedia of Artificial Intelligence* (pp. 404-409). [www.irma-international.org/chapter/constraint-processing/10279](http://www.irma-international.org/chapter/constraint-processing/10279)

### Bio-Inspired Computing through Artificial Neural Network

Nilamadhab Dash, Rojalina Priyadarshini, Brojo Kishore Mishraand Rachita Misra (2017). *Fuzzy Systems: Concepts, Methodologies, Tools, and Applications* (pp. 1285-1313). [www.irma-international.org/chapter/bio-inspired-computing-through-artificial-neural-network/178441](http://www.irma-international.org/chapter/bio-inspired-computing-through-artificial-neural-network/178441)

### Applying CI in Biology through PSO

Rojalina Priyadarshini, Nilamadhab Dash, Brojo Kishore Mishraand Rachita Misra (2016). *Handbook of Research on Computational Intelligence Applications in Bioinformatics* (pp. 119-143). [www.irma-international.org/chapter/applying-ci-in-biology-through-psy/157485](http://www.irma-international.org/chapter/applying-ci-in-biology-through-psy/157485)

### Robotic Innovations in Agriculture: Maximizing Production and Sustainability

Kamran Ikram, Noman Ali Buttar, Muhammad Mohsin Waqas, Muthmainnah Muthmainnah, Muhammad Mubashar Omer, Yasir Niazand Alex Khang (2023). *Handbook of Research on AI-Equipped IoT Applications in High-Tech Agriculture* (pp. 131-154). [www.irma-international.org/chapter/robotic-innovations-in-agriculture/327832](http://www.irma-international.org/chapter/robotic-innovations-in-agriculture/327832)