Knowledge-Based Systems 57 (2014) 41–56

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

Predicting financial distress and corporate failure: A review from the state-of-the-art definitions, modeling, sampling, and featuring approaches

Jie Sun *, Hui Li, Qing-Hua Huang, Kai-Yu He

School of Economics and Management, Zhejiang Normal University, Jinhua, Zhejiang 321004, PR China

ARTICLE INFO

Article history:
Received 13 April 2012
Received in revised form 6 December 2013
Accepted 6 December 2013
Available online 13 December 2013

Keywords:
Definition of financial distress
Sampling methods
Featuring methods
Review
Financial distress prediction
Corporate failure prediction
Case-based reasoning
Ensemble
Group decision-making
Support vector machine
Hybrid modeling
Neural network
Decision tree
Logistic regression
Multiple discriminant analysis

ABSTRACT

As a hot topic, financial distress prediction (FDP), or called as corporate failure prediction, bankruptcy prediction, acts as an important role in decision-making of various areas, including: accounting, finance, business, and engineering. Since academic research on FDP has gone on for nearly eighty years, there are abundant literatures on this topic, which may appear chaotic to the researchers of the field and make them feel confused. This paper contributes to the current review researches by making a full summary, analysis and evaluation on the current literatures of FDP. The current literatures of FDP are reviewed from the following four unique aspects: definition of financial distress in the new century, FDP modeling, sampling approaches for FDP, and featuring approaches for FDP. By considering the new state-of-the-art techniques in this area, FDP modeling are classified and reviewed by the following groups: namely, modeling with pure single classifier, modeling with hybrid single classifier, modeling by ensemble techniques, dynamic FDP modeling, and modeling with group decision-making techniques. Sampling methods for FDP are classified and reviewed by the following paired groups, namely: training sampling and testing sampling, single industry sampling and cross-industry sampling, balanced sampling and imbalanced sampling. Featuring methods for FDP are categorized and reviewed by qualitative selection and combination of qualitative and quantitative selection. We comment on the current researches from the view of each category and propose further research topics. The review paper is valuable to guide research and application of the area.

1. Introduction

The word of "Early Warning" is originated from military area. Nowadays, the word is widely used in some respects such as: macroeconomics, business administration, environmental monitoring, among others. Early warning of financial distress, corporate failure, or bankruptcy is an important research field for corporate finance and its core is financial distress prediction (FDP), which is an extensive ongoing research topic. Generally, FDP is to predict whether or not a company will fall into financial distress based on the current financial data, through mathematical, statistical, or intelligent models. It is also called as financial failure discrimination, bankruptcy prediction, business failure prediction, corporate failure prediction, among others. It plays an important role on managerial decision-making for firms, investment decision-making for investors, credit decision-making for creditors, customer credit rating for banks, and so on.

During the last fifty years, numerous researches focus on this topic and attempt to provide a useful solution for the problem. Literatures on this topic may appear chaotic to the researchers of the field and make them feel confused. A thoughtful review on related articles is useful to help people understand the research and development in FDP. Thus, this paper attempts to review the state-of-the-art researches from a new perspective, which is different from the several review articles on FDP in recent years. Balcaen and Ooghe [11] extensively elaborated on the application of classical statistical techniques in FDP from the following categories, namely: single variable analysis, risk index models, multivariate discriminant analysis (MDA), and conditional probability models. This review made a contribution to the current literature by clarifying the characteristics of the classical statistical techniques in FDP and their related problems. However, intelligent techniques are more frequently used in FDP than statistical techniques in recent years. Thus, Kumar and Ravi [58] considered both statistical

* Corresponding author. Tel.: +86 579 8229 8567.
E-mail address: sunjiehit@gmail.com (J. Sun).

0950-7051/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.knosys.2013.12.006

and intelligent techniques in FDP, and classified articles of FDP from the following groups, namely: statistical techniques, neural networks (NN), case-based reasoning (CBR), decision trees (DT), operational research, evolutionary algorithms (EA), rough set (RS) based techniques, other techniques, and soft computing techniques from the hybridization of all the above-mentioned techniques. They also reviewed each article on the source of data, financial variables, country, time line of the data, and performance of techniques. This review presented that the hybrid modeling is a hot topic in FDP. By considering researches in this category, Bahrammirzaee [10] conducted a comparative review of three famous intelligent techniques for financial applications, i.e., artificial neural networks, expert systems and hybrid intelligent systems. Credit evaluation and financial prediction are two chief problems for financial market. They concluded that these intelligent methods were superior to traditional statistical ones in dealing with financial problems, although such superiority is not absolute. Verikas et al. [125] presented a comprehensive review of hybrid and ensemble-based soft computing techniques for FDP. Their focus was on how different techniques were combined. A technique was named as hybrid if several soft computing approaches were applied and only one predictor was used to make the final prediction. A technique was called as ensemble if outputs of several predictors were combined to obtain a prediction. Lin et al. [74] made a review of FDP literatures between 1995 and 2010. They carried out statistical analysis on the FDP literatures in terms of single and soft classifiers, baseline classifiers, and datasets, and considered soft classification techniques as the direction for future research of FDP. Marques et al. [78] summarized the application of evolution computing to credit scoring, which included classification, variable selection, and parameter optimization. They also reviewed performance evaluation criteria, statistical significance tests, credit databases, and some data preprocessing issues.

Though the six reviews have been published, this review is valuable and significant for the following reasons:

(1) From the above six review articles we can find that hybrid intelligent techniques are potentially useful models for FDP with high predictive performance. Thus, how to construct hybrid models for FDP needs to be analyzed by grouping the current work.
(2) Meanwhile, with the development of computing and modeling techniques, some new techniques may also be useful for solving the problem of FDP, e.g., dynamic prediction of financial distress, class imbalance classification techniques, feature selection approaches, group decision-making techniques, among others. How these new emerged techniques can be used in FDP needs to be analyzed by summarizing existing researches.
(3) Third, none of the existing review articles provides a clear indication on FDP with different sampling techniques and feature selection approaches. This topic needs to be investigated.
(4) Fourth, some review articles in the last century discussed the definition of financial distress, which was seldom focused on in recent reviews. However, the definition of financial distress should be re-considered since some new concept or phenomena come into the area of FDP in the new century. The current reviews need some supplement to help researchers understand the area of FDP more thoughtfully.

This review is to fetch up the gap of the current reviews. We reviewed the current FDP literatures from the following four unique aspects: definition of financial distress in the new century, FDP modeling, sampling methods for FDP, and featuring approaches for FDP. In detail, FDP modeling articles are classified and reviewed by the following groups: namely, modeling with pure single classifier, modeling with hybrid single classifier, modeling by ensemble techniques, dynamic FDP modeling, and modeling with decision implementations. Sampling methods for FDP are classified and reviewed by the following paired groups, namely: training sampling and testing sampling, cross-industry sampling and single industry sampling, balanced sampling and imbalanced sampling. Feature selection methods for FDP are categorized and reviewed by qualitative selection and quantitative selection.

This paper is organized as follows. Section 2 overviews and comments on the definition of financial distress in the new century. Section 3 reviews and comments on FDP modeling methods from the category of pure single classifier methods, hybrid single classifier methods, ensemble methods, dynamic modeling for FDP, and FDP-based decision implementations. Section 4 reviews and comments on sampling methods for FDP. Section 5 presents the overview and comments from the perspective of featuring methods for FDP. Section 6 concludes and analyzes some further research topics in FDP.

2. Development of definition of financial distress in the new century

2.1. Development of definition of financial distress

Financial distress is the situation that an enterprise has certain kind of financial difficulties. In some classical literatures, such financial difficulties include inability to pay debts or preferred dividend and the corresponding consequences such as overdraft of bank deposits, liquidation for interests of creditors, and even entering the statutory bankruptcy proceeding [13,3,29]. Such definition of financial distress is based on the theoretical framework of "cash flow" or "liquid assets" model. As Beaver [13] mentioned, an enterprise is like a reservoir formed by the cash flow, composed of cash inflows and outflows. An enterprise in financial distress is just like a reservoir whose water is drained.

Carminchael [18] believed that financial difficulty is a situation that an enterprise encounters frustration in fulfilling its obligations. These frustrations include: insufficiency of liquidity, insufficiency of equity, default of debt, and insufficiency of liquid capital. Foster [35] defined financial distress as a serious liquidity problem which is unable to be resolved without large-scale restructuring of the operation or structure of economic entities. In Doumpos and Zopounidis [32], financial distress not only contains inability to repay important obligatory payments and the corresponding consequences mentioned above, but also includes the situation of negative net asset value, which means an enterprise's total liabilities exceed its total assets from the view of accounting.

Ross et al. [100] summarized previous studies and concluded that financial difficulties consist of the following four conditions: (1) business failure, that is, a company cannot pay the outstanding debt after liquidation; (2) legal bankruptcy, namely, a company or its creditors applies to the court for a declaration of bankruptcy; (3) technical bankruptcy, namely, a company cannot fulfill the contract on schedule to repay principal and interest; and (4) accounting bankruptcy, namely: a company's book net assets are negative.

In the 21st century, most FDP literatures that collect data from developed countries or areas concentrate on the prediction of bankruptcy, which is the ultimate and most serious form of financial distress [76,106,89,22]. Bose [14] defined financial distress as the condition that a company's stock price is less than 10 cents, which is followed by Ravisankar et al. [96]. When studying FDP for Taiwan companies, Lin [72] defined financial distress as the inability of a firm to pay its financial obligations as they mature.