Available online at www.sciencedirect.com

ScienceDirect

Expert Systems with Applications 36 (2009) 3536–3542

Expert Systems with Applications

www.elsevier.com/locate/eswa

A study of applying data mining approach to the information disclosure for Taiwan's stock market investors

Chi-Lin Lu [a], Ta-Cheng Chen [b,*]

[a] *Department of Business Administration, National Formosa University, Taiwan*
[b] *Department of Information Management, National Formosa University, 64 Wenhua Road, Huwei, Yunlin 632, Taiwan*

Abstract

The financial literature and practices have shown the importance of corporate governance for decades, not only for firm's management but also for investor protection. Information disclosure plays a key role in all of the governance mechanisms. With good information disclosure, the asymmetric information and the agency cost between the insider and the outsider of firms can be reduced effectively. However, the information disclosure status of listed companies is hard to be evaluated or judged by investors before the annual official announcement is reported in the next year. The main purpose of this study is to explore the hidden knowledge of information disclosure status among the listed companies in Taiwan's stock market. In this paper, we employed decision tree-based mining techniques to explore the classification rules of information transparency levels of the listed firms in Taiwan's stock market. Moreover, the multi-learner model constructed by boosting ensemble approach with decision tree algorithm has been applied. The numerical results show that the classification accuracy has been improved by using multi-leaner model in terms of less Type I and Type II errors. In particular, the extracted rules from the data mining approach can be developed as a computer model for the prediction or the classification of good/poor information disclosure potential and like expert systems.

Keywords: Information disclosure; Data mining; Classification rule

1. Introduction

In recent years, a number of financial scandals, such as Enron and WorldCom, have made the corporate governance to become a noticeable topic to both practice and academy in financial management. Most of the problems were caused by the asymmetric information between the insiders and the outsiders of the firms. The Organization for Economic Co-operation and Development (OECD) has identified the corporate governance as a direct monitoring system to maximize the value of firms through the mechanism of transparency and efficiency. OECD also declared six principles for corporate governance in 2004 (http://www.oecd.org/dataoecd/32/18/31557724.pdf). Among these principles, information disclosure plays a key role to ensure whether the internal and the external governance mechanisms work well or not. The two mechanisms defined by World Bank in 1999 (http://www.sovereignglobal.com/media/framework_for_implemenation.pdf) are as follows: internal governance includes board of director's monitor, compensation of manager and incentive plan, etc.; external governance includes external auditing, professional investors and government, etc. Obviously, the above two governance mechanisms have to rely on correct and proper information disclosed by the firms. However, the judgment of information disclosure status is with some incentive-bias problems. For instance, firms prefer releasing good news and concealing the bad ones, such as those firms who violate the regulations or suffer the financial distress. Thus, the whole worldwide stock market authorities keep reviewing the firms' information disclosure status and related regulations to protect the interests of investors and public.

* Corresponding author. Tel.: +886 5 6315740; fax: +886 5 6364127.
E-mail address: tchen@nfu.edu.tw (T.-C. Chen).

0957-4174/$ - see front matter
doi:10.1016/j.eswa.2008.02.007

The evaluation of information disclosure is more complex than just posting their financial statements or annual reports on the publication or website. According to the Public Company Accounting Reform and Investor Protection Act of 2002, except the accurate financial statements, the factors for evaluating the information disclosure status must include the off-balance sheet transaction, pro forma financial information, major stockholder transaction, manager disclosure valuation, code of ethics of CEO, and opinions of financial experts. Obviously, it is difficult for the individuals or even professional investors to understand the full dimensions of a firm's disclosure situation, because we never know whether the firm shows out the whole picture truly or not. Only the authority has the jurisdiction over the authenticity for a firm's disclosure. For better understanding the transparency of listed firms, the Securities and Futures Institute (SFI) of Taiwan initiated the investigation of the information disclosure level among all the listed firms in stock market. Although the investigation report is announced annually, it always takes more than one year to have this official report. So, the investors cannot do anything by this antiquated announcement.

Moreover, the information disclosure also has some significant relationship with some specified firm's business characteristics. For example, Chen and Jaggi (2000) and Eng and Mak (2003) have shown the association among the ownership structure, board composition, firm characteristics and disclosure. In this paper, we developed a transparency evaluation approach by using data mining techniques. It is wished to help investors to identify the corporate governance situation, and then make their investment decisions correctly in time. There are rare studies on the use of this new emerging technology on this topic. Through mining the ownership structure and firm characteristics data by using decision tree-based techniques, we can classify and predict which ownership structure and characteristics will bring out which type of transparency in time, and does not have to wait for the official report until next year.

2. Literature review

2.1. Information disclosure

Modern agency theory shows that the major problem of corporate governance comes from the asymmetric information between the public and the insiders of the firms. Investors who chase the high quality information create the needs of firm transparency. These firms with higher transparency have better performance in stock price. Ball and Brown (1968) indicated that the announcement of accounting earnings, by event study, has a significant impact on stock price. Amir and Lev (1996) and Eng and Mak (2003) showed that the financial information has significant relationship with stock price and the value of firm. Verrecchia (1983) indicated that the information disclosure can be used to improve the corporate governance and reduce the agency problems. Although with some cost, the investors have better confidence with high transparency firms. Bushee and Noe (2000) studied the relationship between the information disclosure and the institutional investors. They found out that the institutional investors prefer the firms with higher transparency than those with lower transparency. Sandeep, Amra, and Liliane (2002) investigated four emerging markets and they also suggested that information disclosure is one of the most important mechanisms to corporate governance. The above studies indicated that information disclosure plays a significant role in corporate governance.

2.2. Ownership structure and information disclosure

The structure of ownership determines the level of monitoring and thereby the level of disclosure. Many studies have shown the strong relationship between them. Cerf (1962) documented the significant characteristics impact on firm such as size, listing, leverage and managerial ownership on disclosure. Recent empirical studies on the association between disclosure and ownership structure (Chen & Jaggi, 2000; Eng & Mak, 2003; Forker, 1992) demonstrated the strong relationship between the corporate governance and the share option disclosure. Chen and Jaggi (2000) examined the association between the independent non-executive directors and the comprehensiveness of financial information disclosures. Eng and Mak (2003) summarized the significant relationship between ownership structure and voluntary disclosure in the financial statements. From the above studies, it is shown that the information disclosure is influenced by the factors of ownership structure, but does not show how it was influenced by those factors.

2.3. Data mining

Data mining usually means the approaches and appliances to nugget the implicit, unknown, potential and valuable knowledge (Frawley, Piatetsky-Shapiro, & Matheus, 1991). It can be looked as a form of knowledge discovery essential for solving problems in a specific domain. Based on the data mining techniques, a response model can be built as a decision model for the prediction or the classification of a domain problem potential like expert systems. In real world, problems are highly nonlinear in nature so that it is hard to develop a comprehensive model taking into account all the independent variables through the conventional statistical methods. As we know, data stored in database are growing at a phenomenal rate since they can be captured and stored easily and inexpensively. It makes the knowledge hidden in the database be more hardly explored and utilized in practice. If the data mining strategies are utilized to infer the relationships between the status of information disclosure and the companies, the stock investors can have much better guideline for their investment. Recently, nonlinear and complex machine learning approaches were presented,