Auditor Change Prediction Using Data Mining and Audit Reports

A

Wikil Kwak
University of Nebraska at Omaha, USA

Xiaoyan Cheng
University of Nebraska at Omaha, USA

Yong Shi
University of Nebraska at Omaha, USA

Fangyao Liu
Southwest Minzu University, China

Kevin Kwak
University of Nebraska at Omaha, USA

INTRODUCTION

The emergence of big data and the increasing use of analytics has brought both opportunities and challenges to the audit profession (Appelbaum, Kogan, & Vasarhelyi, 2018; Appelbaum, Showalter, Sun, & Vasarhelyi, 2021). Audit reports should convey a firm's financial wellbeing to stakeholders including investors, managers, debtholders, and regulators. Thus, audit reports represent the formal communication between the auditor and the interested parties. Auditors have the option to choose four different types of auditor report opinions: unqualified opinion, qualified opinion, disclaimer of opinion, and adverse opinion. Unlike an unqualified opinion, a qualified opinion signals a lack of conformity with Generally Accepted Accounting Principles (GAAP) and casts doubt on the credibility of financial statements. Chow and Rice (1982) document that the audited firms, after receiving a qualified opinion, tend to switch auditors more often than firms receiving a clean opinion. This switch is motivated by opinion shopping (seeking a new auditor who is willing to provide an improved audit opinion). However, Williams (1988) finds no significant relation between the receipt of a qualified opinion and auditor changes. Considerable research (Krishnan & Stephens, 1995; Krishnan, Krishnan, & Stephens, 1996; Chen, Peng, Xue, Yang, & Ye, 2016) has examined the policy implications regarding firms changing auditors after receiving a qualified opinion. In this chapter, we use data mining approaches to revisit the prediction of auditor changes upon the receipt of a qualified opinion in U.S. firms.

A couple decades ago, the American Institute of Certified Public Accountants (AICPA) and Institutes of Internal Auditors (IIA) endorsed data mining as one of the top ten technologies of tomorrow (Koh, 2004). However, there are not many studies in the U.S. to examine the relation between audit opinions and auditor changes using data mining techniques. Several studies investigated audit opinion issues and their overall prediction rates were qualitatively similar in predicting auditor changes using traditional logit or discriminant analysis (Maggina & Tsakianganos, 2011; Fernández-Gámez, García-Lagos, &

DOI: 10.4018/978-1-7998-9220-5.ch001

Sánches-Serrano, 2016). This chapter presents an effort to better understand the application of big data techniques in auditor change decision. A stream of research (Koyuncugil & Ozgulbas, 2012; Zhou, Lu, & Fujita, 2015; Read & Yezegel, 2016) in financial distress modelling uses data mining tools to forecast bankruptcy, and these techniques help auditors improve their judgment when issuing auditor report opinions. Various algorithms and techniques like Support Vector Machine (SVM), decision tree, and k-means are used for our study as they are the most popular in knowledge discovery.

The chapter proceeds as follows: The next section presents relevant prior research. The third section discusses sampling criteria, sources of data, and reports the findings. The last section summarizes and discusses the conclusions of this chapter and provides further research avenues.

BACKGROUND AND PRIOR RESEARCH

Financial distress variables in predicting bankruptcy are useful for auditors to predict going concern evaluation (Gepp, Linnenluecke, O'Neill, & Smith, 2018). A stream of research (Sundgren & Svanstrom, 2014; Fernández-Gámez et al., 2016) in financial distress predicts audit opinion decisions. Sundgren and Svanstrom (2014) find that the phases of auditor careers are related to audit reporting quality, proxied by the auditors' propensity to issue a going-concern opinion prior to bankruptcy filings. Fernández-Gámez et al. (2016) utilized neural networks to predict a qualified audit opinion using corporate governance and financial ratios. Nevertheless, non-Big 4 firms tend to be reluctant to use big data techniques to determine going concern of a firm (Read & Yezegel, 2016).

Another stream of research centers on the prediction of auditor changes. Prior studies reveal that there is a positive association between bankruptcy and auditor changes. Hudaib and Cooke (2005) assume that financial distress may influence a firm's decision to change auditors either directly or indirectly. They found that the probability of switching auditors increases as financial health declines using financial ratio variables or the Z-score variable in their multivariate logistic regression. Dhaliwal, Lamoreaux, Lennox, and Mauler (2015) document that firms with Big-4 auditors have significant influence over auditor selection by management. Moreover, Landsman, Nelson, and Rountree (2009) used client-specific audit and financial risk factors, along with client misalignment, to model auditor switch decisions in pre- and post-Enron periods. In this study, we include Andersen as an audit firm with the Enron scandal. To further examine the accuracy of auditor changes, Sun (2007) suggests that adding industry effect, stock returns, and advanced statistical methods may increase the prediction rate. Bankruptcy is an extreme form of financial distress, with most of the variables used by Altman (1968) and Ohlson (1980) in their bankruptcy prediction studies related to the auditor changes decision. It is assumed that the auditor change decision is made by top management, but the audit committee also participates in the auditor selection and change process.

Classification is the most common data mining method, which uses a set of pre-tagged datasets to train the model for future prediction. Fraud detection and risk analysis are classic examples of applications in this technology. SVM and decision trees are the most well-known algorithms in classification or two category (0, 1) prediction such as auditor change or bankruptcy studies. Support Vector Machine (SVM) was developed by Vladimir Vapnik and is primarily used in the application of classification problems, though it can also be used for regression analysis (Cortes & Vapnik, 1995). The algorithm realizes classifications by finding the most appropriate hyper-planes that can successfully differentiate between classes in a training dataset. The decision of a hyper-plane is to maximize the distances from the closest points in each of the different classes. The algorithm learning process of SVM can be formulated as a

Related Content