Corporate Disclosure Measurement

Md. Salah Uddin Rajib
Jahangirnagar University, Bangladesh

Md. Qutub Uddin Sajib
China University of Geosciences (CUG), China

INTRODUCTION

Researchers have mentioned that about 90% of the entire world's data has been created in the past 2 years (e.g., Kolb & Kolb, 2013). Most of these data are made in electronic format. Corporate disclosures (CDs) are available now in electronic formats through different documents also. As the organizations are disclosing the disclosure annually or in the interim period, the numbers of volumes are increasing gradually. Organizations disclose disclosure through various media, e.g., annual reports, websites, supplementary of financial statements and others.

A number of researchers have investigated the corporate disclosure (CD) scenario in the last decades as it is assumed to have an impact on the capital market (Healy & Palepu, 2001). CD can be categorized into two broad groups – Mandatory disclosures and Voluntary disclosures. Mandatory disclosures are obligatory by the concern organization of a country (e.g., Securities and Exchange Commission). In contrast, voluntary disclosures are made by the firms voluntarily considering the interest of stakeholders. A number of topics on CD have been investigated e.g., financial disclosure (e.g., Malone et al., 1993), human resource capital (e.g., Bontis, 2003), environmental disclosure (e.g., Gamble et al., 1996), strategic disclosure (e.g., Wagenhofer, 1990) and others.

Very often, researchers have used content analysis to explore the scenario of CD practice (e.g., Bontis, 2003; Khan et al., 2011). In content analysis they normally count the frequency of words to explore the scenario. Researchers normally select a list of related terms or keywords to the concern topic and search the terms in the annual reports or other related supplementary where disclosure has been presented by the organizations (e.g., Bontis, 2003).

Researchers have urged for more reliable measurement of CD practice (e.g., Rahman & Post, 2011). Consistent with that, this chapter proposes the use of text mining (TM) approach for measuring the CD. Currently, different areas of researches are applying the approaches and techniques of TM. Researcher of medical science, politics, business and others have used the techniques of TM to extract the information (Moohebat et al., 2015). The purpose of TM is to process unstructured textual information. In large text collection, TM works as a tool for knowledge discovery (Gomez et al., 2002). TM uses several techniques to extract the information, e.g., text classification, text clustering, text summarization, sentiment analysis and others. This chapter investigates the usability and advantages of TM approaches in the measurement of CD. Therefore, the aim of this paper is fourfold. First, it reviews the background of CDs and their measurement and background of TM. Second, it discusses the TM approaches, including general model of TM and methodology. Third, the usability of TM in measuring the CD has been investigated. And finally, it discusses the prospects of research in the area of using the TM in CD measurement.

BACKGROUND

As mentioned earlier, CD can be categorized into two broad groups - mandatory and voluntary disclosure. The demand for CD primarily arises from the perspective of regulation. Therefore, Securities and Exchange Commission (SEC) played an important role in the CD, made by the organizations. History states that because of various crises in US e.g., war, unemployment, racial conflict, environmental degradation and others, authority tempts to examine the fundamental principles underlying the body of corporate law for policy making (Schoenbaum, 1972). Sarbanes-Oxley (SOX) requires more and better information also. Consisting with the situation The Securities Exchange Act of 1934 imposes disclosure requirements on corporations in the US (Schoenbaum, 1972). With time, Generally Accepted Accounting Principles (GAAP) imposes some disclosures on corporations. Voluntary disclosure drives by various factors, e.g., manager-specific influences (e.g., Bamber et al., 2010), corporate governance (e.g., Hermalin & Weisbach, 2012), external pressure groups (e.g., Islam & Deegan, 2010) and others. Therefore, voluntary disclosures are associated with various areas and it is unstructured in nature. Researchers have tempted to find out the relationship of CD with the organization performances (Healy & Palepu, 2001). Therefore, measurement of CD emerges as a research area for the researchers.

This chapter conducts studies on the usability of text mining (TM) for CD measurement. TM is a discipline of natural language processing. The approach of TM has evolved from the simple word processing at the end of the 1990s to now (Antonio & Ferneda, 2008). TM derives much of its inspiration and direction from seminal research on data mining (Feldman & Sanger, 2007). The increased numbers of electronic documents have triggered the use of TM to extract the knowledge from textual data.

TEXT MINING: GENERAL MODEL AND METHODS

The primary objective of TM is to retrieve information from unstructured text and to present the extracted knowledge to the users in a concise form (Ananiadou & McNaught, 2006). It has been mentioned by the researchers that TM can add significant importance and economic value where the data are in electronic form and are in centralized warehouses (e.g., Weiss et al., 1999).

Text Mining (TM)

TM is defined as a tool for extracting the information from electronic documents by using natural language processing computer program. Broadly, TM can be defined as a knowledge-intensive process in which a user interacts with a document collection over time by using a suite of analysis tools (Feldman & Sanger, 2007). Like the data mining, TM extracts useful information from the sources through various approaches. It can collect data both from formalized database records and from unstructured textual data.

General Model of Text Mining

In a TM technique, computer programs scan text in a document and apply model consisting with the objective. The individual documents are used as raw cases for TM. The basic approach of TM can be shown by the following figure -

Figure 1 shows that the documents assemble directory drive from the documents that are raw cases for the TM techniques. After representing the text in the computerized program, the features of the text will be extracted. After extracting the feature, various pre-specified models can be applied. The frequency of the terms can be counted easily or document can be classified considering various aspects.

Related Content

Assessment Approach of Enterprise Readiness to Digital Transformation
Anna Zotovaand Valentina Mantulenko (2019). *Educational and Social Dimensions of Digital Transformation in Organizations (pp. 145-173).*
www.irma-international.org/chapter/assessment-approach-of-enterprise-readiness-to-digital-transformation/215140

Methodology of Climate Change Impact Assessment on Forests
Mostafa Jafari (2018). *Encyclopedia of Information Science and Technology, Fourth Edition (pp. 3114-3130).*
www.irma-international.org/chapter/methodology-of-climate-change-impact-assessment-on-forests/184023

Using Logical Architecture Models for Inter-Team Management of Distributed Agile Teams
Nuno António Santos, Jaime Pereira, Nuno Ferreiraand Ricardo J. Machado (2022). *International Journal of Information Technologies and Systems Approach (pp. 1-17).*
www.irma-international.org/article/using-logical-architecture-models-for-inter-team-management-of-distributed-agile-teams/289996

Target Tracking Method for Transmission Line Moving Operation Based on Inspection Robot and Edge Computing
Ning Li, Jingcai Lu, Xu Chengand Zhi Tian (2023). *International Journal of Information Technologies and Systems Approach (pp. 1-15).*
www.irma-international.org/article/target-tracking-method-for-transmission-line-moving-operation-based-on-inspection-robot-and-edge-computing/321542

Web Navigation Systems for Information Seeking
Guangzhi Zheng (2015). *Encyclopedia of Information Science and Technology, Third Edition (pp. 7693-7701).*
www.irma-international.org/chapter/web-navigation-systems-for-information-seeking/112472