

FRIDAYS: A Financial Risk Information Detecting and Analyzing System

Chi-Han Du,¹ Yi-Shyuan Chiang,^{1*} Kun-Che Tsai,^{1*}
Liang-Chih Liu,² Ming-Feng Tsai,^{3,4} Chuan-Ju Wang¹

¹Research Center for Information Technology Innovation, Academia Sinica, Taiwan

²Department of Information and Finance Management, National Taipei University of Technology, Taiwan

³Department of Computer Science, National Chengchi University, Taiwan

⁴MOST Joint Research Center for AI Technology and All Vista Healthcare, Taiwan

hans@citi.sinica.edu.tw, s107074604@m107.nthu.edu.tw, q40603.cs05@nctu.edu.tw,

lcliu@ntut.edu.tw, mftsai@cs.nccu.edu.tw, cjwang@citi.sinica.edu.tw

Abstract

We present FRIDAYS, a financial risk information detecting and analyzing system that enables financial professionals to efficiently comprehend financial reports in terms of risk and domain-specific sentiment cues. Our system is designed to integrate multiple NLP models trained on financial reports but on different levels (i.e., word, multi-word, and sentence levels) and to illustrate the prediction results generated by the models. The system is available online at <https://cfda.csie.org/FRIDAYS/>.

Introduction

Financial reports, carefully written by accounting professionals and composed of detailed information about the operating conditions of companies and domain-specific terms, usually make for tedious, difficult reading. Given the crucial information contained in such financial reports, they are usually the object of formal requests by administrations for regulation. Auditors and investors across governments and business domains are thus obliged to digest all materials manually, in order to assess the operational risk for official decision-making; yet the nature of financial report texts leads to great difficulties for auditors and investors, who seek to interpret these reports both comprehensively and efficiently.

A further challenge is that the semantics contained in text is far from a simple composition of the meanings of individual words, for which usual keyword matching techniques are often inadequate or even infeasible in some cases. Though this phenomenon is universal when dealing with text information, it becomes a serious problem for financial report analysis as the sentences in reports are usually long and complicated. Thus, a significant amount of time and effort is required for practitioners to decode the information in financial reports.

Given the importance and the great challenge of financial report interpretation, it is surprising that there are few tools to assist professionals in this task. A rare exception is the publicly available RiskFinder (Liu et al. 2018), which however only visualizes the prediction result of text classification models on financial sentences. Therefore, our goal is to build

a comprehensive system for detecting financial risk,* and for visualizing financial reports to reduce the burden on financial personnel or other professionals such as educators or researchers and to bridge the gap between technical results and useful interpretation. FRIDAYS, the proposed system, is an NLP system that analyzes 10-K filings – the annual financial reports of companies subject to SEC regulations – and specializes in visualizing and integrating the four sentiment classes of financial words (FWs) and financial multiword expressions (FMWEs) (i.e., positive, negative, litigious, and uncertain), and the financial risk levels of sentences (FSs) (including high, natural, and low risk levels). To address the challenges of revealing information between the different text granularity levels, the proposed system integrates multiple NLP models. These models either detect the existence of FWs and FMWEs, classify their sentiments (Schneider and Smith 2015), or classify FS risk (Joulin et al. 2017). Our system aims to showcase three different kinds of analytical information at the same time in an effective way. FRIDAYS provides two types of user interfaces: one for the assessment of a single report, and the other for an overview of a given company. The system helps users to efficiently comprehend financial sentiment and risk delivered by financial reports.

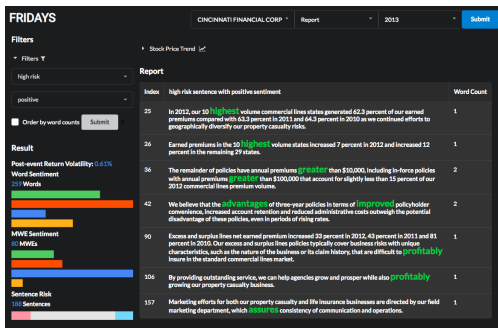
With the proposed system, we provide the following potential usage scenarios for professionals: (1) Financial professionals (e.g., accountants) can easily identify problematic (e.g., potentially high-risk) sentences, and use it as a double-check tool for use in decision-making; (2) Accounting educators can verify the assumptions and conclusions in financial NLP research, and can also utilize the system to facilitate their teaching; (3) Machine learning researchers can use the system to assess prediction results and thus locate potential problems in proposed models to improve performance.

User Interface

FRIDAYS enables users to investigate financial risk using two different methods: (1) Single report assessment, analyzing the risk composition and sentiment cues of individual reports; (2) Company overview, comparing a company's reports over the years. To assess a single report, as shown in Fig. 1(a), by

*These authors contributed equally to this work.
Copyright © 2019, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

*Financial risk is often understood to include the potential for financial loss and uncertainty about its extent.



(a) Single report assessment



(b) Company overview

Figure 1: System interface

using the drop-down menus in the top-right corner, users first select a company, select ‘report’, and then select a specific year; the system then displays the sentences from the corresponding report in the right panel. The three bar graphs in the left panel indicate FW occurrence, FMWE occurrence, and FS composition. On clicking on any bar in the charts, the corresponding texts within the report are emphasized for illustration; e.g., users see the negative FMWEs in the report by clicking on the orange bar in the MWE sentiment chart. To assess a company’s reports across different years, as shown in Fig. 1(b), users choose a company of interest and then select ‘overview’; the system then displays two interactive charts: (1) a chart showing the stock trend over the years, as a quantitative overview of the selected company, as well as the corresponding release dates for the reports; (2) the distributions of FS and FW for each report as an overview of the textual information in the company’s reports. Key advanced functionalities by which FRIDAYS surpasses its predecessors include the following:

User-friendly selection via multiple drop-down menu enables users to search for companies, to view single or multiple report(s), and to specify the year of interest.

Report assessment offers a visualization of an individual report for a fiscal year. Users interact with the left-hand chart to render the report by directly clicking on the bar charts. Unlike previous work, the system visualizes the report in terms of sentences along with different levels of text granularity.

Company overview reveals key information about a company including the overall stock trend, the future annualized post-event return volatilities after the report release date, the FS compositions, and FW occurrence data.

Synergistic filter located at the top-left corner enables users to render reports by both FS and FW, to analyze if FW occurrence affects the risk level of a FS, and vice versa.

Use Cases

With the system-suggested risk levels and sentiments, FRIDAYS enables financial professionals to quickly navigate through sentences, considering for example higher risk and litigious words that would require more attention. Moreover, the advanced filter could be of use to both accounting educators and researchers to further study the usage of FWs and FMWEs and their contexts.

For example, the authors of the financial dictionary proposed in (Loughran and McDonald 2011) claim that the occurrence of positive words in financial reports might be not significantly correlated with excess returns; therefore, researchers and educators attempt to examine the causes of this phenomenon. With the advanced filter, one of the known situations where positive words fail to work, i.e., high-risk sentences formed by positive words, can be easily examined. Take sentence 235 of the company **SAUCONY INC**’s annual report in 2003 for example, which reads: ‘*A technological breakthrough (pos.) or marketing or promotional success (pos.) by one of our competitors could adversely (neg.) affect our competitive position.*’ From this sentence, consisting of two positive and one negative FWs but is determined by the model to be high-risk, we detect from the sentence level that the two positive words in fact reveal a negative tone due to their context and regarding Saucony’s business status. Therefore, determining the tone of a sentence merely by counting the number of positive words is likely to lead to biased conclusions; similar scenarios usually require further investigation. FRIDAYS would be the ideal system to equip users with a better understanding and provide a clearer illustration.

Acknowledgments

This research was partially supported by the Ministry of Science and Technology of Taiwan under the grants MOST 107-2628-E-001-006-MY3 and MOST 107-2218-E-009-052.

References

- Joulin, A.; Grave, E.; Bojanowski, P.; and Mikolov, T. 2017. Bag of tricks for efficient text classification. In *EACL*, 427–431.
- Liu, Y.-W.; Liu, L.-C.; Wang, C.-J.; and Tsai, M.-F. 2018. Riskfinder: A sentence-level risk detector for financial reports. In *NAACL*, 81–85.
- Loughran, T., and McDonald, B. 2011. When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The Journal of Finance* 66(1):35–65.
- Schneider, N., and Smith, N. A. 2015. A corpus and model integrating multiword expressions and supersenses. In *NAACL*, 1537–1547.