Evaluating Sentiment Changes in BIST Dividend Index Excluded Companies: A DistilBERT Analysis Approach

Dr. Abdullah Kürşat Merter 0000-0001-6874-1890
Dr. Yavuz Selim Balcıoğlu 0000-0001-7138-2972
Research assistant Sedat Çerez 0000-0002-6443-6319
Prof. Dr. Gökhan Özer 0000-0002-3255-998X

Abstract

The purpose of this study is to investigate whether the annual reports of companies that were excluded from the BIST Dividend Index in 2022 have shifted from positive to negative sentiment. To accomplish this, a sophisticated sentiment analysis was conducted using a DistilBERT model, a variant of the BERT model that is pre-trained on financial datasets. The study involved a meticulous data preparation process where texts were extracted from 12 annual activity reports of these companies, six from 2021 and another six from 2022. After the text extraction process, the data was thoroughly cleaned to remove unnecessary information and standardize the format so that it would be suitable for analysis. The 2021 reports exhibited an average sentiment score of +0.40, suggesting a predominantly positive sentiment. In contrast, the 2022 reports showed a negative shift, with an average sentiment score of -0.175. This shift from positive to negative sentiment is particularly noteworthy, as it mirrors the companies' removal from the BIST Dividend Index. The study also explored the correlation between the sentiment of the reports and the market performance of the companies. In general, an average decrease of 0.575 points was observed in the annual reports of companies that were included in the Dividend Index in 2021 but were not included in 2022. These findings were presented with a clear distinction between optimistic performance in 2021 and downturn in 2022.

1   Introduction

Since each business is different from each other, investors need access to as much information as possible about the activities of the relevant business in their decision making. This detailed information is expressed in qualitative rather than quantitative form (Hajek et al., 2014). Understanding these textual expressions helps to assess the financial position of a company, as well as improve its management, and set clear objectives. Annual reports are an important document that provides a detailed overview of a company's financial position and operational performance over a given period of time (Merter et al., 2024). These reports provide a wealth of information about the company's management strategies, operational activities, risks, financial results, and future plans (Stittle, 2003).

As a communication tool between investors and companies in capital markets, annual reports have always been an important tool for all stakeholders (Alduais, 2022). Based on the information presented in annual reports, investors can make decisions about their investments by learning about the risks and financial status of listed companies (Chakraborty and Bhattacharjee, 2020; Li and Zhang, 2014; Soepriyanto e. al., 2021). They provide an important reference point for academics, analysts, investors, and other stakeholders to evaluate the company's performance (Sai e. al., 2019; Li et al., 2018). In general, annual reports contain two different types of data: quantitative and qualitative (Hájek and Olej, 2013). Quantitative data represent accounting and finance-based data representing the firm's financial indicators, while qualitative data include textual information about the firm's organizational structure, forward-looking strategies, and activities.

In modern financial markets, sentiment analysis research has gained importance as a technique to understand investor behavior and predict changes in financial asset prices (Dwivedi et al., 2021). In particular, the abundance of textual data from various sources, including news websites, social media platforms, corporate reports, and financial news, provides a fertile environment for the creation and use of sentiment analysis techniques. In this context, traditional financial analysis approaches have gained a new dimension due to the effective use of artificial intelligence and machine learning techniques in assessing sentiment changes in financial markets. Özer et al. (2023), in their study based on the sustainability reports of companies operating in the Borsa Istanbul Sustainability 25 index, mentioned that corporate sustainability performance can be predicted more quickly and accurately with the artificial intelligence text analysis method.

Due to the topic of the study, the research sample consists of firms that make dividend payments. There are two indices in Borsa Istanbul according to the dividend payments of firms: BIST Dividend and BIST Dividend 25. While the BIST Dividend index includes firms that have distributed cash dividends in the last three years, the BIST Dividend 25 index consists of firms with both high dividend yields and market capitalization (Mazgit, 2013; Şit, 2021). Therefore, the BIST Dividend 25 index includes more liquid firms compared to the BIST Dividend index. To obtain more accurate analysis results, the study was conducted on the BIST Dividend Index in order to determine a larger sample.

Understanding investors' views and emotions is crucial, and sentiment analysis is particularly useful in predicting the performance of stocks of companies listed on major stock exchanges such as Borsa Istanbul (BIST). This study aims to analyze the sentiment changes observed in the annual reports of companies that were listed on the BIST Dividend Index in 2021 but dropped out of the index in 2022. The purpose of this analysis is to understand the changes in the financial performance of these organizations and to determine how these changes can be understood in light of the emotional tones in their annual reports. To this end, we hope to improve the ability of financial analysts and investors to predict the future performance of these companies and make more informed investment decisions. We use a DistilBERT-based analysis approach to assess sentiment changes in companies. DistilBERT represents a scaled-down version of the BERT model.

Following the first introduction, the second part of the study includes previous studies on the subject. The third section explains the methodology, including the sample and the methodology used in the study. The fourth section presents the empirical findings obtained as a result of the analysis, and the fifth and final section concludes the study with a general evaluation.

2   Literature Review

The literature reports a strong relationship between textual data, corporate performance, and stock returns. Previous research that emphasizes the importance of sentiment in annual reports has looked at how financial market ratios such as volatility, trading volumes, and stock prices are affected by sentiment.

Li (2006) examined the effect of risk-related sensitivity in annual reports on stock returns. It measured the frequency of words related to risk and uncertainty in the activity reports of businesses operating between 1994 and 2005 and evaluated them according to the relationship between them and the earnings of the following year. According to the analysis results, it was concluded that there was a negative change in the next year's earnings of companies with high-risk sensitivity. Hájek et al. (2014), in their study of US companies, observed that sentiment scores in annual reports are one of the important determinants of financial performance. Sai et al. (2019) examined the relationship between emotional tone and financial performance of a total of 36 annual reports of 12 companies operating in India between 2015 and 2018. According to the results obtained, they stated that the emotional tone of the annual reports is related to the future performance of the company and that these unstructured texts can be used effectively by stakeholders.

Myšková and Hájek (2020) examined the impact of the emotional tone of risk-related words in annual reports on firm performance. As a result of the analysis, they observed that risk-related emotional tone is an important factor in predicting financial performance and that proactive and opportunistic risk management has a positive effect on financial performance. Studying 102 EU companies and the relationship between corporate social responsibility disclosures in annual reports and firm size, Mućko (2021) observed that large companies avoid words with a negative tone and reports with a positive tone have a positive relationship with firms' financial performance.

Alduais (2022), who examines the relationship between the readability of annual reports of firms operating in China between 2008 and 2021 and their performance, observes that firms with more readable annual reports are more profitable and have lower agency costs. Iqbal et al. (2023), who examine the emotional tone of the annual reports of 33 banks operating in six Islamic countries between 2006 and 2020 and their performance using the panel data analysis method, observe that there is a strong positive relationship between the tone of the statements in the reports and financial performance. Özer et al. (2023) applied text mining analysis techniques such as sentiment analysis and topic modelling to sustainability reports of companies operating in the BIST Sustainability 25 index in 2021. According to the findings, they observed that 81% of the firms had positive sentiment and 19% had negative sentiment in the relevant period.

3   Methodology

3.1 Sample

The study sample consists of firms in the Borsa Istanbul Dividend Index. There are a total of 107 firms in the BIST dividend index. However, since the topic of the study is the evaluation of the sentiment of the firms exiting the dividend index, the firms that were included in the dividend index in 2021 but not in 2022 were included in the sample. Under this condition, six firms were included in the Dividend Index in 2021 but left the index in 2022. The analysis was carried out on six companies. The annual reports of these firms were obtained from their websites and the Public Disclosure Platform (KAP). The results of the firms are coded from A1 to A6.

3.2 Sentiment Analysis

Sentiment analysis is a natural language processing technique that helps to find and understand the emotional tones of text data (Joshy and Sundar, 2022). Moreover, sentiment analysis is the process of identifying emotional states, drawing conclusions, and categorizing emotional expressions (Tam et al., 2021). It is used to identify emotional states and extract the meaning of emotional content found in text data.

Artificial intelligence and machine learning techniques are frequently used in sentiment analysis. Text data are processed by natural language processing algorithms that are used to identify emotional tones. These algorithms use elements such as sentence structure, meaning inference, and the identification of emotional expressions to classify and analyze emotional content.

In general, in sentiment analysis, the emotional content of texts is divided into three categories: positive, negative, and neutral (Mućko, 2021). For example, when analyzing a company's annual report, a positive weighted result as a result of sentiment analysis means that the company evaluated the relevant activity period well and achieved positive results. In contrast, a negative result indicates that the firm experienced difficulties and faced negative situations during the operating period. In addition, in some studies, other sub-emotions (sadness, surprise, etc.) can also be measured.

Sentiment analysis can be applied in various fields, such as financial markets. It is used in many areas, such as marketing, customer relationship management, social media analysis, financial markets, and the health sector. For example, sentiment analysis of social media comments on a brand's products can be used to obtain information about brand reputation, or sentiment analysis of news articles in financial markets can be used to evaluate their impact on stock prices.

Before using the DistilBERT algorithm to perform sentiment analysis in annual reports, the appropriate data set must be collected and prepared. The first step in this process is to obtain annual reports from company websites or financial reporting platforms. Then, pre-processing procedures are applied to the collected reports. The data set is prepared by performing operations such as stopping word removal, word stemming, and text cleaning. The DistilBERT algorithm is then used to label the sentiment of each activity report in the prepared data set. In this step, the texts are categorized by sentiment (positive, negative, or neutral). Once the results are obtained, they are interpreted, and additional analysis can be performed. Examples include whether a particular emotional state is related to financial performance or whether emotional trends can be detected in a particular industry.
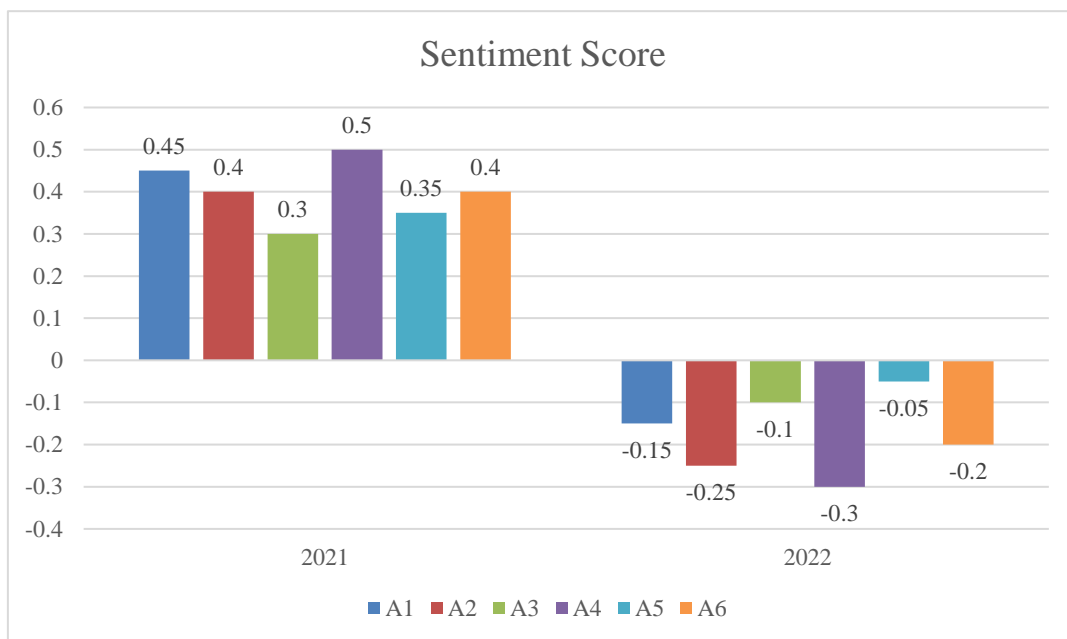

Figure 1. *Sentiment Scores of Annual Reports for 2021 and 2022*

4   Research Findings

The sentiment scores for each firm are given below with explanations.

A1: The A1 report in 2021 showed a positive sentiment score of +0.45, reflecting optimism in the telecommunications and digital transformation sectors. In 2022, the sentiment score decreased to -0.15, reflecting concerns about rising operating costs and competitive pressures.

A2: The A2 report in 2021 reflected positivity (+0.40), supported by strong trends in the real estate market. The change in 2022 decreased to -0.25, with volatility in the real estate market and regulatory changes affecting the sector.

A3: It showed a modest positive sentiment (+0.30) in 2021, driven by advances in chemical manufacturing processes. In 2022, the sentiment score decreased to -0.10, indicating concerns associated with difficulties in the availability of raw materials and environmental regulations.

A4: 2021 sentiment score is highly positive (+0.50), with successes in renewable energy projects. In 2022, the sentiment score drops to -0.30, driven by volatility in energy prices and project delays.

A5: In 2021, it showed a stable positive sentiment (+0.35) thanks to innovations in glass manufacturing. The sentiment score for the 2022 report decreased to -0.05, indicating minor concerns about export markets and exchange rates.

A6: Maintained a positive sentiment (+0.40) in 2021 with successful engineering and construction projects. In 2022, the sentiment score dropped to -0.20, indicating cost overruns and project delays.

Figure 1 was created below to more clearly observe the changes in sentiment in the annual reports of the firms that were included in the BIST Dividend Index in 2021 but not in 2022.

As can be seen in Figure 1, there is a significant decrease in sentiment scores in the annual reports of all firms that were included in the dividend index in 2021 but not in 2022. This indicates that firms that dropped out of the index faced negative situations during the relevant activity period. The largest decline among the firms belongs to A4 with a 0.8-point decrease.

5   Conclusions

Since each business is different, investors should have access to as much information as possible about the operations of that business to make an informed decision. This detailed information is expressed qualitatively and quantitatively. Understanding these textual statements helps to assess a company's financial position, improve management, and set specific goals (Hajek et al., 2014). Annual reports are an important document that provides a comprehensive summary of an organization's financial position and operational performance over a given period of time. The company's management strategies, operational activities, risks, financial results, and future plans are discussed in detail in these reports (Merter et al., 2023).

The purpose of this study is to analyze the emotional changes in the annual reports of companies that were included in the BIST Dividend Index in 2021 but were removed from the index in 2022. The purpose of this analysis is to understand the changes in the financial performance of these companies and to determine how these changes can be explained by the emotions of their annual reports. In doing so, we aim to improve the ability of investors and financial analysts to predict the future performance of these companies and make better informed investment decisions. We use a DistilBERT-based analytical technique to assess sentiment shifts in businesses.

The sentiment analysis clearly shows a significant shift from optimism in 2021 to pessimism in 2022 among the companies excluded from the BIST dividend index. This sentiment shift is closely aligned with their delisting and reflects the broader market and business challenges faced by these companies during these periods. The findings highlight the potential of advanced sentiment analysis techniques such as DistilBERT to provide deep insights into market and company performance based on textual analysis of corporate disclosures.

References

- Alduais, F. 2022. Textual Analysis of the Annual Report and Corporate Performance: Evidence from China. Journal of Financial Reporting and Accounting.

- Chakraborty, B., & Bhattacharjee, T. 2020. A Review on Textual Analysis of Corporate Disclosure According to the Evolution of Different Automated Methods. Journal of Financial Reporting and Accounting, **184**, p. 757.

- Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., . . . Williams, M. D. 2021. Artificial Intelligence AI: Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. International Journal of Information Management, **57**, 101994.

- Hájek, P., & Olej, V. 2013. Evaluating Sentiment in Annual Reports for Financial Distress Prediction Using Neural Networks and Support Vector Machines. Communications in Computer and Information Science.

- Hajek, P., Olej, V., & Myskova, R. 2014. Forecasting Corporate Financial Performance Using Sentiment in Annual Reports for Stakeholders' Decision-Making. Technological and Economic Development of Economy, **204**, p. 721.

- Joshy, A., & Sundar, S. 2022. Analyzing the Performance of Sentiment Analysis using BERT, DistilBERT, and RoBERTa. IEEE International Power and Renewable Energy Conference IPRECON.

- Li, F. 2006. Do Stock Market Investors Understand the Risk Sentiment of Corporate Annual Reports? SSRN.

- Li, Y., & Zhang, L. 2014. Short Selling Pressure, Stock Price Behavior, and Management Forecast Precision: Evidence from a Natural Experiment. Journal of Accounting Research, **531**, p. 79.

- Li, Z., Wang, B., Fu, Y., Shi, Y., & Su, X. 2018. Different Types of Investor Reactions to Annual Reports. Emerging Markets Finance and Trade, **563**, p. 626.

- Mazgit, İ. 2013. Endeks Kapsamında Olmanın Hisse Senedi Getirilerine Etkisi: BIST Temettü 25 Endeksi Üzerine Bir Uygulama. Sosyoekonomi, **2020**, p. 225.

- Merter, A. K., Balcıoğlu, Y. S., Çerez, S., & Özer, G. 2024. The Role of the Bert Algorithm In Unraveling Brand Value Declines: A Study On Annual Reports Of Turkish Companies. 7. International Marmara Scientific Research and Innovation Congress p. 449. İstanbul: International Science and Art Research Center.

- Merter, A. K., Çerez, S., Balcıoğlu, Y. S., & Özer, G. 2023. **Topic Modelling of Annual Reports of Borsa Istanbul**. Empirical and Theoretical Studies in Economics and Finance, p. 125.

- Mućko, P. 2021. Sentiment Analysis of CSR Disclosures in Annual Reports of EU Companies. Procedia Computer Science, **192**, p. 3351.

- Myšková, R., & Hájek, P. 2020. Mining Risk-Related Sentiment in Corporate Annual Reports and Its Effect On Financial Performance. Technological and Economic Development of Economy, **266**, p. 1422.

- Özer, G., Balcıoğlu, Y. S., Merter, A. K., & Çerez, S. 2023. Metin Analizi Yöntemi ile Kurumsal Sürdürülebilirlik Raporlarının İncelenmesi: BİST Sürdürülebilirlik 25 Endeksinde Bir Uygulama. Social Science Development Journal, **838**, p. 178.

- Sai, P. K., Gupta, P., & Fernandes, S. F. 2019. Analysing Performance of Company through Annual reports using Text Analytics. International Conference on Digitization ICD, p. 21.

- Soepriyanto, G., Tjokroaminoto, S., & Zudana, A. E. 2021. Annual Report Readability and Accounting Irregularities: Evidence from Public Listed Companies in Indonesia. Journal of Financial Reporting and Accounting, **195**, p. 793.

- Stittle, J. 2003. **Annual reports: Delivering Your Corporate Message to Stakeholders**. Gower Publishing, Ltd.

- Şit, A. 2021. Kar Dağıtım Politikaları Firma Değeri Üzerinde Etkili Midir? BIST- Temettü 25 Endeksi Üzerine Bir Uygulama. Van Yüzüncü Yıl Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, **612**, p. 159.

- Tam, S., Said, R. B., & Tanriöver, Ö. Ö. 2021. A ConvBILSTM Deep Learning Model-Based Approach for Twitter Sentiment Classification. IEEE Access, p. 41283.