Are Delayed Annual Reports a Sign of Bad News? A Predictive Analysis Using Machine Learning and Sentiment Analysis

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

This research offers a pioneering investigation into the potential implications of late annual report announcements in the context of BIST100 companies. Annual reports serve as a pivotal source of comprehensive information, enabling stakeholders to evaluate a company's financial health, growth prospects, operational strategies, and potential risks. Recognizing this, our study harnesses artificial intelligence (AI) to scrutinize the timeliness of annual report announcements and their inherent sentiment. Using AI, we constructed a predictive model based on the announcement dates of annual reports between 2009 and 2021. The model estimated the release dates for 2022, facilitating the identification of companies that released their reports later than expected. These delayed reports were then subjected to an in-depth text mining, sentiment analysis, and schema analysis. Our text mining process utilized the robust Hidden Dirichlet Allocation (LDA) Subject Salient Terms (TST) method, known for its efficacy in revealing concealed topics in large text volumes. Our findings were striking; around 87% of statements in the delayed annual reports reflected a negative sentiment, while only 13% displayed a positive tone. Thus, late annual report announcements tend to have a generally pessimistic outlook, indicating they might indeed be bearers of adverse news. This research offers a unique perspective on the relationship between the timeliness of financial reports and the contained sentiment, ultimately contributing to the enhanced transparency and informed decision-making in financial markets. The study underscores the necessity for companies to maintain timely communication and suggests potential areas of interest for future investigations.

1 Introduction

In the prevailing data-centric environment, proficiency in extracting actionable insights from voluminous datasets has become indispensable, particularly within the financial sector. The timeliness and accuracy of the information have a significant impact on the distinction between failure and financial success. In this context, annual reports emerge as crucial documents, furnishing a holistic view of a company's financial health, growth potential, strategic initiatives, and associated risks (Stittle, 2003). The timeliness of annual reports, along with the conveyed sentiment, holds significant value for research purposes (Huang et al., 2013). These documents are essential for investors, analysts, and a broad spectrum of stakeholders, who rely on them for informed decision-making. Thus, the prompt dissemination of these reports is imperative for ensuring market stability and preserving investor confidence.

The delay in the publication of an annual report can precipitate concern and engender skepticism among market participants, indicative of concealed operational or financial challenges within the company (Boyne and Law, 1991). These challenges might encompass operational inefficiencies, financial instabilities, regulatory compliance issues, or substantial shifts in management or strategic direction, necessitating a prolonged period for thorough disclosure. Such delays hamper the seamless transmission of information, vital for the efficacy and transparency of financial markets, and consequently erode investor trust. This erosion is perceived as a departure from established corporate governance practices, possibly portending negative developments. In the era of digital transformation, advancements in artificial intelligence (AI) have unveiled innovative avenues for examining the narratives within these postponed disclosures (Dwivedi et al., 2021). The use of sentiment analysis and machine learning techniques has made it possible to find patterns, get important insights, and judge the emotional tone of these reports, which has helped us learn more about what these delays mean.

This research embarks on an exploratory endeavor to ascertain the potential effects of delays in the publication of annual reports amongst companies listed on the BIST100 index, utilizing the advancements in artificial intelligence (AI) for predictive analysis. By constructing a predictive model anchored in the historical announcement dates of annual reports spanning from 2009 to 2021, this study aspires to project the publication timelines for 2022, identifying firms that deviated from expected release schedules. This effort not only enriches the existing academic literature by scrutinizing the sentiment and thematic content of delayed reports via sophisticated AI methodologies but also enhances our understanding of the correlation between the timing of these disclosures and the sentiment embedded within. With a predominant share of statements in these deferred reports manifesting a negative sentiment, the results suggest that such delays are likely indicative of unfavorable news, casting light on the intricate relationship between corporate disclosure practices and investor sentiment. This investigation not only fills a void in academic discourse but also provides investors, regulatory bodies, and corporate executives within the BIST100 context with actionable intelligence, steering towards a regime of increased transparency and informed decision-making in the financial markets.

2 Literature Review

The role of corporate annual reports in informing stakeholders and the wider public has been an area of significant research interest in the field of accounting and finance. Annual reports represent an essential communication tool between a firm and its stakeholders, serving to provide a comprehensive picture of a company's financial health, strategies, and prospects (Beattie et al., 2004). The timeliness of the information communicated through annual reports is recognized as a key factor in its perceived relevance and usefulness. According to the Efficient Market Hypothesis, the value of information decreases as it ages due to its rapid incorporation into share prices (Fama, 1970). In line with this, Lang and Lundholm (1996) emphasized the importance of timely reporting in maintaining investor confidence and market stability. Delayed release of annual reports has attracted scholarly attention due to the potential negative implications for stakeholder perception and company performance. For instance, Bhattacharya et al. (2017) found that delayed financial reporting can induce uncertainty among stakeholders and increase information asymmetry, potentially leading to higher trading costs. Moreover, delays may be interpreted as a sign of managerial inefficiency or possible financial distress, negatively impacting a firm's reputation (DeFond et al., 2015).

Recent advances in artificial intelligence (AI) have enabled more sophisticated approaches to analyzing corporate disclosures. Among them, text mining techniques, such as the Hidden Dirichlet Allocation (LDA) model, have been successfully applied to detect underlying topics in large volumes of text (Blei et al., 2003). Moreover, sentiment analysis, a technique originating from computational linguistics, has been employed to detect and quantify the emotional content of these disclosures (Loughran and McDonald, 2011). Hájek (2013) found that sentiment analysis improved the accuracy of neural networks and support vector machines in predicting future financial distress. Chiong (2018) found that sentiment analysis and support vector machines predicted stock market trends better than deep learning. Khan (2020) found that sentiment analysis and political situation features improved the accuracy of machine learning algorithms in predicting Pakistan's stock market by up to 20% over 7 days. However, Seals (2020) found that sentiment analysis only improved stock forecasting accuracy by 0-3% over 7 days.

The research on this topic provides mixed evidence on whether delayed annual reports signal bad news to investors. Some studies suggest delayed reports indicate poor performance and negative market reactions. Bartov (2013) found the market reacts more negatively to quarterly report delays than annual report delays, indicating deeper problems. Duarte-Silva (2013) showed delay announcements produce an average 6% stock drop, especially when due to accounting issues, signaling deteriorating performance. Haw (2000) and Cheng (2006) found good news firms release reports earlier while loss firms delay, and the market reacts accordingly. However, other research found no systematic relation between report timing and earnings news. Wang (2006) documents significant improvements in Chinese report timeliness, though bad news still meant delays. Waymire (1984) found no significant stock drops for "bad news" management forecasts.

3 Methodology

Utilizing the historical data of annual report announcements by companies listed on the BIST100 index between 2009 and 20211, a sophisticated predictive model, harnessing the capabilities of machine learning algorithms, was developed to forecast a specific 15-day period anticipated for the publication of their annual reports in the year 2022. The model's design was such that it allowed for the identification of deviations from the normative reporting schedule, thus enabling a systematic classification. Entities that exceeded this forecasted timeframe, failing to issue their annual reports within the designated 15-day window, were systematically categorized under the designation "Late annual report-releasing companies". The 2022 annual reports of these companies were manually retrieved from the public disclosure platform. Table 1 shows the annual report words and characters count for each company.

REPORTING YEAR		***2022***
ŞİRKET	***Word Count***	***Character Count***
EREGL	15733	124665
ISDMR	12911	102233
OYAKC	13304	101503
SASA	14530	116339
SOKM	47417	379977
TCELL	146215	1126756
TTKOM	111187	860250
VESBE	77012	609653
TOTAL	**438309**	**3421376**

Table 1. *Word Count in Annual Reports*

4 Findings

Analyzed the sentiment of annual reports using the NRCLex library in Python program. Figure 1 displays the sentiment levels of the annual reports of companies who publish their annual reports late.

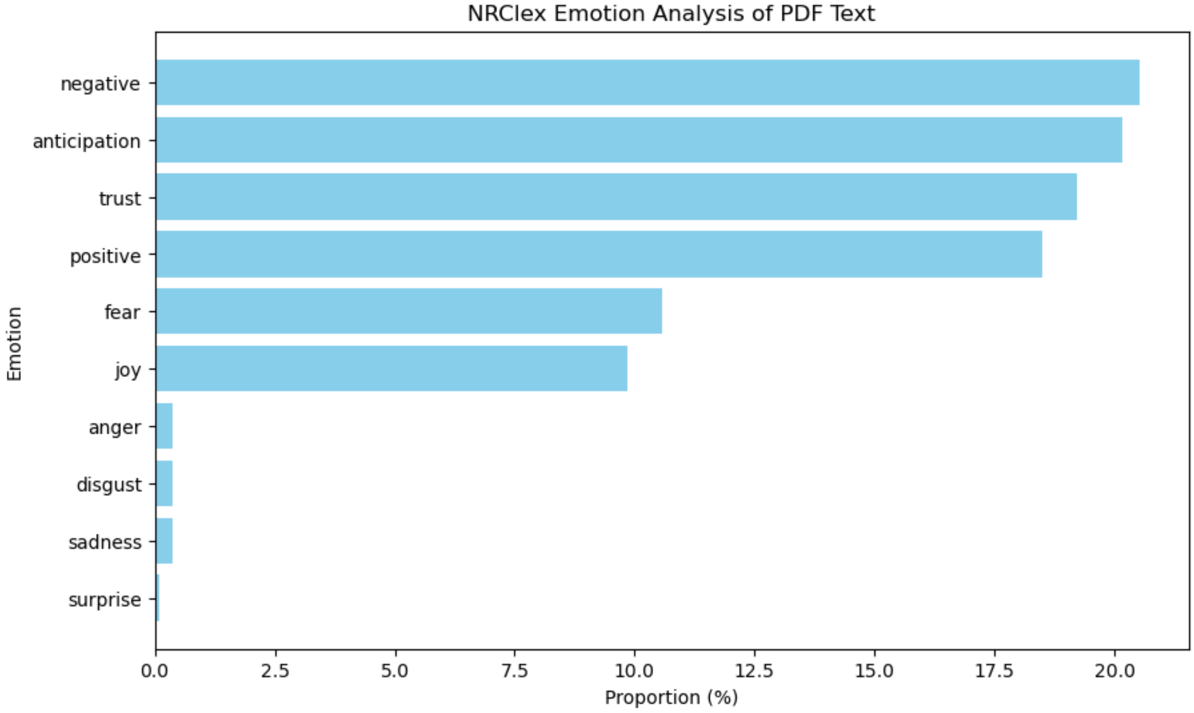

Figure 1. *Sentiment analysis of annual reports*

Sentiments	Percentage
Negative	20.1
Anticipation	19.2
Trust	17.9
Positive	17.6
Fear	10.6
Joy	9.5
Anger	2.5
Disgust	2.4
Sadness	0.2
Surprise	0.0

Table 2. *Late annual report-releasing companies*

The data in Table 2 represents the distribution of various sentiments within the "Late annual report-releasing companies " category. The analysis of this data reveals several insights:

- **Negative Emotions**: The dominant emotion is "Negative," constituting 20.1% of the emotional content. This prominent presence of negative emotions may indicate concerns, challenges, or areas of dissatisfaction that are highlighted in the reports.
- **Positive Emotions**: Positive emotions, such as "Anticipation" (19.2%), "Trust" (17.9%), and "Positive" (17.6%), are also significant. The close proportions of these emotions suggest a balanced representation of optimism, confidence, and future expectations within the reports.
- **Fear and Joy**: "Fear" is present at 10.6%, and "Joy" is at 9.5%. The presence of fear may reflect underlying uncertainties or risks, while joy may indicate specific achievements or positive developments.
- **Lower Prevalence Emotions**: "Anger" (2.5%) and "Disgust" (2.4%) are relatively low, indicating that these emotions may not be central to the thematic content. "Sadness" is almost negligible at 0.2%, and "Surprise" is absent, indicating a lack of unexpected elements or a minimal focus on sorrowful aspects.
- **Overall Interpretation**: The distribution of emotions in "Geç Açıklayanlar" paints a complex emotional landscape. While there are significant negative emotions, positive emotions like trust, anticipation, and

positivity are also prominent. This combination may reflect a nuanced understanding of the situation, where challenges and opportunities coexist.

- **Contextual Consideration**: To fully comprehend the implications of these emotions, further contextual analysis would be necessary. Understanding the nature of the reports, the specific content, and the context in which these emotions are expressed would provide a more comprehensive interpretation.

5 Conclusion

This study is a fundamental investigation of the consequences of late annual report releases in BIST100 listed companies. Artificial intelligence was used to examine the timeliness and and sentiment of these disclosures. Through constructing a predictive model based on historical disclosures dates from 2009 to 2021, the study identified companies releasing their reports outside a forecasted 15-day date range in 2022 as "Late annual report-releasing companies." Subsequent sentiment analysis revealed a predominantly negative sentiment in these delayed reports, with approximately 20.1% of statements reflecting a negative emotion. Findings suggests a relationship between the delay in annual report releases and the likelihood of such reports harboring unfavorable news.

5.1 Practical Implications

This research emphasizes the significance of companies to provide timely communication to maintain investor confidence and ensure market stability. The evidence suggesting a negative sentiment in delayed reports could serve as a red flag for investors, analysts, and regulatory bodies, prompting a closer examination of companies with lagged disclosures. Moreover, using AI to analyze annual reports provides stakeholders with a model for more efficiently evaluating corporate transparency and decision-making processes. Companies within the BIST100 index could leverage these insights to enhance their reporting strategies and mitigate the adverse effects of lagged disclosures.

5.2 Theoretical Implications

The study contributes theoretically to the growing body of research on the relationship between market perception, company emotions expression, and timely annual reporting. The integration of advanced AI methodologies for predictive analysis and sentiment examination bridges existing gaps in understanding the perception of delayed financial disclosures and their potential implications. This study provides up opportunities for more research into the predictive potential of sentiment analysis in annual reporting and highlights the significance of timely and transparent communication in the financial markets. Future studies could extend this research by examinig the causative factors behind delayed reports, the sector-specific impacts of such delays, and the long-term effects on company performance and stock valuation.

References

- Bartov, E., DeFond, M. L., Konchitchki, Y., 2013. “The consequences of untimely quarterly and annual financial reporting”, Working Paper, New York University
- Beattie, V., McInnes, W., Fearnley, S., 2004. “A methodology for analysing and evaluating narratives in annual reports: a comprehensive descriptive profile and metrics for disclosure quality attributes”, *Accounting forum*, **28(3)**, p. 205-236.
- Bhattacharya, N., Desai, H., Venkataraman, K., 2017. “Does earnings quality affect information asymmetry? Evidence from trading costs”, *Contemporary Accounting Research*, **34(2)**, p. 1072-1103.
- Blei, D. M., Ng, A. Y., Jordan, M. I., 2003. “Latent dirichlet allocation. Journal of machine Learning research”, **3**, p. 993-1022.
- Boyne, G., Law, J., 1991. “Accountability and local authority annual reports: The case of welsh district councils”, *Financial Accountability & Management*, **7(3),** p. 179-194.
- Cheng, F. F., 2006. “Timeliness of annual report releases in relation to the direction and magnitude of earnings and share price evaluation: the case of Malaysia”, *Sunway Academic Journal*, **3**, p. 117-131.
- Chiong, R., Fan, Z., Hu, Z., Adam, M. T., Lutz, B., Neumann, D., 2018. “A sentiment analysis-based machine learning approach for financial market prediction via news disclosures”, in Proceedings of the genetic and evolutionary computation conference companion, p. 278-279
- DeFond, M. L., Hung, M., Li, S., Li, Y., 2015. “Does mandatory IFRS adoption affect crash risk?”, *The Accounting Review*, **90(1)**, p. 265-299.
- Duarte-Silva, T., Fu, H., Noe, C. F., Ramesh, K., 2013. “How do investors interpret announcements of earnings delays?”, *Journal of Applied Corporate Finance*, **25(1)**, p. 64-71.

- 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,** p. 1-47.
- Fama, E. F., 1970. "Efficient capital markets: A review of theory and empirical work", *The journal of Finance*, **25(2)**, p. 383-417.
- Haw, I. M., Qi, D., Wu, W., 2000. "Timeliness of annual report releases and market reaction to earnings announcements in an emerging capital market: The case of China", *Journal of International Financial Management & Accounting*, **11(2),** p. 108-131.
- Hájek, P., 2018. "Combining bag-of-words and sentiment features of annual reports to predict abnormal stock returns", *Neural Computing and Applications*, **29**, p. 343-358.
- Huang, A. H., Zang, A. Y., Zheng, R., 2014. "Evidence on the information content of text in analyst reports", *The Accounting Review*, **89(6),** p. 2151-2180.
- Khan, W., Malik, U., Ghazanfar, M. A., Azam, M. A., Alyoubi, K. H., Alfakeeh, A. S., 2020. "Predicting stock market trends using machine learning algorithms via public sentiment and political situation analysis", *Soft Computing*, **24**, p. 11019-11043.
- Lang, M., Lundholm, R., 1996. "Corporate disclosure policy and analyst behavior", *The Accounting Review*, **71(4)**, p. 467-492.
- Loughran, T., McDonald, B., 2011. "When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks", *The Journal of Finance*, **66(1)**, p. 35-65.
- Seals, E., Price, S. R., 2020. "Preliminary Investigation in the use of Sentiment Analysis in Prediction of Stock Forecasting using Machine Learning", in SoutheastCon, p. 1-2
- Stittle, J., 2003. "**Annual reports: Delivering your corporate message to stakeholders**", Gower Publishing, Ltd.
- Wang, J., Song, L., 2006. "Timeliness of annual reports of Chinese listed companies", *Journal of Chinese economic and business studies*, **4(3)**, p. 241-257.
- Waymire, G., 1984. "Additional evidence on the information content of management earnings forecasts", *Journal of Accounting Research*, **22(2),** p. 703-718.