Sentiment Analysis of Chairperson's Message and Its Influence on Financial Performance : Study on NIFTY 50 Companies

Manoj Kumar Jena [1]
Srikant Das [2]
Brajaballav Kar [3]

Abstract

Purpose : Return on assets (ROA) is one of the most important financial parameters used for multi-level analysis. The investment in assets also depends on the future perspective of the management, which is reflected in the chairperson's message in the annual reports. Thus, the chairperson's message is likely to indicate future investments in assets and expected returns from such assets. The message is important for the investors as this is one of the major sources of qualitative information about the business strategy. However, the information the business managers share may include selective information from a strategic perspective. Are there any tendencies in the language used in the message by the chairpersons as well as the accompanying sentiments? How does a company's overall perception evolve? How has the worldwide epidemic influenced chairpersons' messages?

Design/Methodology/Approach : The study investigated these questions using the information from the annual reports of the top 50 companies (NIFTY 50) listed in the National Stock Exchange of India. The methodology of this study is mostly grounded on text mining, text analytics, and regression analysis.

Findings : The negative tones in the chairperson's speech did not show much significance for the future performance of NIFTY50 companies. However, the positive sentiments did have an impact and were significantly associated with the ROA of the companies.

Practical Implications : The paper would add value to the academicians, researchers in the area of NLP and text analytics, and the investors' community who access the annual reports for making investment decisions.

Originality : The use of text analytics to understand sentiments and apply them to elicit a deeper understanding of the company's strategy is a novel attempt.

Keywords : text analytics, text mining, chairman message, corporate governance, signaling theory

JEL Classification Codes : G17, G14, E37, E70

Paper Submission Date : September 25, 2023 ; **Paper sent back for Revision :** May 2, 2024 ; **Paper Acceptance Date :** August 20, 2024 ; **Paper Published Online :** November 15, 2024

[1] *Assistant Professor (Corresponding Author),* KIIT School of Management, Campus - 7, KIIT University, Bhubaneswar - 751 024, Odisha. (Email : manoj.jena@ksom.ac.in) ; ORCID iD : https://orcid.org/0000-0001-6409-8033
[2] *Associate Professor,* KIIT School of Management, Campus - 7, KIIT University, Bhubaneswar - 751 024, Odisha. (Email : srikant@kiit.ac.in) ; ORCID iD : https://orcid.org/0000-0003-4379-4384
[3] *Associate Professor,* KIIT School of Management, Campus - 7, KIIT University, Bhubaneswar - 751 024, Odisha. (Email : brajkar@gmail.com) ; ORCID iD : https://orcid.org/0000-0002-2127-1147

DOI : https://doi.org/10.17010/ijf/2024/v18i11/174639

The Companies Act, 2013 mandates certain disclosures for listed companies such as financial performance, state of company's affairs, dividends, subsidiary details, loans and investments, change in nature of business, change in share capital, board composition, statutory auditors' remarks, consolidated financial statements, and directors' responsibility. The disclosures in the directors' report are mandated for greater transparency in corporate governance. This is useful for shareholders to access unbiased information about the business. Apart from the above disclosures, companies publish the chairman's letter to the shareholders which includes the strategic discussions from the head of the organization.

It is perceived that the chairman's message to shareholders includes information about the prospects of the business in an unbiased manner. However, the unbiasedness of the disclosed texts is examined by researchers from various perspectives. It is observed that CEOs attempt to influence readers through different rhetoric and project a positive corporate image in company annual reports (Kar & Das, 2022; Merkl-Davies & Koller, 2012; Smith & Taffler, 1992). Moreover, the amount of textual material may not provide transparency. Furthermore, from a reader's standpoint, assimilation of information while comparing year-on-year statistics across numerous papers becomes extremely hard and vulnerable to prejudice. With a large volume of qualitative messages, it is crucial to uncover the hidden insights within the text and see how they align with quantitative data from a forward-looking perspective. This approach enables capital market participants to enhance their understanding, leading to more informed decision-making.

Research Gap

The chairman's message to the shareholders and its correlation to firm performance has not been sufficiently discussed in the literature. One of the major reasons for inadequate research on the area is perhaps due to the unstructured data type of the verbal information contained in the chairman's report. The issue here is deriving the objective information from unstructured data and bringing an analogy between the same with the structured data in the form of numbers from the financial performance figures of the company. Ha et al. (2019), Ittoo et al. (2016), Jun and Park (2017), Kim et al. (2018), Lacity and Janson (1994), McLellan et al. (2003), Pineda-Jaramillo and Pineda-Jaramillo (2022), Singh and Kalra (2024), Thomas (2014), and Trombeta and Cox (2022) have established various frameworks to develop a computational analysis of textual information in the last decades. Among all the methods to analyze unstructured data, text mining has been used more predominantly for computational text analysis.

Research Aim and Contribution

The present study is focused on deriving the sentiments in the chairman's speech using the financial lexicon and text mining method. The sentiments are further compared with the financial performance of the respective companies under the NIFTY50 index (Indian index comprising 50 companies under the National Stock Exchange). The research will significantly contribute to the investors' and shareholders' community in the capital markets domain to decipher the large volume of qualitative information posted by the company management in an unbiased manner. This will further help in making investment decisions based on strategic information. The text analytics methods with R programming used in the research work will immensely help the academic and researcher community to work in a similar domain more effectively.

Presently, research in financial texts includes the application of analytics to derive unbiased information from a large volume of data. This helps in further data mining operations and derives sentiments in the texts. Our research is related to deriving the sentiments from the chairperson's message using the most recent applications related to text analytics based on lexicon methods. We further compare the sentiments with the financial performance measures, which are not only in trends but also extremely useful in the world of financial data and information.

Literature Review

Background

Corporates publish annual reports consisting of information based on regulatory requirements passed through a due audit process. The report speaks about management priorities, risks, and strategic operations. The annual report is a credible database used by accounting professionals for financial communication between management and shareholders. The research has shown that private and institutional investors have extensively used the chairman's statement or President's letter in annual reports to access business information (Bartlett & Chandler, 1997; Bhuyan & Baid, 2020; Clatworthy & Jones, 2006; Day, 1986; Kar & Jena, 2019; Qiu et al., 2014; Smith & Taffler, 1992). Over four to five decades, it has been observed that the readability of the chairman's address, footnotes to the financial statements, and other verbal parts of the annual report have remained below the fluent comprehension level (Clatworthy & Jones, 2001; Courtis, 1995). The subjective interpretation of the abundance of textual content remains an issue. It leads to the application of text analytics technologies to help the reader deduce the sentiment and information contained in the chairman's speech.

Theoretical Basis

The present research derives the motivation from the aspects of the signaling theory. The theory is fundamentally concerned with reducing the information asymmetry between two parties (Spence, 2002). Signaling theory works when there are two parties involved, and each party has unequal access to information. Typically, the sender chooses how to communicate or signal the information, and the receiver chooses how to interpret the signal (Connelly et al., 2011). In this research, the company chairpersons and management are the senders of information. Investors other than the promoter groups are the ones who interpret the information for making investment decisions.

Further, the signaling hypothesis assumes that managers may use their discretion power over narrative reporting to enhance the signal value of earnings and present a better impression of quality (Al-Sayani & Al-Matari, 2023; Bozos et al., 2011; Klein & Peterson, 1989; Yasar et al., 2020). Thus, the tone or the sentiment of the content in the chairperson's message may serve as an indicator of the financial performance of the company.

Past Research and Findings

The Securities and Exchange Commission of the US requires chief executive officers (CEOs) of large publicly traded companies to certify their financial statements. A study on corporate governance shows how CEOs signal the unobservable quality of their firms to potential investors via the observable quality of their financial statements (Zhang & Wiersema, 2009).

A probit analysis of 106 listed companies examined the impact of corporate size, profitability, solvency, operational complexity, industry, and auditors on compliance with non-mandatory accounting pronouncements. The research showed that companies that are large, profitable, and highly geared are more likely to comply with the non-mandatory accounting pronouncements than their counterparts (Ng & Koh, 1994). Abrahamson and Park (1994) conducted a content analysis of over 1,000 presidents' letters to shareholders to investigate how corporate officials purposely hide poor organizational outcomes from shareholders. The results revealed that major institutional investors, outside directors, and accountants minimized such concealment, while small investors and outside shareholder directors concealed the unfavorable organizational impact.

Different stakeholders, such as investors, analysts, and auditors, use the tone of corporate financial disclosures to infer management's private information on the company's prospects, risks, value, and operational outcomes. The tone of financial disclosures is typically dependent on financial performance, company size and growth, management compensation types, and management characteristics (Luo & Zhou, 2020). The research suggests that firms with better current performance, lower accruals, smaller size, less return volatility, and longer history tend to have more forward-looking statements in the Management Discussion & Analysis section of the annual reports (Li, 2010). According to research, financial graphs and images in annual reports are commonly used for impression management. Company management selectively employs graphs to provide a more favorable perception of financial success (Beattie & Jones, 1999, 1992; Courtis, 1997; Falschlunger et al., 2015; Frownfelter-Lohrke & Fulkerson, 2001).

The literature review reveals that, on a basic level, the better the financial outcome, the more favorable the tone of the disclosures in the narrative portions. However, there remains a vacuum in how textual narratives or disclosures, driven by legal constraints, contribute to transparent representation. While these disclosures aim to enhance transparency, they can also be used as a management tool for impression management. Most research work focuses on companies in the US, UK, and other non-Indian regions. Since the financial reporting environment and disclosure regulations vary widely across countries, it may not be entirely appropriate to generalize the above research findings to the Indian context. This study takes a step forward in adding value to the existing hypothesis within an Indian framework.

Hypotheses Development

As already stated, investors, analysts, and researchers use the chairperson's statement to access the business trends and strategic information related to the company's performance. The relationship between the financial performance and the chairperson's statement is never straightforward, given the former is about numbers and the latter is about the unstructured content of the textual data. To describe the impact of textual content, we shall explore a representative quantitative number in the form of sentiments of the text. The difficulty in finding the impact of textual information is accentuated when we speak about any causal relationship between the statements and performance. This provides a background to develop a model for finding the relation between the sentiments in the texts and the numerical data through various financial parameters. Based on this, we discuss two hypotheses to test as follows:

- **H01 :** The sentiments do not show trends concerning financial performance.
- **Ha1 :** The sentiments show trends concerning financial performance.

Research Methodology

Research Type

The present research includes the descriptive analytics of the qualitative data taken from chairpersons' messages. The research also includes correlation analysis, where the numeric representations of the text data are tested for the correlation with the financial performance of the company.

Data Description

Sentiment analysis is performed by extracting textual information from the chairman's message in the annual reports. The population for this research includes six years of data for the NIFTY 50 companies. The rationale for considering NIFTY 50 companies are as follows: (a) these companies represent the larger companies accounting for 13 sectors of the Indian economy which also represent the major stock index of the country. Thus, the research is not limited to any specific sector, (b) the focus on the larger companies may serve as a benchmark for others, given these companies follow better corporate governance practices and publish the disclosures as per the regulatory requirement, and (c) the NIFTY 50 index represents about 62% of the free float market capitalization of the stocks listed on NSE as on September 30, 2022 (https://www.nseindia.com/products-services/indices-nifty50-index).

The NIFTY 50 index firms are evaluated every six months. To maintain dataset consistency, we used the January 2022 list of 50 firms. Annual reports for each of these companies from fiscal year 2017 (FY17) to FY22 can be accessed from the Bombay Stock Exchange website. The annual reports for FY23 are not yet available during the writing of this article. Hence, the chairpersons' messages for FY23 are not included in the research. Many times, annual reports have issues related to font type, image format, and other word recognition problems. To avoid complications with the annual report's presentations, we copied the information into separate Word files. In some circumstances, the messages are handwritten and saved in a Word file as part of the data-gathering procedure.

The quantitative data for the company's financial metrics are downloaded from the Centre of Monitoring Indian Economy (CMIE) ProwessIQ database. The dataset includes financial characteristics such as the company's return on assets (ROA), debt-to-equity ratio (LEVERAGE), and total assets (SIZE). We collected the aforementioned data for five years, from FY18 to FY22.

Data Cleaning

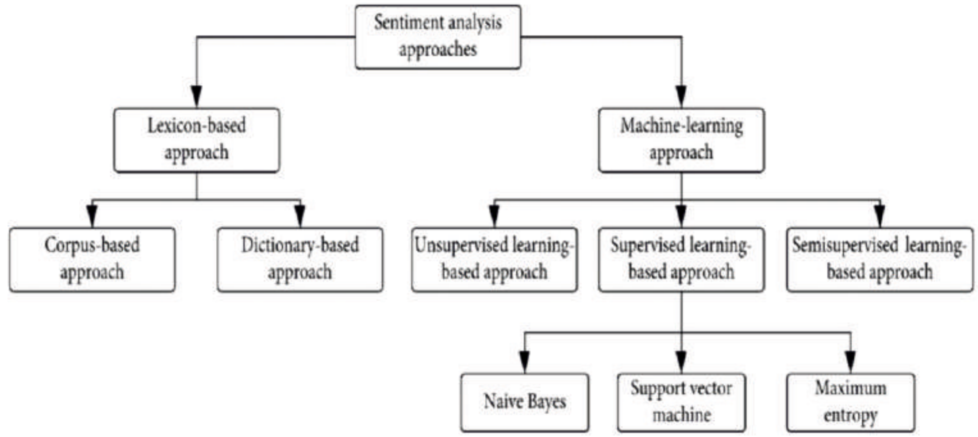
HDFC Bank's annual report for the financial years FY17, FY18, and FY19 did not contain the chairman's message. The director's report statements are taken as a replacement for the chairman's statement for these years of HDFC Bank. Britannia Industries did not publish the chairman's message in the annual report. Hence, the qualitative information from the director's report was extracted for the analysis for the period FY17–FY22. The Annual reports of HCL Technology for FY17, FY18, and FY19 did not contain the chairperson's message. Hence, this was removed from the analysis. However, for FY18, the chairman's speech at the AGM was recorded and noted for analysis. For FY19, passages from the Management Discussion and Analysis section were used in the research. Tech Mahindra did not have a Chairman's speech; thus, the MD and CEO's message was evaluated instead.

Nestle India follows the financial year ending in December. Thus, the annual data ending December 2021 is used as FY22 data, and so on. The organization has been tracking this for the past six years. The data preparation and cleaning stage included many operations, such as merging the financial data with the derived sentiment data, cleaning the column names, and converting the data from long to wide format.

Methodology

Among various text analysis methods, sentiment analysis is widely used to interpret textual information in different sections of annual reports, including corporate social responsibility (CSR) disclosures, auditors' statements, Chairman's letters, and notes to financial statements (Ranjan et al., 2018). Sentiment analysis can be

Figure 1. ***Methods in Sentiment Analysis***



Source : Che et al. (2020).

conducted using either machine learning or lexicon-based approaches. The specific methods within each approach are illustrated in Figure 1.

Lexicon-Based Sentiment Classifier

We have used a lexicon-based approach to extract the sentiment across the chairperson's message. In this approach, each word or token is analyzed with a predefined set of rules to evaluate the sentiment associated with the word. Essentially, the words of the speech are matched with the terms in a selected dictionary, where the terms are categorized according to a type of sentiment.

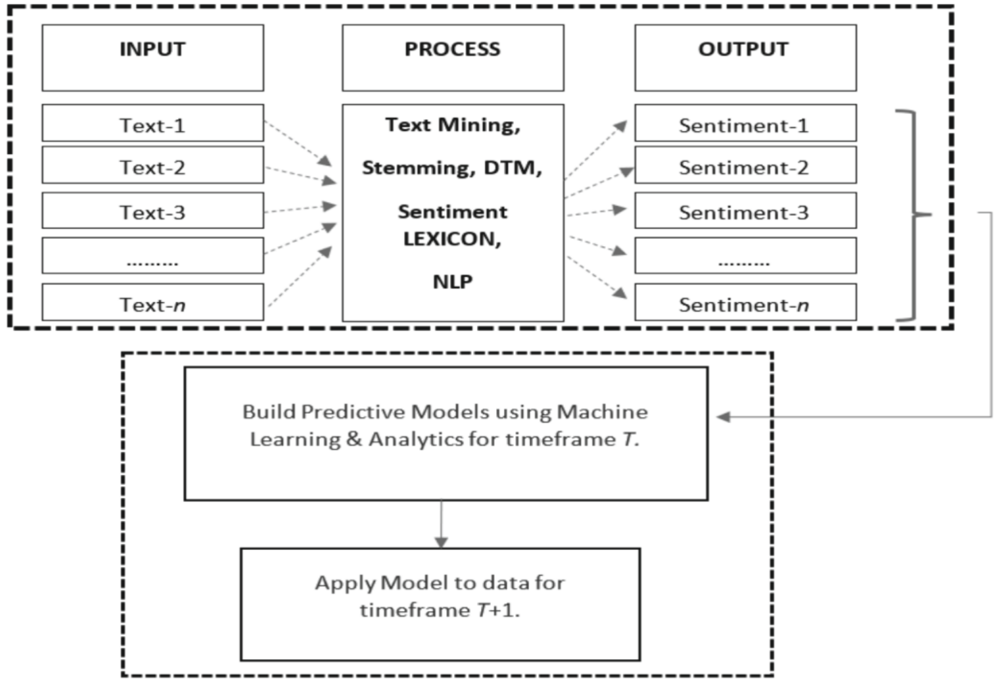
Loughran and McDonald (LM) developed a specific dictionary for the financial domain and business context by examining a large sample of annual reports in the US from 1994 to 2008 (Loughran & McDonald, 2011). The word lists in the LM dictionary are quite extensive. The dictionary categorizes the words into six different sentiments such as constraining, litigious, negative, positive, superfluous, and uncertainty. The list contains 2,355 negative and 354 positive words. We have extensively used the LM word list to measure the sentiment of various financial news articles and business reports (Dougal et al., 2012; Frankel et al., 2022; García, 2013; Huang et al., 2014; Karalevicius et al., 2018; Liu et al., 2017; Liu & McConnell, 2013; Picault & Renault, 2017; Solomon et al., 2014).

In the research framework shown in Figure 2, Text-1, Text-2, and Text-3 each represent the chairperson's statement of a company. Sentiment-1, Sentiment-2, and Sentiment-3 each correspond to the quantitative values of sentiments expressed within the chairperson's speech. The model is built using the sentiments and other financial parameters as independent variables to check the impact on financial performance in the subsequent time.

Sentiment as a Non-Dimensional Number

We have observed differential patterns of reporting in the chairman's statement contingent upon whether the companies are profitable or unprofitable. Unprofitable companies use fewer quantitative results, fewer personal references, and more passive statements and focus more on the future (Clatworthy & Jones, 2006; Rosa & Kawshalya, 2021). The chairman's speech includes both positive and negative sentiments. It is observed from data that the length of chairpersons' speech across different companies varies significantly. Thus, the number of

Figure 2. *Research Framework*



positive and negative words in the report varies accordingly. Instead of separately considering the count of positive and negative words, it is more appropriate to use a non-dimensional number as a measure of the overall tone of the speech. An aggregate tone measure is used in earlier research (Huang et al., 2014, 2018).

Calculation of the Sentiment Score

We used the following steps to calculate the net sentiment score for each of the chairperson's speeches for NIFTY 50 companies from FY17 to FY22 :

- Downloaded the respective companies' annual reports from the BSE website.
- The chairperson's statement was copied into an Excel file respective to the company names and financial year.
- The paragraphs and sentences in the speech were segmented and tokenized to individual words using the *R* code with the help of the "tidytext" library and the "unnest_tokens" function.
- The tokenized words were matched with the positive and negative sentiment words from the LM dictionary, as stated above.
- The positive tone is found by dividing the total count of positive sentiment words by the total word count after removing the stop words.

$$\text{Net Positive Tone} = \frac{\text{Count of positive words}}{\text{Total word count after stop words}}$$

- The negative tone is found by dividing the total count of negative sentiment words by the total word count after removing the stop words.

Table 1. ***Descriptive Statistics of the Sentiment Analysis Data***

Sentiment Types	Year	Mean	*SD*	Median	Min.	Max.	Q1	Q3	IQR	MAD	Skewness	Kurtosis
Positive	2017	0.044	0.015	0.042	0.019	0.109	0.034	0.049	0.015	0.013	1.567	4.72
	2018	0.041	0.014	0.04	0.018	0.077	0.03	0.048	0.017	0.012	0.707	–0.154
	2019	0.042	0.014	0.04	0.015	0.088	0.03	0.047	0.017	0.014	0.759	1.116
	2020	0.042	0.017	0.04	0.015	0.103	0.031	0.048	0.018	0.014	1.37	2.312
	2021	0.039	0.013	0.038	0.014	0.062	0.031	0.05	0.019	0.014	0.054	–1.018
	2022	0.043	0.015	0.041	0.018	0.081	0.035	0.047	0.013	0.01	0.786	0.136
Negative	2017	0.018	0.012	0.014	0	0.056	0.009	0.022	0.012	0.009	1.245	1.216
	2018	0.017	0.011	0.016	0.002	0.049	0.009	0.02	0.011	0.009	1.099	0.83
	2019	0.016	0.009	0.017	0	0.036	0.012	0.021	0.009	0.007	0.04	–0.619
	2020	0.026	0.013	0.024	0	0.066	0.017	0.033	0.016	0.012	0.78	0.844
	2021	0.026	0.011	0.024	0.008	0.056	0.02	0.035	0.014	0.009	0.704	0.255
	2022	0.021	0.01	0.02	0.007	0.048	0.013	0.025	0.012	0.01	0.722	–0.082
Net Sentiment (NPT)	2017	0.026	0.019	0.026	–0.016	0.094	0.017	0.037	0.019	0.015	0.503	1.648
	2018	0.024	0.019	0.025	–0.019	0.069	0.014	0.035	0.02	0.016	–0.013	0.326
	2019	0.025	0.015	0.026	–0.006	0.063	0.016	0.034	0.018	0.013	0.284	0.256
	2020	0.016	0.022	0.014	–0.038	0.092	0.002	0.027	0.025	0.019	0.952	2.123
	2021	0.013	0.017	0.013	–0.027	0.052	0.005	0.025	0.019	0.014	–0.206	–0.237
	2022	0.022	0.017	0.02	–0.014	0.062	0.012	0.031	0.017	0.013	0.249	–0.06

$$\text{Net Positive Tone} = \frac{\text{Count of negative words}}{\text{Total word count after stop words}}$$

➯ To find the overall tone measure or sentiment in the chairman's speech, we subtracted the count of negative words from the count of positive words and divided the result by the total word count in the report.

$$\text{Sentiment Measure} = \frac{\text{Count of positive words} - \text{Count of negative words}}{\text{Total word count after stop words}}$$

Table 1 shows the net positive tone for each of the companies. The chairperson's speeches for DivisLab from FY17 to FY20, as well as HCLTech from FY17, were not available.

Controlled Variables

The study employed the multiple linear regression method to create a predictive model and analyze if the tone of the chairperson's remark might operate as an essential indicator to forecast the short-term and long-term financial performance of organizations. The positive sentiment score, the negative sentiment score, and the interaction effect between the two scores for each company were used as independent variables controlling for the company's financial leverage, size, and age (Sheng et al., 2011).

Prediction Model

We began with the hypothesis that the sentiments expressed in the chairperson's speech might be related to the company's financial performance. To test this, we quantified positive sentiment, negative sentiment, and net sentiment for use in a regression model. As previously mentioned, our hypotheses were: (H01) Sentiments do not exhibit trends in relation to financial performance, and (Ha1) Sentiments do exhibit trends in relation to financial performance. We conducted empirical tests using ROA as an indicator of financial performance. The regression models for different time frames are formulated below:

Model - I :

$$ROA_{i,t} = \beta_0 + \beta_1 \, POSITIVE_{i,t} + \beta_2 \, NEGATIVE_{i,t} + \beta_3 \, POSITIVE_{i,t} * NEGATIVE_{i,t} + \beta_4 \, LEVERAGE_{i,t} + \beta_5 \, SIZE_{i,t} + \epsilon_{i,t}$$
------------- Eq (1)

Model - II :

$$ROA_{i,t+1} = \beta_0 + \beta_1 \, POSITIVE_{i,t} + \beta_2 \, NEGATIVE_{i,t} + \beta_3 \, POSITIVE_{i,t} * NEGATIVE_{i,t} + \beta_4 \, LEVERAGE_{i,t} + \beta_5 \, SIZE_{i,t} + \epsilon_{i,t}$$
------------- Eq (2)

Model I is to check the relationship within the same financial year, and Model-II is for investigating the impact of sentiment on future financial performance.

$ROA_{i,t+1}$ is the return on assets of the *i*-th company for t+1 th financial year. This is defined as the net income before interest and taxes divided by the total assets. $NEGATIVE_{i,t}$ is the fraction of negative sentiment words in the chairperson's speech after removing the stop words from the total word count. $POSITIVE_{i,t}$ is the fraction of positive sentiment words in the respective chairperson's speech after removing the stop words from the total words. The subscript (i,t) represents the *i*-th company for the *t*-th financial year. $LEVERAGE_{i,t}$ is the financial leverage of the *i*-th company for the *t*-th financial year. This is the debt-to-equity ratio data collected from the CMIE database. $SIZE_{i,t}$ is the size of the *i*-th company in the *t*-th financial year defined as the natural logarithm of the total assets of the company.

Analysis and Results

R statistical programming software was used with *R*-studio as an integrated development environment for all the data pre-processing and predictive modeling.

Descriptive Statistics

Descriptive statistics of the financial variables in the study are shown in Table 2.

Table 2. ***Descriptive Statistics of the Financial Variables***

Financial Parameters	Year	Mean	*SD*	Min.	Median	Max.
Debt to Equity Ratio	2016	0.96	1.15	0	0.48	4.99
	2017	0.95	1.13	0	0.44	5.24
	2018	0.9	0.95	0	0.72	4.13
	2019	0.98	1.06	0	0.78	5.1

	2020	0.97	0.94	0	0.79	3.92
	2021	0.85	0.86	0	0.61	3.47
	2022	0.82	0.88	0	0.65	3.69
ROA	2016	9.28	8.73	–7.37	5.97	29.72
	2017	9.07	7.87	–0.19	6.58	27.15
	2018	8.31	7.58	–0.4	4.73	29.61
	2019	8.29	7.79	–2.25	5.66	32.89
	2020	7.84	7.93	–2.82	4.11	34.1
	2021	7.41	6.88	–6.32	5.49	25.64
	2022	8.34	6.66	–3.14	6.5	27.5
Total Assets	2016	13.47	1.53	10.47	13.54	17.24
	2017	13.6	1.52	10.64	13.64	17.36
	2018	13.75	1.51	10.87	13.82	17.4
	2019	13.9	1.49	11.06	14.02	17.48
	2020	14	1.49	11.28	14.16	17.55
	2021	14.11	1.46	11.29	14.25	17.7
	2022	14.21	1.47	11.23	14.41	17.8

Regression Results

The regression coefficients for FY20, FY21, and FY22 are shown in Table 3 for both the same-year performance and future-year performance. For the same-year performance, the empirical results of ROA show that the coefficient estimate of POSITIVE and NEGATIVE sentiments are not significant for the same-year prediction: i.e., the sentiments associated do not show a significant relation with the financial performance of that year. In Model II, for the study of future year performance, the coefficients of the POSITIVE sentiment variable show significant predictive power for ROA. For FY20, the coefficient estimate is 278.3, with a statistical significance of up to a level of 4%. For FY22, the POSITIVE sentiment variable shows predictive power for ROA within a

Table 3. ***Regression Coefficients and Significance***

	MODEL – I (Same Year Performance)			MODEL-II (Future Year Performance)		
Variables	**FY20**	**FY21**	**FY22**	**FY20**	**FY21**	**FY22**
Positive	–20.39	257.86	49.99	278.3	–20.19	356
	(0.8236)	(0.047)*	(0.66)	(0.0404)*	(0.7966)	(0.007)**
Negative	–49.64	255.52	226.9	375.7	–107.98	365
	(0.7776)	(0.183)	(0.308)	(0.1832)	(0.4746)	(0.059)
Positive * Negative	–488.33	–7503.6	–1761.2	–12760	1445.68	–10270
	(0.8968)	(0.083)	(0.6983)	(0.0480)	(0.6545)	(0.019)*
Leverage	–3.39	–3.22	–3.04	–2.6	–2.848	–2.83
	(0.0039)**	(0.0013)**	(0.0017)**	(0.0059)**	(0.0047)**	(0.0041)**
Size (Total Asset)	–2.29	–2.18	–1.616	–2.48	–1.927	–2.042
	(0.00281)**	(0.00035)***	(0.0055)**	(0.0005)***	(0.0033)**	(0.0007)***

significance level of 1%. Broadly, the NEGATIVE sentiment measure was not able to provide good explanatory power for ROA for the next financial year. The POSITIVE * NEGATIVE interaction factor remained statistically insignificant baring the year 2022, where it shows statistical significance up to a level of 2%. We found consistent findings for both models for the LEVERAGE and SIZE indices of financial performance. LEVERAGE and SIZE show a substantial and unfavorable relationship with the company's current and future year financial performance in terms of ROA.

Discussion

The study is conducted to analyze the textual data from the company's management within a framework of signaling theory. This research analyzed the impact of sentiments in the textual data on the financial performance of the companies. This area of research is still in the development stage, as one of the major challenges is establishing the connection between unstructured data in the form of texts and structured data in the form of financial numbers. We study the NIFTY 50 companies for a period extending from 2018 to 2022. We used the word count method to perform text segmentation and tokenization. Then, we used the dictionary method of sentiment extraction (specifically the LM dictionary for financial texts) for studying the sentiments and subsequently used the positive and negative sentiment for the regression analysis. Our hypothesis is about the sentiments having explanatory power for predicting financial performance. The hypothesis is rejected for Model-I, which represents that the sentiments in the chairperson's speech have no impact on the present-year performance of the company. Model II studies the impact of sentiments on the future year's financial performance. The negative sentiments do not have much significance in explaining the ROA numbers for the next financial year. The hypothesis is valid for positive sentiments for FY20 and FY22, i.e., the positive sentiments have predictive power within a significance level of 1% to 5%.

For FY21 or for the period from April 2020 to March 2021, both positive and negative sentiments are not significant in explaining the dependent variable. The average net positive tone (NPT) for FY21 remains the lowest compared to other years. For FY21, the NPT was 0.013, whereas for other years, the NPT was around 0.024, a ~46% fall in aggregate positive sentiment. It may be noted that the majority of institutional lockdowns happened during FY21 as a part of the COVID-19 restrictions. The average ROA for FY21 also fell to 7.4% from an average of 8.3% in other years. It could be because of the reason that as the negative sentiment rises, the correlation with performance loses significance.

Implications

The research is extremely useful for the investor community operating in capital markets. The text analytics methods used in the research, specifically for the financial texts, can be of help for deriving the sentiments across the documents and using them in further analysis. The methods can also be generalized to derive the sentiments from texts involved in non-financial documents. The predictive analysis and results obtained in the research will be immensely helpful in making more informed decisions related to investments. The research results also add to the robustness of the signaling theory. The algorithm of the text analytics process used in the R-code will assist the academicians working on similar methods.

Limitations of the Study and Scope for Further Research

A limitation on this aspect may be noted related to this empirical result. The NIFTY50 companies are generally the stable performers across the industries in India. Given the size and fundamentals of the companies, the performance may not be significantly lower. Therefore, it is recommended to use a larger set of companies to check the impact of sentiments on performance in future studies.

In the present research, we have extracted unigrams from the chairpersons' speech and conducted the text summarization and analytics. Analysis of bigrams or entire sentences can be done in further studies. Any link between sentiments and the share price of a company may be studied to establish the correlations. A similar method can be applied to the texts from investors' conference calls for studying the impact on the quarterly performance of the companies.

Conclusion

The negative sentiment in the chairpersons' speech is not very relevant for the financial performance prediction for the present and future years. However, the positive tone in the speech indicates a better performance of the companies in the future year but not in the current year. The results are also in line with the signaling theory from the perspective that the firms majorly project a positive impression while masking the underperformance situations.

Authors' Contribution

Manoj Kumar Jena conceptualized the study, collected data, developed codes, and wrote the first draft. Srikant Das conceptualized the research and reviewed and modified the draft. Brajaballav Kar conceptualized and discussed the research and reviewed and modified the draft.

Conflict of Interest

The authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Funding Acknowledgment

The authors received no financial support for the research, authorship, and/or for the publication of this article.

References

Abrahamson, E., & Park, C. (1994). Concealment of negative organizational outcomes: An agency theory perspective. *Academy of Management Journal*, *37*(5), 1302–1334. https://doi.org/10.2307/256674

Al-Sayani, Y. M., & Al-Matari, E. M. (2023). Corporate governance characteristics and impression management in financial statements. A further analysis. Malaysian evidence. *Cogent Social Sciences*, *9*(1), Article ID 2191431. https://doi.org/10.1080/23311886.2023.2191431

Bartlett, S. A., & Chandler, R. A. (1997). The corporate report and the private shareholder: Lee and Tweedie twenty years on. *The British Accounting Review*, *29*(3), 245–261. https://doi.org/10.1006/bare.1996.0044

Beattie, V., & Jones, J. M. (1992). The use and abuse of graphs in annual reports: Theoretical framework and empirical study. *Accounting and Business Research*, *22*(88), 291–303. https://doi.org/10.1080/00014788.1992.9729446

Beattie, V., & Jones, J. M. (1999). Australian financial graphs: An empirical study. *Abacus*, *35*(1), 46–76. https://doi.org/10.1111/1467-6281.00034

Bhuyan, A. K., & Baid, D. (2020). Chairman's letter: Does it communicate something? *Indian Journal of Finance*, *14*(1), 29–44. https://doi.org/10.17010/ijf/2020/v14i1/149857

Bozos, K., Nikolopoulos, K., & Ramgandhi, G. (2011). Dividend signaling under economic adversity: Evidence from the London Stock Exchange. *International Review of Financial Analysis*, *20*(5), 364–374. https://doi.org/10.1016/j.irfa.2011.07.003

Che, S., Zhu, W., & Li, X. (2020). Anticipating corporate financial performance from CEO letters utilizing sentiment analysis. *Mathematical Problems in Engineering*, 1–17. https://doi.org/10.1155/2020/5609272

Clatworthy, M. A., & Jones, M. J. (2001). The effect of thematic structure on the variability of annual report readability. *Accounting, Auditing & Accountability Journal*, *14*(3), 311–326. https://doi.org/10.1108/09513570110399890

Clatworthy, M. A., & Jones, M. J. (2006). Differential patterns of textual characteristics and company performance in the chairman's statement. *Accounting, Auditing & Accountability Journal*, *19*(4), 493–511. https://doi.org/10.1108/09513570610679100

Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. *Journal of Management*, *37*(1), 39–67. https://doi.org/10.1177/0149206310388419

Courtis, J. K. (1995). Readability of annual reports: Western versus Asian evidence. *Accounting, Auditing & Accountability Journal*, *8*(2), 4–17. https://doi.org/10.1108/09513579510086795

Courtis, J. K. (1997). Corporate annual report graphical communication in Hong Kong: Effective or misleading? *The Journal of Business Communication (1973)*, *34*(3), 269–284. https://doi.org/10.1177/002194369703400304

Day, J. F. (1986). The use of annual reports by UK investment analysts. *Accounting and Business Research*, *16*(64), 295–307. https://doi.org/10.1080/00014788.1986.9729330

Dougal, C., Engelberg, J., García, D., & Parsons, C. A. (2012). Journalists and the stock market. *The Review of Financial Studies*, *25*(3), 639–679. https://doi.org/10.1093/rfs/hhr133

Falschlunger, L. M., Eisl, C., Losbichler, H., & Greil, A. M. (2015). Impression management in annual reports of the largest European companies: A longitudinal study on graphical representations. *Journal of Applied Accounting Research, 16*(3), 383–399. https://doi.org/10.1108/JAAR-10-2014-0109

Frankel, R., Jennings, J., & Lee, J. (2022). Disclosure sentiment: Machine learning vs. dictionary methods. *Management Science*, *68*(7), 5514–5532. https://doi.org/10.1287/mnsc.2021.4156

Frownfelter-Lohrke, C., & Fulkerson, C. L. (2001). The incidence and quality of graphics in annual reports: An international comparison. *The Journal of Business Communication (1973)*, *38*(3), 337–357. https://doi.org/10.1177/002194360103800308

García, D. (2013). Sentiment during recessions. *The Journal of Finance*, *68*(3), 1267–1300. https://doi.org/10.1111/jofi.12027

Ha, H. T., Van, V. T., & Hung, D. N. (2019). Impact of social responsibility information disclosure on the financial performance of enterprises in Vietnam. *Indian Journal of Finance*, *13*(1), 20–36. https://doi.org/10.17010/ijf/2019/v13i1/141017

Huang, X., Krishnan, S., & Lin, P. (2018). Tone analysis and earnings management. *Journal of Accounting and Finance*, *18*(8), 46–61. https://doi.org/10.33423/jaf.v18i8.110

Huang, X., Teoh, S. H., & Zhang, Y. (2014). Tone management. *The Accounting Review*, *89*(3), 1083–1113. https://doi.org/10.2308/accr-50684

Ittoo, A., Nguyen, L. M., & van den Bosch, A. (2016). Text analytics in industry: Challenges, desiderata and trends. *Computers in Industry*, *78*, 96–107. https://doi.org/10.1016/j.compind.2015.12.001

Jun, S.-P., & Park, D.-H. (2017). Visualization of brand positioning based on consumer web search information: Using social network analysis. *Internet Research*, *27*(2), 381–407. https://doi.org/10.1108/IntR-02-2016-0037

Kar, B., & Das, P. P. (2022). Impact of sector and region on survival and closure. *Indian Journal of Finance*, *16*(2), 8–23. https://doi.org/10.17010/ijf/2022/v16i2/160397

Kar, B., & Jena, M. K. (2019). Performance and age of companies listed on the Bombay Stock Exchange. *Indian Journal of Finance*, *13*(5), 52–67. https://doi.org/10.17010/ijf/2019/v13i5/144185

Karalevicius, V., Degrande, N., & De Weerdt, J. (2018). Using sentiment analysis to predict interday Bitcoin price movements. *Journal of Risk Finance*, *19*(1), 56–75. https://doi.org/10.1108/JRF-06-2017-0092

Kim, H., Lee, T., Ryu, S., & Kim, N. (2018). A study on text mining methods to analyze civil complaints: Structured association analysis. *Journal of the Korea Industrial Information Systems*, *23*(3), 13–24. https://doi.org/10.9723/jksiis.2018.23.3.013

Klein, L. S., & Peterson, D. R. (1989). Earnings forecast revisions associated with stock split announcements. *Journal of Financial Research*, *12*(4), 319–328. https://doi.org/10.1111/j.1475-6803.1989.tb00525.x

Lacity, M. C., & Janson, M. A. (1994). Understanding qualitative data: A framework of text analysis methods. *Journal of Management Information Systems*, *11*(2), 137–155. https://doi.org/10.1080/07421222.1994.11518043

Li, F. (2010). The information content of forward-looking statements in corporate filings—A Naïve Bayesian machine learning approach. *Journal of Accounting Research*, *48*(5), 1049–1102. https://doi.org/10.1111/j.1475-679X.2010.00382.x

Liu, B., & McConnell, J. J. (2013). The role of the media in corporate governance: Do the media influence managers' capital allocation decisions? *Journal of Financial Economics*, *110*(1), 1–17. https://doi.org/10.1016/j.jfineco.2013.06.003

Liu, B., McConnell, J. J., & Xu, W. (2017). The power of the pen reconsidered: The media, CEO human capital, and corporate governance. *Journal of Banking & Finance*, *76*, 175–188. https://doi.org/10.1016/j.jbankfin.2016.12.004

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, *66*(1), 35–65. https://doi.org/10.1111/j.1540-6261.2010.01625.x

Luo, Y., & Zhou, L. (2020). Textual tone in corporate financial disclosures: A survey of the literature. *International Journal of Disclosure and Governance*, *17*(2–3), 101–110. https://doi.org/10.1057/s41310-020-00077-y

McLellan, E., MacQueen, K. M., & Neidig, J. L. (2003). Beyond the qualitative interview: Data preparation and transcription. *Field Methods*, *15*(1), 63–84. https://doi.org/10.1177/1525822X02239573

Merkl-Davies, D. M., & Koller, V. (2012). 'Metaphoring' people out of this world: A critical discourse analysis of a chairman's statement of a UK defence firm. *Accounting Forum*, *36*(3), 178–193. https://doi.org/10.1016/j.accfor.2012.02.005

Ng, E. J., & Koh, H. C. (1994). An agency theory and probit analytic approach to corporate non-mandatory disclosure compliance. *Asia-Pacific Journal of Accounting*, *1*(1), 29–44. https://doi.org/10.1080/10293574.1994.10510464

Picault, M., & Renault, T. (2017). Words are not all created equal: A new measure of ECB communication. *Journal of International Money and Finance*, *79*, 136–156. https://doi.org/10.1016/j.jimonfin.2017.09.005

Pineda-Jaramillo, J., & Pineda-Jaramillo, D. (2022). Analysing travel satisfaction of tourists towards a metro system from unstructured data. *Research in Transportation Business & Management*, *43*, Article ID 100746. https://doi.org/10.1016/j.rtbm.2021.100746

Qiu, X. Y., Srinivasan, P., & Hu, Y. (2014). Supervised learning models to predict firm performance with annual reports: An empirical study. *Journal of the Association for Information Science and Technology*, *65*(2), 400–413. https://doi.org/10.1002/asi.22983

Ranjan, S., Singh, I., Dua, S., & Sood, S. (2018). Sentiment analysis of stock blog network communities for prediction of stock price trends. *Indian Journal of Finance*, *12*(12), 7–21. https://doi.org/10.17010/ijf/2018/v12i12/139888

Rosa, S. J., & Kawshalya, M. D. (2021). *Differential patterns of textual characteristics in the chairman's statement and company performance: Evidence from Sri Lanka*. Available at SSRN. https://doi.org/10.2139/ssrn.4117510

Sheng, S., Zhou, K. Z., & Li, J. J. (2011). The effects of business and political ties on firm performance: Evidence from China. *Journal of Marketing*, *75*(1), 1–15. https://doi.org/10.1509/jm.75.1.1

Singh, A. K., & Kalra, S. (2024). IPO forecasting using machine learning methodologies: A systematic review apropos financial markets in the digital era. *Arthshastra Indian Journal of Economics & Research*, *13*(2), 23–40. https://doi.org/10.17010/aijer/2024/v13i2/173502

Smith, M., & Taffler, R. (1992). The chairman's statement and corporate financial performance. *Accounting & Finance*, *32*(2), 75–90. https://doi.org/10.1111/j.1467-629X.1992.tb00187.x

Solomon, D. H., Soltes, E., & Sosyura, D. (2014). Winners in the spotlight: Media coverage of fund holdings as a driver of flows. *Journal of Financial Economics*, *113*(1), 53–72. https://doi.org/10.1016/j.jfineco.2014.02.009

Spence, M. (2002). Signaling in retrospect and the informational structure of markets. *American Economic Review*, *92*(3), 434–459. https://doi.org/10.1257/00028280260136200

Thomas, D. A. (2014). Searching for significance in unstructured data: Text mining with Leximancer. *European Educational Research Journal*, *13*(2), 235–256. https://doi.org/10.2304/eerj.2014.13.2.235

Trombeta, G., & Cox, S. M. (2022). The textual-visual thematic analysis: A framework to analyze the conjunction and interaction of visual and textual data. *The Qualitative Report*, *27*(6), 1557–1574. https://doi.org/10.46743/2160-3715/2022.5456

Yasar, B., Martin, T., & Kiessling, T. (2020). An empirical test of signalling theory. *Management Research Review*, *43*(11), 1309–1335. https://doi.org/10.1108/MRR-08-2019-0338

Zhang, Y., & Wiersema, M. F. (2009). Stock market reaction to CEO certification: The signaling role of CEO background. *Strategic Management Journal*, *30*(7), 693–710. https://doi.org/10.1002/smj.772

About the Authors

Prof. Manoj Kumar Jena has over 10 years of corporate experience spanning diverse sectors such as manufacturing, IT, automotive, and finance. His areas of interest include data analytics & machine learning, predictive modeling, text mining, and analytics.

Dr. Srikant Das is a senior academician with significant experience in teaching, research, and corporates. He is presently working as the Chief Technology Officer of KIIT University. His area of expertise includes IT and project management.

Prof. Brajaballav Kar has over 30 years of experience in corporate and academics. He has a working knowledge of different ERP systems, SCM, and CRM software, as well as expertise in operations management. His papers have been published in many different national and international journals.