

Article

NLP Sentiment Analysis and Accounting Transparency: A New Era of Financial Record Keeping

Alessio Faccia ^{1,*}, Julie McDonald ¹ and Babu George ²

¹ School of Business, University of Birmingham Dubai, Dubai 341799, United Arab Emirates; j.a.mcdonald@bham.ac.uk

² School of Business, Alcorn State University, Lorman, MS 39096, USA; bgeorge@alcorn.edu

* Correspondence: a.faccia@bham.ac.uk

Abstract: Transparency in financial reporting is crucial for maintaining trust in financial markets, yet fraudulent financial statements remain challenging to detect and prevent. This study introduces a novel approach to detecting financial statement fraud by applying sentiment analysis to analyse the textual data within financial reports. This research aims to identify patterns and anomalies that might indicate fraudulent activities by examining the language and sentiment expressed across multiple fiscal years. The study focuses on three companies known for financial statement fraud: Wirecard, Tesco, and Under Armour. Utilising Natural Language Processing (NLP) techniques, the research analyses polarity (positive or negative sentiment) and subjectivity (degree of personal opinion) within the financial statements, revealing intriguing patterns. Wirecard showed a consistent tone with a slight decrease in 2018, Tesco exhibited marked changes in the fraud year, and Under Armour presented subtler shifts during the fraud years. While the findings present promising trends, the study emphasises that sentiment analysis alone cannot definitively detect financial statement fraud. It provides insights into the tone and mood of the text but cannot reveal intentional deception or financial discrepancies. The results serve as supplementary information, enriching traditional financial analysis methods. This research contributes to the field by exploring the potential of sentiment analysis in financial fraud detection, offering a unique perspective that complements quantitative methods. It opens new avenues for investigation and underscores the need for an integrated, multidimensional approach to fraud detection.



Citation: Faccia, A.; McDonald, J.; George, B. NLP Sentiment Analysis and Accounting Transparency: A New Era of Financial Record Keeping.

Computers **2024**, *13*, 5. <https://doi.org/10.3390/computers13010005>

Academic Editor: Paolo Bellavista

Received: 20 August 2023

Revised: 18 December 2023

Accepted: 20 December 2023

Published: 23 December 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: financial statement fraud; sentiment analysis; natural language processing (NLP); fraud detection; forensic accounting; Wirecard; Tesco; Under Armour

1. Introduction

Transparency in financial reporting is a fundamental principle that underpins confidence and trust in financial markets [1]. It allows investors, creditors, regulators, and other stakeholders to make informed decisions [2] by providing clear and accurate information about a company's financial health. Transparency ensures financial information is readily available, understandable, and comparable across different organisations [3]. In an era where capital markets are highly interconnected, transparency plays a vital role in maintaining market stability and fostering economic growth [4]. It helps reduce information asymmetry, enhance corporate governance, and promote accountability and ethical conduct [5]. Despite the rigorous regulations and standards to ensure transparency, fraudulent financial statements still need to be solved [6]. This type of fraud involves the deliberate misrepresentations, omissions, or manipulations of financial information to deceive stakeholders.

Fraudulent financial statements can take various forms, including revenue recognition fraud, expense manipulation, and asset valuation manipulation [7,8]. The motivations behind such fraud may range from meeting performance targets, obtaining financing,

inflating stock prices, or concealing financial difficulties [9]. Detecting financial statement fraud is often challenging [10], as it may involve sophisticated schemes carefully designed to evade traditional auditing and monitoring techniques [11–13]. The consequences of financial statement fraud are far-reaching [14]. For investors, it can lead to significant financial losses, eroded trust in financial markets, and reduced investment activity. For the companies involved, it may result in legal penalties, reputational damage, loss of market value, and even bankruptcy in severe cases.

The ripple effect of financial statement fraud can also impact the broader economy [15,16] by undermining confidence in financial institutions and regulatory bodies, thereby affecting capital allocation and economic stability.

This study explores the application of sentiment analysis as a novel approach to detecting potential financial statement fraud. Sentiment analysis involves the computational analysis of textual data to identify and understand subjective patterns and emotions. By analysing the language and sentiment expressed in financial statements, this research aims to uncover hidden patterns and anomalies that might indicate fraudulent activities. Unlike traditional numerical analysis, sentiment analysis provides a unique lens to scrutinise the psychological aspects of financial reporting, potentially revealing inconsistencies, exaggerations, or subtle manipulations that may signal fraud. The central hypothesis of this study is that variations in language and sentiment within financial statements can be significant indicators of fraudulent activities. By applying advanced sentiment analysis techniques to a carefully selected dataset of proven fraudulent financial statements, this research seeks to verify to what extent sentiment analysis could serve as an effective tool in signalling financial statement fraud.

Natural Language Processing (NLP) can identify patterns or cues suggesting concealed or obscured information within a text [17,18]. However, the capability to detect concealed information is not straightforward [19] and depends on various factors, including the nature of the concealment, the quality of the data, and the specific techniques used. Among the NLP-related techniques potentially available, there are:

- Detecting Euphemisms and Metaphors [20]: Concealment can sometimes occur through euphemisms or metaphors. NLP techniques that recognise semantic relations can identify these linguistic phenomena, although deciphering their exact meaning may still be complex.
- Analysing Sentiment [21]: By examining the sentiment of a text, NLP might identify inconsistencies or subtle cues that something might be hidden. For instance, if the sentiment within a document suddenly changes without apparent reason, it may indicate an attempt to conceal information.
- Steganography Detection [22]: Steganography is hiding information within other information. Although commonly associated with images, it can also be applied to text. Specialised algorithms can sometimes detect patterns in text that might indicate steganography.
- Identifying Non-Natural Language Patterns [23]: If information is concealed through coded language or special patterns, advanced NLP techniques may detect the non-natural use of language. By modelling what is considered 'typical' language usage, deviations from this norm can be flagged for further investigation.
- Using Contextual Analysis [24]: Sometimes, what is not said is as important as what is said. Analysing the context in which information is presented and cross-referencing it with known facts can reveal inconsistencies that might indicate concealment.
- Challenges and Limitations [25]: Recognising that detecting concealed information is a highly complex task is crucial. It may require domain-specific knowledge and careful tuning of algorithms. Furthermore, false positives are possible, where the algorithms mistakenly identify concealed information where there is none.
- Legal and Ethical Considerations [26]: These techniques must also be aligned with legal and ethical standards, especially concerning privacy and consent.

In this research context, we will focus on sentiment analysis only to explore potential financial statement frauds.

The Fraud Tree [27] is a conceptual framework that categorises fraud into three main branches, each representing a different type of fraudulent activity. These branches are financial statement fraud, corruption, and asset misappropriation.

Financial statement fraud involves intentionally manipulating, altering, or falsifying financial statements to create a false appearance of a company's financial health. Examples include revenue recognition fraud, earnings management, and hiding liabilities. It can significantly impact investors, creditors, and other stakeholders who rely on accurate financial information.

Corruption refers to unethical practices involving bribery, kickbacks, conflicts of interest, or other forms of influence-peddling. It often involves abusing power (or position) for personal gain (or to benefit others) at the organisation's expense. Corruption can take place at various levels within an organisation or government body.

Asset misappropriation involves the theft or misuse of an organisation's assets. It can include skimming revenues, stealing inventory, payroll fraud, or fraudulent disbursements. Asset misappropriation is the most common type of fraud but usually involves smaller amounts than financial statement fraud.

Our analysis focused on financial statement fraud because it is the most relevant to the task. Financial statement fraud directly pertains to manipulating these documents and often has far-reaching consequences, affecting the organisation and the broader financial markets, investors, and regulators. It may lead to a loss of investor confidence, legal sanctions, and significant financial loss. Financial statement fraud often involves sophisticated schemes and intricate manipulations of financial data. Analysing the language and sentiment within financial statements could provide supplementary insights into the tone and presentation of the financial information.

While corruption and asset misappropriation are significant areas of fraud, they were not the focus of our analysis due to the specific context of the request and the nature of the documents analysed. Financial statement fraud provided the most pertinent avenue for investigation concerning the sentiment analysis of financial reports and the known fraudulent activities of the companies in question.

In summary, this study pioneers the application of Natural Language Processing (NLP) in financial statement analysis, specifically focusing on sentiment analysis to detect financial statement fraud. Our research distinctively bridges the fields of linguistics, computer science, and finance, applying advanced NLP techniques to scrutinise the language and sentiment expressed in financial reports. The study meticulously analyses the financial statements of three major companies—Wirecard, Tesco, and Under Armour—known for financial statement fraud. Examining shifts in polarity and subjectivity within these reports reveals significant patterns that may indicate attempts to manipulate financial portrayals. While acknowledging the limitations of sentiment analysis as a standalone tool, our findings significantly contribute to the field by offering a novel, multidimensional perspective on fraud detection. This research enriches the existing fraud detection toolkit and opens new avenues for future investigations, underscoring the imperative for an integrated approach that combines traditional financial analysis with linguistic insights to enhance the efficacy of fraud detection mechanisms.

2. Related Work in the Literature

The area of sentiment analysis and financial statement fraud detection is a rich and evolving study area, reflecting the growing complexity of financial markets and the advanced technologies employed to navigate them. Sentiment analysis, a field at the intersection of linguistics, computer science, and artificial intelligence, has become increasingly relevant in financial contexts. Its evolution from a rudimentary analysis of word polarity to sophisticated Natural Language Processing (NLP) and machine learning algorithms underscores its potential in dissecting complex financial narratives.

Applying sentiment analysis in financial reporting and market prediction is a noteworthy development. Studies have demonstrated how the sentiment in financial news, earnings calls, and reports can offer predictive insights into market movements and company performance. This approach transcends traditional numerical analysis, offering a nuanced view of market sentiments that often drive investor behaviour.

Simultaneously, financial statement fraud—a critical concern in the corporate world—has seen a shift in detection methodologies. Traditionally, this detection relied heavily on ratio analysis, trend analysis, and forensic accounting. However, while useful, these methods have limitations in dealing with the subtleties and complexities of modern financial fraud. The advent of technology in fraud detection, particularly the integration of data analytics and advanced computational methods, has opened new avenues for more effective and nuanced fraud detection mechanisms.

The convergence of sentiment analysis and financial statement fraud detection represents a burgeoning area of research. Some studies have begun to explore how sentiment analysis can augment traditional fraud detection methods. For instance, analysing the tone and sentiment of financial reports and executive communications has provided new indicators of potential fraudulence. However, these methodologies are still nascent and often face challenges such as the need for large datasets and sophisticated algorithms capable of interpreting complex financial jargon and nuanced expressions.

Despite these advancements, there remains a gap in the literature specifically addressing the combined use of sentiment analysis in financial statement fraud detection. Many studies focus on either aspect in isolation, leaving a significant opportunity for research that integrates these two fields. This gap suggests potential for future investigations that could develop comprehensive models encompassing sentiment analysis and fraud detection techniques. Such research could offer more robust tools for financial analysts and regulators, enhancing their ability to detect and prevent fraudulent activities.

In sentiment analysis, recent research has made significant strides in enhancing the accuracy and depth of emotion and aspect detection in textual data. Huang et al. [28] made a notable contribution by introducing aspect-specific context position information in their study. This innovative approach underscores the importance of contextual details in determining the sentiment related to specific aspects within a text. Meanwhile, Fei et al. [29] explored the concept of latent emotion memory. Their research focused on the complexities of classifying multiple emotions within a single text, highlighting the intricate interplay between different emotional states. Furthering this exploration, Fei et al. [30] introduced a multiplex cascade framework. This framework integrates various layers of sentiment analysis, offering a more comprehensive approach that benefits from the accumulated knowledge in the field. Together, these studies represent significant advancements in sentiment analysis, demonstrating how nuanced contextual understanding and innovative methodologies can profoundly enhance our ability to interpret and classify emotions and aspects in text.

Sentiment analysis in the financial domain has garnered significant attention, with researchers employing advanced techniques to extract and interpret sentiment from financial texts. Sohangir et al. [31] demonstrate how deep learning models can effectively apply to financial data, offering insights into market trends and investor sentiments. This study underscores the potential of big data and deep learning in transforming financial analysis and decision-making processes. Building on this foundation, Xiang et al. [32] introduce an innovative neural model that integrates semantic and syntactic information. This approach enhances the accuracy of sentiment analysis in financial texts, reflecting the complexity and specificity of financial language.

Further advancing the field, Shang et al. [33] present a novel “Lexicon Enhanced Collaborative Network for targeted Financial Sentiment Analysis”. Their research emphasises the importance of a lexicon-based approach in conjunction with neural networks to better capture the nuanced sentiments in financial contexts, particularly when targeting specific aspects or entities in financial reports or news. Collectively, these studies highlight

the evolving landscape of sentiment analysis in finance, where the integration of deep learning, semantic and syntactic analysis, and lexicon-based approaches are paving the way for more accurate, nuanced, and context-specific interpretations of financial sentiments. This progression in the field is crucial for investors, analysts, and policymakers, offering enhanced tools for understanding market dynamics and investor behaviour.

2.1. A Unique Application of Existing Techniques

While sentiment analysis and NLP are well-established, their application in detecting financial statement fraud, particularly in the manner used in this article, offers new insights. This approach differs from traditional methods in forensic accounting and fraud detection. In particular, this study intersects various fields—finance, linguistics, and computer science. This interdisciplinary nature is a novel contribution, combining insights from different domains to address financial statement fraud uniquely. The choice of case studies (Wirecard, Tesco, and Under Armour) and the specific period analysed also contribute to the novelty. This study also opens up new avenues for future research. If our methodology and findings suggest unexplored areas or potential advancements in fraud detection, these should be highlighted as contributions to the field.

2.2. Rationale behind Case Study Selection

The case studies—Wirecard, Tesco, and Under Armour—were meticulously chosen to represent various scenarios in financial statement fraud. Each case embodies unique characteristics of fraud in different sectors and geographical contexts, providing a comprehensive understanding of the phenomenon. These cases are historically significant and offer contemporary lessons in financial fraud. Their analysis yields insights relevant to current financial markets and regulatory environments. Each case study presents different methods of fraud and varying impacts on stakeholders, offering a broad spectrum for analysis. This diversity ensures a thorough examination of the applicability of sentiment analysis in varying fraudulent contexts.

The analysis of each case was not limited to a superficial review. Instead, it involved a deep dive into financial statements, market reactions, and the linguistic nuances of corporate communications. This comprehensive approach ensures a robust analysis of each case. The study combines qualitative insights from corporate communications with quantitative financial data, offering a more nuanced understanding of each case's context and the dynamics of the fraud involved.

Using sophisticated NLP techniques to analyse financial reports and communications sentiment adds depth to the experimental approach. This is not a mere application of basic sentiment analysis—it involves complex linguistic and statistical methodologies. The findings from the case studies are cross-validated with established theories and models in financial fraud detection, ensuring that the results are grounded in recognised academic frameworks.

The study offers insights that traditional financial fraud detection methods might overlook. This includes subtle linguistic cues in corporate communication that could indicate potential fraud. Finally, the results have significant implications for investors, regulators, and financial analysts, providing them with a new toolset to detect and understand financial statement fraud.

3. Materials and Methods

There are several compelling justifications for adopting sentiment analysis as the central approach: (a) Identifying inconsistencies and anomalies: sentiment Analysis involves examining the emotional tone and polarity expressed in text. Sudden shifts in sentiment within financial statements could indicate attempts to conceal information or mislead stakeholders. A drastic change in sentiment without an apparent reason might indicate underlying fraudulent activities. (b) Uncovering subtle clues: fraudulent activities often involve complex strategies to obscure their true nature. Sentiment analysis has the potential

to uncover subtle cues and inconsistencies in language usage that might be indicative of concealed information. It can help unveil hidden motives or intentions that might not be immediately evident. (c) Complementing traditional approaches: traditional methods of detecting financial fraud might focus on numerical and quantitative analysis. Incorporating sentiment analysis as a complementary technique allows for a more holistic assessment of financial documents. By considering quantitative data and qualitative language cues, your research can offer a more comprehensive fraud detection framework. (d) Scalability and accessibility: sentiment analysis techniques have advanced significantly in recent years, making them more accessible and applicable to large datasets. This scalability is crucial when dealing with extensive financial documents, as it allows for efficient processing and analysis of a substantial amount of textual information. (e) Interdisciplinary approach: by integrating NLP techniques like sentiment analysis into financial fraud detection, your research will showcase an interdisciplinary approach that leverages linguistic and financial expertise. This approach has the potential to yield unique insights that might be missed by relying solely on traditional financial analysis methods.

Potential for automation: Automating fraud detection processes is a growing necessity in today's fast-paced business landscape. Sentiment analysis can be integrated into automated systems that continuously monitor financial documents, alerting stakeholders to suspicious shifts in sentiment patterns that warrant further investigation (See Figures 1 and 2).

Step 1: Extracting Text from the PDF Document

Step 2: Cleaning the Extracted Text

Step 3: Sentiment Analysis

Performed a basic sentiment analysis using a pre-trained model (TextBlob), and then discuss what further steps could be taken to detect potential fraud.

Figure 1. Sentiment analysis steps.

```
python

# Importing TextBlob for sentiment analysis
from textblob import TextBlob

# Function to perform sentiment analysis using TextBlob
def analyze_sentiment_textblob(text):
    blob = TextBlob(text)
    sentiment_scores = {
        'polarity': blob.sentiment.polarity,
        'subjectivity': blob.sentiment.subjectivity
    }
    return sentiment_scores

# Performing sentiment analysis on the cleaned text using TextBlob
sentiment_scores_textblob = analyze_sentiment_textblob(cleaned_text)
sentiment_scores_textblob
```

Figure 2. Python coding for the sentiment analysis performance.

In our study, the choice of TextBlob as the primary tool for sentiment analysis warrants a thorough justification, particularly given its known limitations and the specific demands of analysing financial documents. TextBlob, being a lexicon-based and rule-based system, offers a straightforward and accessible approach to sentiment analysis. However, we acknowledge that its general approach may not fully encapsulate the nuanced and technical language typical of financial reporting.

Several factors influenced the decision to employ TextBlob. Firstly, its simplicity and ease of implementation allowed for a preliminary exploration of sentiment in financial texts. TextBlob's rule-based system, while less sophisticated than machine learning models, provides a clear and transparent mechanism for sentiment analysis, which is beneficial for initial explorations in a new application domain like financial fraud detection.

Furthermore, we recognise that TextBlob, a general-purpose tool, might not be ideally suited for the specific lexicon and complex semantics of financial reporting. The financial domain encompasses a range of specific terminologies and expressions that general sentiment analysis tools may not accurately interpret. This limitation is a significant aspect of our study and serves as a basis for future research to explore more specialised sentiment analysis methods tailored to financial contexts.

As part of our methodological rigour, we considered various sentiment analysis tools before finalising TextBlob. However, due to the scope and resources available for this study, an exhaustive comparison with other tools was not feasible. Tools such as VADER and Stanford's NLP toolkit were considered; each offers its strengths in sentiment analysis. VADER, for instance, is known for its effectiveness in handling social media text, while Stanford's toolkit offers advanced machine learning-based approaches.

Our decision not to engage in a detailed comparison of these tools in the current study was based on a strategic choice to establish a foundational understanding of sentiment analysis in financial fraud detection. It was deemed more beneficial to focus on exploring the applicability of sentiment analysis in this domain rather than conducting a comparative analysis of various tools.

We acknowledge the importance of such a comparison for future research. A detailed examination of different sentiment analysis tools, particularly those tailored to financial texts, could provide valuable insights into the most effective methodologies for detecting financial statement fraud. This future work would significantly contribute to developing a more nuanced and accurate sentiment analysis framework for this domain.

While TextBlob was a starting point in applying sentiment analysis to financial statement fraud detection, we acknowledge its limitations and the need for more specialised tools in future research. This study's findings should be viewed as preliminary, providing a basis for more in-depth investigations that employ advanced sentiment analysis techniques better suited to the complex and specific nature of financial reporting.

While the sentiment analysis provides insight into the overall tone of the document, it does not directly indicate financial statement fraud. Financial statements are typically factual and objective, so the sentiment may not reveal fraudulent activities.

To better assess the potential for financial statement fraud, a more detailed analysis would be required, including:

- Analysing financial ratios: investigating key financial ratios and metrics to identify inconsistencies or anomalies.
- Comparing with industry benchmarks: comparing the company's financial performance with industry standards and peers.
- Reviewing auditor's statements: analysing the auditor's comments and any qualifications in the report.
- Using specialised tools: employing specialised forensic accounting tools and techniques to detect potential fraud.

Sentiment analysis can be a part of a broader investigation into financial statement integrity, but it should be combined with other analytical methods and expert review to provide a comprehensive assessment.

Three case studies are considered for the application of the sentiment analysis described in this section: Wirecard (Germany), Tesco (UK), and Under Armour (US). All the cases are confirmed cases of financial statement fraud.

The Wirecard case already led to convictions [34] and charges over “deliberate money laundering in combination with aggravated fraud and deliberate violation of accounting duties” [35]. The fraud can be traced back to 2011 [36]. Therefore, we consider the most recent full financial statements available (2014–2018) for the analyses already affected by financial statement fraud.

Concerning the Tesco case, a GBP 250 million accounting black hole was unveiled in 2014 [37], also acknowledged by the company [38], which also appropriately restated the comparative fiscal year 2014 column in the subsequent annual report 2015 under IFRS requirements. For our analyses, we considered the periods between 2011 and 2014, which include the year affected by financial statement fraud (2014) and three previous years (2011–2013) as a baseline.

For the Under Armour case, we analysed the period between 2012 and 2016, where 2015–2016 was affected by financial statement fraud (Under Armour will pay USD 9 million to settle federal regulators’ charges that it misled investors about its sales growth in 2015 and 2016) [39] and the previous three (2012–2014) as a baseline.

The three cases were selected to ensure geographical variety (Wirecard in Germany, Tesco in the UK, and Under Armour in the US) because they could ensure specific confirmed and (relatively) recent financial statement frauds that affected identified accounting periods.

4. Results

The analyses were performed on the three case studies and returned the following results.

4.1. Wirecard

Fiscal Year 2014

- Polarity: A polarity score of 0.0674 indicates a slightly positive sentiment in the text. It may reflect optimistic language to present positive company performance and growth aspects.
- Subjectivity: The subjectivity score of 0.3525 represents moderate subjectivity in the text. This score suggests a combination of objective financial data and subjective interpretations or opinions.

Fiscal Year 2015

- Polarity: A polarity score of 0.0697 indicates a slightly positive sentiment in the text. It aligns with typical financial reports, where positive language may emphasise growth or favourable performance.
- Subjectivity: The subjectivity score of 0.3656 reflects moderate subjectivity in the text. It may indicate a combination of objective financial data and subjective interpretations or opinions.

Fiscal Year 2016

- Polarity: With a score of 0.0745, the document exhibits a slightly positive sentiment. It aligns with using positive language in financial reports to emphasise growth, success, or favourable conditions.
- Subjectivity: The subjectivity score of 0.3661 indicates moderate subjectivity in the text, reflecting a mix of objective financial data and subjective interpretations or opinions.

Fiscal Year 2017

- Polarity: This score indicates a slightly positive sentiment in the text. A score closer to 1 would signify a strong positive sentiment, while a score closer to -1 would indicate a strong negative sentiment. The given score suggests a mildly positive tone in the document.

- Subjectivity: This score reflects a moderate level of subjectivity in the text. A score closer to 1 would indicate high subjectivity (personal opinions or feelings), while a score closer to 0 would indicate objectivity (factual information).

Fiscal Year 2018

- Polarity: This score ranges from -1 to 1 , where -1 indicates a negative sentiment, 1 indicates a positive sentiment, and 0 indicates a neutral sentiment. The score of 0.0716 suggests a slightly positive sentiment in the text.
- Subjectivity: This score ranges from 0 to 1 , where 0 is objective and 1 is subjective. A score of 0.3807 indicates a moderate level of subjectivity in the text.

4.2. Tesco

Fiscal Year 2011

- Polarity: A polarity score of 0.1071 indicates a slightly positive sentiment in the text. It is consistent with typical financial reports, where positive language may be used to present an optimistic view of the company's performance.
- Subjectivity: The subjectivity score of 0.3715 represents moderate subjectivity in the text. It may suggest a mix of objective financial data and subjective interpretations or opinions.

Fiscal Year 2012

- Polarity: A polarity score of 0.1018 indicates a slightly positive sentiment in the text. It may reflect the positive language used in annual financial reports to convey success, growth, or achievements.
- Subjectivity: The subjectivity score of 0.3649 represents moderate subjectivity in the text. This score suggests a blend of objective financial data and subjective interpretations or opinions.

Fiscal Year 2013

- Polarity: A polarity score of 0.1186 indicates a positive sentiment in the text. It is consistent with typical financial reports, where positive language may convey an optimistic view of the company's performance and achievements.
- Subjectivity: The subjectivity score of 0.5984 is higher than previous reports, reflecting a substantial level of subjectivity in the text. It may indicate the presence of more opinions, interpretations, or subjective statements in addition to the objective financial data.

Fiscal Year 2014

- Polarity: A polarity score of 0.1479 indicates a positive sentiment in the text. It suggests that the document contains language emphasising positive aspects, achievements, or favourable conditions. It is a common practice in annual reports to convey an optimistic view of the company's performance.
- Subjectivity: The subjectivity score of 0.4607 reflects a higher subjectivity level than previous reports. It may indicate the presence of more opinions, interpretations, or subjective statements in the text in addition to the objective financial data.

4.3. Under Armour

Fiscal Year 2012

- Polarity: A polarity score of 0.0506 indicates a slightly positive sentiment in the text. It likely reflects the language used to emphasise Under Armour's commitment to innovation, athlete insights, and product development.
- Subjectivity: The subjectivity score of 0.3592 represents moderate subjectivity in the text. It suggests a blend of objective data and subjective interpretations or opinions.

Fiscal Year 2013

- Polarity: A polarity score of 0.0507 indicates a slightly positive sentiment in the text. It reflects the language used to present the brand’s global expansion, product launches, and plans for growth in a positive light.
- Subjectivity: The subjectivity score of 0.3622 represents moderate subjectivity in the text. It may suggest a combination of objective facts and subjective interpretations or opinions.

Fiscal Year 2014

- Polarity: A polarity score of 0.0353 indicates a slightly positive sentiment in the text. It may reflect the language used to highlight achievements, endorsements, and growth in the brand’s connected fitness systems.
- Subjectivity: The subjectivity score of 0.3612 represents moderate subjectivity in the text. It suggests a mix of factual financial information and subjective expressions or opinions.

Fiscal Year 2015

- Polarity: A polarity score of 0.0407 indicates a slightly positive sentiment in the text. It may reflect the language used to convey success, awards, and positive achievements related to the brand’s performance in the footwear industry.
- Subjectivity: The subjectivity score of 0.3655 represents moderate subjectivity in the text. It may suggest a combination of factual data and subjective expressions or interpretations.

Fiscal Year 2016

- Polarity: A polarity score of 0.0446 indicates a slightly positive sentiment in the text. It suggests the document contains language emphasising positive aspects, achievements, or favourable conditions. It is common for annual reports to convey a positive view of the company’s performance.
- Subjectivity: The subjectivity score of 0.3534 reflects moderate subjectivity in the text. It may indicate a blend of objective data and subjective statements or interpretations.

All the results are summarised in Table 1 below.

Table 1. NLP—Sentiment analysis (fiscal years in red affected by financial statement fraud).

Company	Fiscal Year	Polarity	Subjectivity	Links
Wirecard	2018	0.0716	0.3807	https://www.wirecard.com/wp-content/uploads/2020/12/Annual-Report-2018.pdf
Wirecard	2017	0.0772	0.3668	https://www.wirecard.com/wp-content/uploads/2020/12/Annual-Report-2017.pdf
Wirecard	2016	0.0745	0.3661	https://www.wirecard.com/wp-content/uploads/2020/12/Annual-Report-2016.pdf
Wirecard	2015	0.0697	0.3656	https://www.wirecard.com/wp-content/uploads/2020/12/Annual-Report-2015.pdf
Wirecard	2014	0.0674	0.3525	https://www.wirecard.com/wp-content/uploads/2020/12/Annual-Report-2014.pdf
TESCO	2014	0.1479	0.4607	https://www.annualreports.com/HostedData/AnnualReportArchive/t/LSE_TESCO_2014.pdf
TESCO	2013	0.1186	0.5984	https://www.annualreports.com/HostedData/AnnualReportArchive/t/LSE_TESCO_2013.pdf
TESCO	2012	0.1018	0.3649	https://www.annualreports.com/HostedData/AnnualReportArchive/t/LSE_TESCO_2012.pdf
TESCO	2011	0.1071	0.3715	https://www.annualreports.com/HostedData/AnnualReportArchive/t/LSE_TESCO_2011.pdf

Table 1. Cont.

Company	Fiscal Year	Polarity	Subjectivity	Links
UNDER ARMOUR	2016	0.0446	0.3534	https://underarmourinc.gcs-web.com/static-files/9801ef64-9915-4b77-af34-03e26689d3ab
UNDER ARMOUR	2015	0.0407	0.3655	https://underarmourinc.gcs-web.com/static-files/4f141906-99bf-46e4-b103-76c8f3fe6c6c
UNDER ARMOUR	2014	0.0353	0.3612	https://underarmourinc.gcs-web.com/static-files/13a42846-a519-469e-978b-b3199de9fbe8
UNDER ARMOUR	2013	0.0507	0.3622	https://underarmourinc.gcs-web.com/static-files/13a42846-a519-469e-978b-b3199de9fbe8
UNDER ARMOUR	2012	0.0506	0.3592	https://underarmourinc.gcs-web.com/static-files/90c5ec87-deb1-4f4f-9807-98634f83a697

5. Discussion

In our analysis of the three companies known for financial statement fraud—Wirecard, Tesco, and Under Armour—it is crucial to recognise the distinct nature of each case. While a common trend in fraudulent practices was initially sought, the diversity in the methods of fraud employed by these companies has instead revealed a broader spectrum of fraudulent behaviour. This lack of a unifying trend, rather than undermining the study’s validity, actually underscores the complexity and varied nature of financial statement fraud. It highlights the need for a nuanced approach to fraud detection that can adapt to different contexts and methods of deceit.

Furthermore, the unique characteristics of each case study provide valuable insights into the multifaceted nature of financial fraud. By dissecting each case independently, we gain a deeper understanding of how different types of fraud manifest and the varying linguistic and financial indicators that may signal their presence. This understanding is vital in developing more sophisticated and adaptable tools for fraud detection.

We conducted a detailed examination of the financial reports of three prominent companies (Wirecard, Tesco, and Under Armour) to explore the potential relationship between sentiment analysis and financial statement fraud. Each company has been associated with specific financial statement fraud cases, providing a valuable context for our analysis.

Utilising sentiment analysis, we assessed the polarity and subjectivity of the language used in the financial reports across multiple fiscal years. Polarity measures the overall positivity or negativity of the text, while subjectivity quantifies the degree of personal opinion and bias. The analysis yielded intriguing patterns that may indicate the companies’ attempts to portray their financial situations in particular ways.

The main highlights that were identified are remarkable:

- Wirecard: Known for fraudulent financial statements, the analysis revealed a consistent tone across the years, with a slight decrease in polarity and subjectivity in 2018.
- Tesco: With fraud related to the year 2014, Tesco exhibited a marked increase in polarity and subjectivity during the fraud year.
- Under Armour: During the fraud years of 2015–2016, Under Armour showed a decrease in polarity, while subjectivity remained moderately consistent.

Implications and Limitations

While the findings present interesting trends, it is crucial to recognise that sentiment analysis alone cannot definitively detect financial statement fraud. It offers insights into the tone and mood of the text but cannot uncover financial discrepancies or intentional deception. The results serve as supplementary information, providing a unique perspective on how companies may use language to shape perceptions of their financial health. However, a comprehensive assessment of financial statement fraud requires a multifaceted approach, integrating forensic accounting, financial ratio analysis, regulatory compliance, and more.

In the Wirecard Case (fraudulent financial statements), the polarity recorded a slight upward trend from 2014 to 2017, followed by a decrease in 2018, while in terms of sub-

jectivity, it can be identified as a consistent moderate subjectivity with a small decrease in 2018. The consistent polarity and subjectivity across the years may reflect a persistent tone and approach, possibly engineered to maintain investor confidence. However, the slight decrease in both scores in 2018 might warrant further investigation. It displayed a consistent tone over time yet showed a slight decline in both polarity and subjectivity in 2018. The stable polarity and subjectivity might suggest an attempt to sustain investor trust. The dip in 2018 could hint at underlying issues, meriting further probes (See Figure 3).

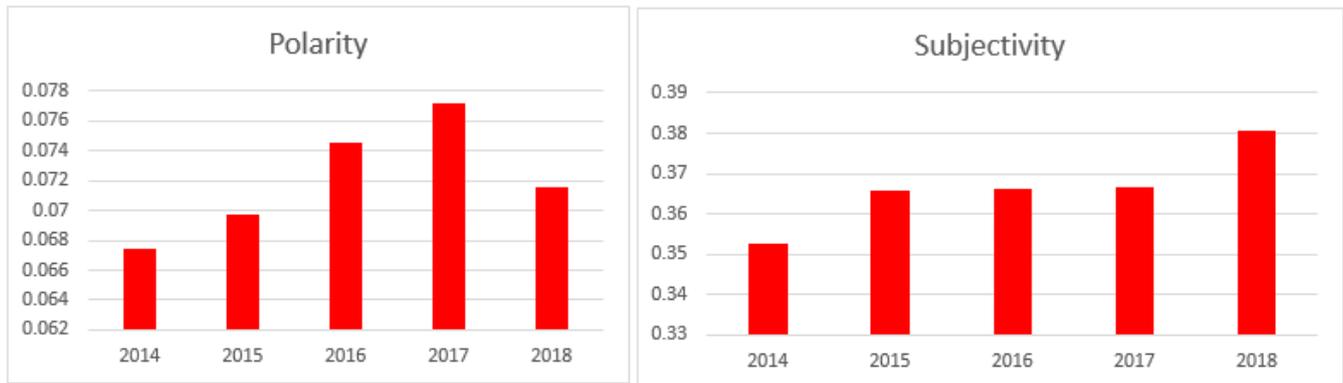


Figure 3. Wirecard Trends in polarity and subjectivity (fiscal years affected by financial statement frauds in red).

In the case of Tesco (Fraud Related to the Year 2014), polarity recorded a notable increase in polarity in 2014, followed by a decrease, while subjectivity had a significant increase in subjectivity in 2013, reaching its peak in 2014. The marked increase in polarity and subjectivity in 2014 (fraud year) might indicate an attempt to portray a more optimistic picture than the actual financial situation warranted. The fraud year 2014 marked a pronounced rise in both polarity and subjectivity. The peak in sentiment metrics in 2014 possibly indicates a more positive portrayal than the actual financial health (See Figure 4).

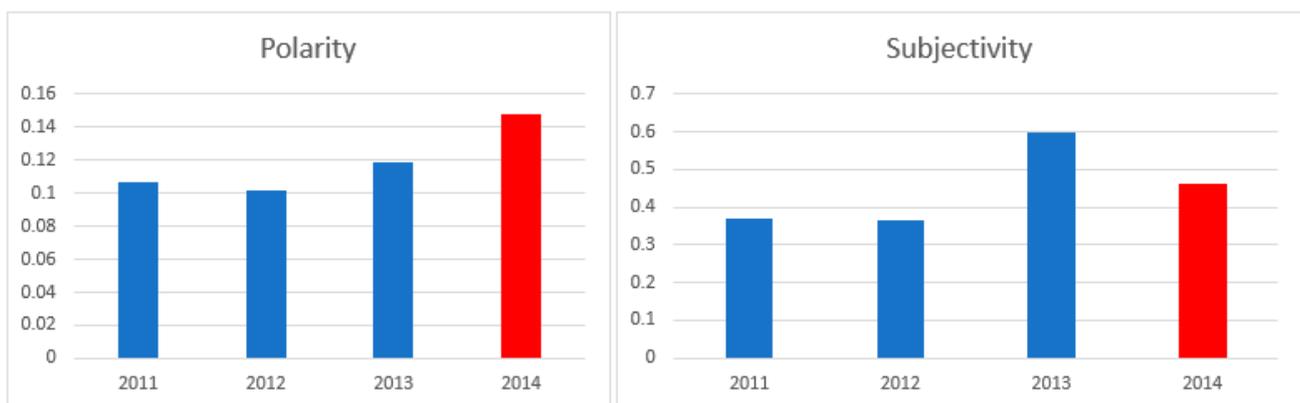


Figure 4. Tesco trends in polarity and subjectivity (fiscal years affected by financial statement frauds in red).

For Under Armour (fraud performed in the Years 2015–2016), polarity decreased during the fraud years (2015–2016), followed by an increase, while the subjectivity was moderate with slight fluctuations, but no clear pattern can be identified in this case. The decrease in polarity during the fraud years may reflect a subtle shift in language, perhaps due to the complexities of maintaining the facade. However, the subjectivity scores do not show a clear pattern. The 2015–2016 (fraud years) revealed a drop in polarity, while subjective scores remained fairly stable. The diminished polarity during 2015–2016 might

reflect challenges in upholding a deceptive front, but the subjectivity does not exhibit a discernible trend (See Figure 5).

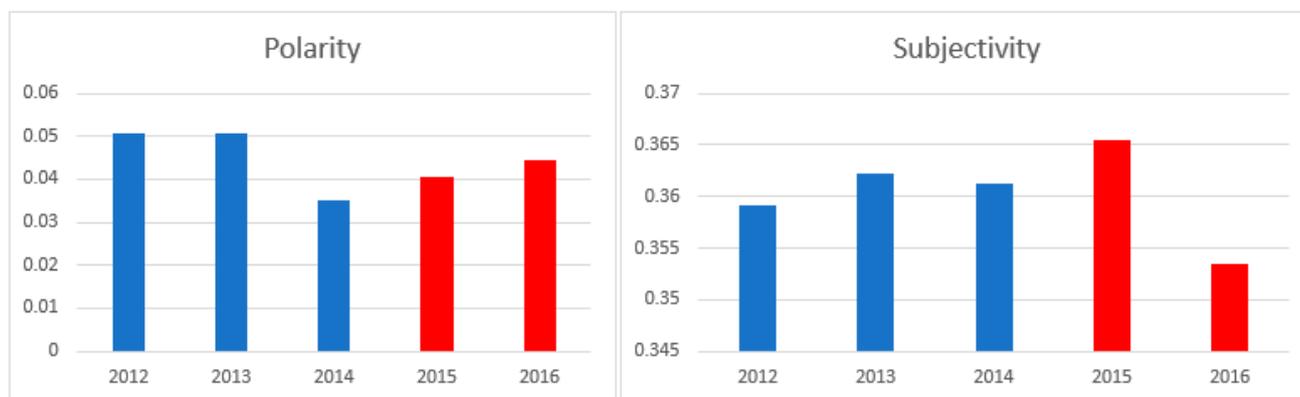


Figure 5. Under Armour trends in polarity and subjectivity (fiscal years affected by financial statement frauds in red).

While there are some interesting trends and patterns in the sentiment analysis scores, it is essential to note that sentiment analysis alone is not a reliable indicator of financial statement fraud. Sentiment analysis measures the tone and mood of the text but cannot detect discrepancies in financial data, accounting irregularities, or intentional deception.

The findings here may serve as supplementary information in a broader investigation, but financial statement fraud detection typically requires a multifaceted approach involving forensic accounting, financial ratio analysis, the examination of auditor's statements, and adherence to regulatory compliance [34]. Sentiment analysis can provide insights into how companies present their financials and may highlight unusual shifts in tone or subjectivity that warrant further scrutiny. However, it should be used with other robust financial analysis tools and techniques to provide a comprehensive assessment.

Suppose fraud investigators and auditors place disproportionate emphasis on the emotional tone and subjectivity of financial statements. In that case, they risk introducing significant errors in their assessments. Overreliance on sentiment could lead to false positives, where non-fraudulent statements are wrongly flagged due to transient negative tones. Conversely, false negatives might arise when deceptive entities craft statements with neutral or positive sentiments, masking underlying discrepancies. Moreover, financial language's inherent complexity and objectivity could mean sentiment analysis occasionally misinterprets context. By prioritising sentiment, there is also the danger of overshadowing crucial quantitative indicators, potentially overlooking tangible financial irregularities. While the results offer a nuanced lens through which one might view company financial disclosures, a holistic financial fraud detection approach demands integrating multiple methods like forensic accounting, financial ratio analysis, and regulatory compliance checks.

Our study conducted sentiment analysis at the document level, providing an overall sentiment score for all financial statements. However, while offering a broad overview, this approach does not capture the nuanced sentiment variations often present within different sections or sentences of a document. By their nature, financial statements can contain a mix of positive and negative sentiments, reflecting the multifaceted reality of a company's financial position and outlook.

For instance, a statement like "Our financial outlook is negative, several financial pointers are pointing downward. However, we are optimistic about the near future illustrating this complexity." Here, negative sentiments about the current financial status are juxtaposed with a positive outlook, making it challenging to accurately define the overall sentiment of the document. This example underscores the need for a more refined approach to dissect and interpret the varying sentiments more granularly.

Financial language is often intricate, employing industry-specific jargon and complex constructions that can convey multiple sentiments within a single sentence or paragraph. This complexity necessitates an analysis beyond the document level, focusing on paragraph or sentence-level interpretations to gauge the sentiment accurately. Moreover, using industry jargon, technical terms, and financial euphemisms can further complicate sentiment analysis. Terms that might typically carry a negative connotation in general language use could have a neutral or positive implication in a financial context, and vice versa.

In this study, the relationship between sentiment and fraudulent behaviour in financial statements has been a central focus, yet the connection has not been sufficiently elucidated. Our research hypothesises that certain sentiment patterns within financial communications can indicate underlying fraudulent activities. However, it is crucial to acknowledge that this relationship is not straightforward and that negative and excessively positive sentiments can be signals of fraud. An overly positive sentiment, for instance, might be employed to mask underlying financial issues or artificially inflate a company's perceived performance, thereby misleading stakeholders. This aspect of sentiment analysis in financial documents warrants a more comprehensive exploration, particularly considering the complexities of human psychology and cognitive biases.

Cognitive biases, such as overconfidence bias, can play a significant role in how corporate communications are framed [40,41]. Overconfidence in financial reporting, for instance, may lead to overly optimistic representations of a company's financial health or future prospects [42]. This bias could inadvertently or deliberately lead to misrepresentation, which may be detectable through analysing the sentiment expressed in these communications. In light of this, our study would benefit from a deeper examination of how cognitive biases influence the sentiment in financial statements and how this, in turn, correlates with fraudulent behaviour. Integrating a psychological perspective into our analysis could provide a more holistic understanding of the interplay between sentiment, cognitive biases, and financial fraud [43–45]. This approach would strengthen the theoretical foundation of our research and enhance the practical applicability of sentiment analysis in detecting financial statement fraud. In future work, a more nuanced exploration of these aspects, supported by empirical evidence, is necessary to establish a more definitive link between sentiment patterns and fraudulent activities in financial reporting.

6. Conclusions

The conclusions of this study stem from an explorative examination of financial reports from three prominent companies known for their fraudulent financial statements. This research sought to determine whether sentiment analysis could unveil hidden clues indicative of financial fraud.

The analysis of Wirecard's financial statements revealed consistent polarity and subjectivity over the years, followed by a slight decrease in 2018. This pattern may reflect a persistent tone intended to maintain investor confidence.

Tesco's marked increase in polarity and subjectivity in 2014, the fraud year, might symbolise an attempt to project an optimistic picture inconsistent with the actual financial situation. This change in tone could be a subtle sign of underlying manipulation.

Under Armour's results were more ambiguous, with decreased polarity during the fraud years but no clear pattern in subjectivity. This inconsistency could reflect the complexities of maintaining a fraudulent façade, though it is less conclusive.

The findings suggest that sentiment analysis could serve as a complementary tool to traditional methods for detecting financial fraud. By analysing the emotional tone of financial statements, subtle clues or inconsistencies in language might be exposed, providing additional insights into potential fraudulent activities.

The study's limitations are noteworthy. Sentiment analysis alone cannot detect financial fraud since it focuses on the textual tone rather than numerical discrepancies or accounting irregularities. Financial statements are often factual and objective, so sentiment may not always reveal the underlying fraudulent activities. The complexity of financial

language and the potential for false positives also pose challenges, as does the lack of domain-specific knowledge inherent in generic sentiment analysis techniques.

The intriguing results of this study pave the way for further research. Combining sentiment analysis with quantitative financial metrics could yield a more comprehensive view of potential fraud. Creating domain-specific sentiment analysis models tailored to financial contexts might enhance detection accuracy. Extending the research to different industries and regions or exploring other Natural Language Processing techniques could provide deeper insights. Integrating automated systems incorporating sentiment analysis for continuous monitoring represents another promising avenue.

Additionally, the proactive use of emerging technologies like blockchain may play a crucial role in fraud prevention rather than merely detecting them *ex post*. Blockchain's decentralised and immutable nature can add layers of security and transparency to financial transactions, thereby reducing opportunities for fraud [46–48].

This research underscores the potential value of sentiment analysis in the complex and critical field of financial fraud detection. While not a standalone solution, it offers a novel perspective that could enrich traditional analysis methods. The study's limitations highlight the challenges of this interdisciplinary approach, but they also outline a path for future research that could lead to more nuanced and effective tools for ensuring financial transparency and integrity. By continuing to explore the intersection of linguistics, finance, and technology, new opportunities for understanding and combating financial statement fraud may emerge.

Future Research

To enhance the robustness of our findings, a comparative analysis with companies with correct financial statements is essential. To address this, we propose extending our research to include a control group comprising companies with a history of accurate and transparent financial reporting [49,50]. This comparison will enable us to identify any overlapping corporate communications and financial reporting characteristics between the fraudulent and non-fraudulent companies [51,52].

By analysing the sentiment and linguistic patterns in the financial statements of these 'clean' companies, we aim to establish a baseline for normal corporate communication and financial reporting behaviour [53–55]. This baseline will serve as a point of comparison to more effectively identify deviations [56] in the reports of companies engaging in fraudulent activities. Including such a comparative analysis is expected to yield a more comprehensive understanding of the linguistic and financial indicators of fraud and will strengthen the study's overall conclusions.

This comparative approach will also allow us to explore the hypothesis that certain linguistic patterns, initially presumed to indicate fraudulent activity, may also be present in legitimate reports [57–60]. Identifying these patterns will refine the accuracy of sentiment analysis as a tool for fraud detection, ensuring that it is sensitive to genuine indicators of fraud while minimising false positives.

Our study acknowledges the need for a more sophisticated approach to sentiment analysis in financial documents. Future research should focus on developing and applying methods that can perform sentiment analysis at the sentence or paragraph level [61–67]. This approach would enable a more accurate understanding of the sentiments conveyed in financial statements, accounting for the complex interplay of positive and negative statements within a document. Furthermore, incorporating examples and case studies that highlight these semantic challenges would significantly enhance the depth of the research [68,69]. A detailed examination of specific instances where mixed sentiments are present in financial statements would provide a more comprehensive understanding of the challenges and potential solutions [70] in applying sentiment analysis to financial documents.

Author Contributions: Conceptualisation, A.F. and J.M.; methodology, A.F.; software, A.F.; validation, J.M. and B.G.; formal analysis, A.F. and B.G.; investigation, A.F. and J.M.; resources, A.F.; data curation, A.F.; writing—original draft preparation, A.F.; writing—review and editing, J.M. and B.G.;

visualisation, A.F.; supervision, A.F.; project administration, A.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data used are publicly available. We used financial statements available on the companies' websites. All links are provided in the related sections.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Tonkiss, F. Trust, confidence and economic crisis. *Intereconomics* **2009**, *44*, 196–202. [CrossRef]
2. Fung, B. The demand and need for transparency and disclosure in corporate governance. *Univers. J. Manag.* **2014**, *2*, 72–80. [CrossRef]
3. Gardi, B.; Abdalla Hamza, P.; Sabir, B.Y.; Mahmood Aziz, H.; Sorguli, S.; Abdullah, N.N.; Al-Kake, F. Investigating the effects of financial accounting reports on managerial decision-making in small and medium-sized enterprises. *Turk. J. Comput. Math. Educ.* **2021**, *12*, 2134–2142. [CrossRef]
4. Currie, W.L.; Gozman, D.P.; Seddon, J.J. Dialectic tensions in the financial markets: A longitudinal study of pre-and post-crisis regulatory technology. *J. Inf. Technol.* **2018**, *33*, 304–325. [CrossRef]
5. Rezaee, Z. *Corporate Governance and Ethics*; John Wiley & Sons: Hoboken, NJ, USA, 2008.
6. Rezaee, Z. Causes, consequences, and deterrence of financial statement fraud. *Crit. Perspect. Account.* **2005**, *16*, 277–298. [CrossRef]
7. Triani, N. Fraudulent Financial Reporting Detection Using Beneish M-Score Model in Public Companies in 2012–2016. *Asia Pac. Fraud J.* **2019**, *4*, 27–42.
8. Anderson, J.R.; Tirrell, M.E. Too Good to Be True CEOs and Financial Reporting Fraud. *Consult. Psychol. J. Pract. Res.* **2004**, *56*, 35. [CrossRef]
9. Tirole, J. *The Theory of Corporate Finance*; Princeton University Press: Princeton, NJ, USA, 2010.
10. Asare, S.K.; Wright, A.; Zimbelman, M.F. Challenges facing auditors in detecting financial statement fraud: Insights from fraud investigations. *J. Forensic Investig. Account.* **2015**, *7*, 63–111.
11. Rezaee, Z. *Financial Statement Fraud: Prevention and Detection*; John Wiley & Sons: Hoboken, NJ, USA, 2002.
12. Segal, S. Accounting frauds—review of advanced technologies to detect and prevent frauds. *Econ. Bus. Rev.* **2016**, *2*, 45–64. [CrossRef]
13. Dorminey, J.; Fleming, A.S.; Kranacher, M.J.; Riley, R.A., Jr. The evolution of fraud theory. *Issues Account. Educ.* **2012**, *27*, 555–579. [CrossRef]
14. Craja, P.; Kim, A.; Lessmann, S. Deep learning for detecting financial statement fraud. *Decis. Support Syst.* **2020**, *139*, 113421. [CrossRef]
15. Leuz, C.; Wysocki, P.D. Economic Consequences of Financial Reporting and Disclosure Regulation: A Review and Suggestions for Future Research. 2008. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1105398 (accessed on 10 March 2021).
16. Roychowdhury, S.; Shroff, N.; Verdi, R.S. The effects of financial reporting and disclosure on corporate investment: A review. *J. Account. Econ.* **2019**, *68*, 101246. [CrossRef]
17. Luitel, D. A Language-Model-Based Approach for Detecting Incompleteness in Natural-Language Requirements. Ph.D. Dissertation, University of Ottawa, Ottawa, ON, Canada, 2023.
18. Kang, Y.; Cai, Z.; Tan, C.W.; Huang, Q.; Liu, H. Natural language processing (NLP) in management research: A literature review. *J. Manag. Anal.* **2020**, *7*, 139–172. [CrossRef]
19. Lewis, D.D.; Jones, K.S. Natural language processing for information retrieval. *Commun. ACM* **1996**, *39*, 92–101. [CrossRef]
20. Felt, C.; Riloff, E. Recognising euphemisms and dysphemisms using sentiment analysis. In Proceedings of the Second Workshop on Figurative Language Processing, Online, July 2020; pp. 136–145.
21. Goel, S.; Uzuner, O. Do sentiments matter in fraud detection? Estimating semantic orientation of annual reports. *Intell. Syst. Account. Financ. Manag.* **2016**, *23*, 215–239. [CrossRef]
22. Chen, Z.; Huang, L.; Yu, Z.; Yang, W.; Li, L.; Zheng, X.; Zhao, X. Linguistic Steganography Detection Using Statistical Characteristics of Correlations between Words. In *Proceedings of the Information Hiding: 10th International Workshop, IH 2008, Santa Barbara, CA, USA, 19–21 May 2008*; Revised Selected Papers 10; Springer: Berlin/Heidelberg, Germany, 2008; pp. 224–235.
23. Hirsch, T.; Hofer, B. Detecting non-natural language artifacts for de-noising bug reports. *Autom. Softw. Eng.* **2022**, *29*, 52. [CrossRef] [PubMed]
24. Iqbal, S.; Hassan, S.U.; Aljohani, N.R.; Alelyani, S.; Nawaz, R.; Bornmann, L. A decade of in-text citation analysis based on natural language processing and machine learning techniques: An overview of empirical studies. *Scientometrics* **2021**, *126*, 6551–6599. [CrossRef]
25. Caldarini, G.; Jaf, S.; McGarry, K. A literature survey of recent advances in chatbots. *Information* **2022**, *13*, 41. [CrossRef]
26. Rogers, A.; Baldwin, T.; Leins, K. Just What do You Think You're Doing, Dave? A Checklist for Responsible Data Use in NLP. *arXiv* **2021**, arXiv:2109.06598.

27. Aghili, S. *Fraud Auditing Using CAATT: A Manual for Auditors and Forensic Accountants to Detect Organisational Fraud*; CRC Press: Boca Raton, FL, USA, 2019.
28. Huang, B.; Guo, R.; Zhu, Y.; Fang, Z.; Zeng, G.; Liu, J.; Wang, Y.; Fujita, H.; Shi, Z. Aspect-level sentiment analysis with aspect-specific context position information. *Knowl. -Based Syst.* **2022**, *243*, 108473. [[CrossRef](#)]
29. Fei, H.; Zhang, Y.; Ren, Y.; Ji, D. Latent Emotion Memory for Multi-Label Emotion Classification. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; Volume 34, pp. 7692–7699.
30. Fei, H.; Li, F.; Li, C.; Wu, S.; Li, J.; Ji, D. Inheriting the Wisdom of Predecessors: A Multiplex Cascade Framework for Unified Aspect-Based Sentiment Analysis. In Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI, Vienna, Austria, 23–29 July 2022; pp. 4096–4103.
31. Sohangir, S.; Wang, D.; Pomeranets, A.; Khoshgoftaar, T.M. Big Data: Deep Learning for financial sentiment analysis. *J. Big Data* **2018**, *5*, 3. [[CrossRef](#)]
32. Xiang, C.; Zhang, J.; Li, F.; Fei, H.; Ji, D. A semantic and syntactic enhanced neural model for financial sentiment analysis. *Inf. Process. Manag.* **2022**, *59*, 102943. [[CrossRef](#)]
33. Shang, L.; Xi, H.; Hua, J.; Tang, H.; Zhou, J. A Lexicon Enhanced Collaborative Network for targeted financial sentiment analysis. *Inf. Process. Manag.* **2023**, *60*, 103187. [[CrossRef](#)]
34. Faccia, A.; Petratos, P. NLP and IR Applications for Financial Reporting and Non-Financial Disclosure. Framework Implementation and Roadmap for Feasible Integration with the Accounting Process. In Proceedings of the 6th International Conference on Natural Language Processing and Information Retrieval, Bangkok, Thailand, 16–18 December 2022; pp. 117–124.
35. Faccia, A.; Mosteanu, N.R. Accounting and blockchain technology: From double-entry to triple-entry. *Bus. Manag. Rev.* **2019**, *10*, 108–116.
36. Mosteanu, N.R.; Faccia, A. Digital systems and new challenges of financial management—FinTech, XBRL, blockchain and cryptocurrencies. *Qual.-Access Success* **2020**, *21*, 159–166.
37. Faccia, A.; Petratos, P. Blockchain, enterprise resource planning (ERP) and accounting information systems (AIS): Research on e-procurement and system integration. *Appl. Sci.* **2021**, *11*, 6792. [[CrossRef](#)]
38. Faccia, A.; Mosteanu, N.R. Tax evasion, information systems and blockchain. *J. Inf. Syst. Oper. Manag.* **2019**, *13*, 65–74.
39. Mosteanu, N.R.; Faccia, A. Fintech frontiers in quantum computing, fractals, and blockchain distributed ledger: Paradigm shifts and open innovation. *J. Open Innov. Technol. Mark. Complex.* **2021**, *7*, 19. [[CrossRef](#)]
40. Tetlock, P.E. Cognitive biases and organizational correctives: Do both disease and cure depend on the politics of the beholder? *Adm. Sci. Q.* **2000**, *45*, 293–326. [[CrossRef](#)]
41. Langevoort, D.C. Disasters and Disclosures: Securities Fraud Liability in the Shadow of a Corporate Catastrophe. *Geo. LJ* **2018**, *107*, 967.
42. French, S.; Maule, J.; Papamichail, N. *Decision Behaviour, Analysis and Support*; Cambridge University Press: Cambridge, UK, 2009.
43. Hirshleifer, D. Psychological bias as a driver of financial regulation. *Eur. Financ. Manag.* **2008**, *14*, 856–874. [[CrossRef](#)]
44. Raval, V. A disposition-based fraud model: Theoretical integration and research agenda. *J. Bus. Ethics* **2018**, *150*, 741–763. [[CrossRef](#)]
45. Repenning, N.; Löhlein, L.; Schäffer, U. Emotions in accounting: A review to bridge the paradigmatic divide. *Eur. Account. Rev.* **2022**, *31*, 241–267. [[CrossRef](#)]
46. Faccia, A.; Al Naqbi, M.Y.K.; Lootah, S.A. August. Integrated Cloud Financial Accounting Cycle: How Artificial Intelligence, Blockchain, and XBRL Will Change the Accounting, Fiscal and Auditing Practices. In Proceedings of the 2019 3rd International Conference on Cloud and Big Data Computing, Oxford, UK, 28–30 August 2019; pp. 31–37.
47. Faccia, A.; Pandey, V.; Banga, C. Is permissioned blockchain the key to support the external audit shift to entirely open innovation paradigm? *J. Open Innov. Technol. Mark. Complex.* **2022**, *8*, 85. [[CrossRef](#)]
48. Petratos, P.N.; Faccia, A. Fake news, misinformation, disinformation and supply chain risks and disruptions: Risk management and resilience using blockchain. *Ann. Oper. Res.* **2023**, *327*, 735–762. [[CrossRef](#)]
49. Armstrong, C.S.; Guay, W.R.; Weber, J.P. The role of information and financial reporting in corporate governance and debt contracting. *J. Account. Econ.* **2010**, *50*, 179–234. [[CrossRef](#)]
50. Doyle, J.; Ge, W.; McVay, S. Determinants of weaknesses in internal control over financial reporting. *J. Account. Econ.* **2007**, *44*, 193–223. [[CrossRef](#)]
51. Dalnial, H.; Kamaluddin, A.; Sanusi, Z.M.; Khairuddin, K.S. Accountability in financial reporting: Detecting fraudulent firms. *Procedia-Soc. Behav. Sci.* **2014**, *145*, 61–69. [[CrossRef](#)]
52. Zainudin, E.F.; Hashim, H.A. Detecting fraudulent financial reporting using financial ratio. *J. Financ. Report. Account.* **2016**, *14*, 266–278. [[CrossRef](#)]
53. Ho, J.; Cheng, W. Metaphors in financial analysis reports: How are emotions expressed? *Engl. Specif. Purp.* **2016**, *43*, 37–48. [[CrossRef](#)]
54. Türegün, N. Text mining in financial information. *Curr. Anal. Econ. Financ.* **2019**, *1*, 18–26.
55. El-Haj, M.; Rayson, P.; Walker, M.; Young, S.; Simaki, V. In search of meaning: Lessons, resources and next steps for computational analysis of financial discourse. *J. Bus. Financ. Account.* **2019**, *46*, 265–306. [[CrossRef](#)]
56. Busk, P.L.; Serlin, R.C. Meta-Analysis for Single-Case Research. In *Single-Case Research Design and Analysis (Psychology Revivals)*; Routledge: Oxfordshire, UK, 2015; pp. 187–212.

57. Bolton, R.J.; Hand, D.J. Statistical fraud detection: A review. *Stat. Sci.* **2002**, *17*, 235–255. [[CrossRef](#)]
58. Ras, I.A. A Corpus-Assisted Critical Discourse Analysis of the Reporting on Corporate Fraud by UK Newspapers 2004–2014. Ph.D. Dissertation, University of Leeds, Leeds, UK, 2017.
59. Neu, D.; Warsame, H.; Pedwell, K. Managing public impressions: Environmental disclosures in annual reports. *Account. Organ. Soc.* **1998**, *23*, 265–282. [[CrossRef](#)]
60. Boiral, O. Sustainability reports as simulacra? A counter-account of A and A+ GRI reports. *Account. Audit. Account. J.* **2013**, *26*, 1036–1071. [[CrossRef](#)]
61. Appel, O.; Chiclana, F.; Carter, J.; Fujita, H. A hybrid approach to the sentiment analysis problem at the sentence level. *Knowl.-Based Syst.* **2016**, *108*, 110–124. [[CrossRef](#)]
62. Poria, S.; Cambria, E.; Winterstein, G.; Huang, G.B. Sentic patterns: Dependency-based rules for concept-level sentiment analysis. *Knowl.-Based Syst.* **2014**, *69*, 45–63. [[CrossRef](#)]
63. Nasukawa, T.; Yi, J. Sentiment Analysis: Capturing Favorability Using Natural Language Processing. In Proceedings of the 2nd International Conference on Knowledge Capture, Sanibel Island, FL, USA, 23–25 October 2003; pp. 70–77.
64. Zucco, C.; Calabrese, B.; Agapito, G.; Guzzi, P.H.; Cannataro, M. Sentiment analysis for mining texts and social networks data: Methods and tools. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2020**, *10*, e1333. [[CrossRef](#)]
65. Zhang, L.; Wang, S.; Liu, B. Deep learning for sentiment analysis: A survey. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2018**, *8*, e1253. [[CrossRef](#)]
66. Bengfort, B.; Bilbro, R.; Ojeda, T. *Applied Text Analysis with Python: Enabling Language-Aware Data Products with Machine Learning*; O’Reilly Media, Inc.: Sebastopol, CA, USA, 2018.
67. Simperl, E. Reusing ontologies on the Semantic Web: A feasibility study. *Data Knowl. Eng.* **2009**, *68*, 905–925. [[CrossRef](#)]
68. Zhang, Y.; Sidibé, D.; Morel, O.; Mériaudeau, F. Deep multimodal fusion for semantic image segmentation: A survey. *Image Vis. Comput.* **2021**, *105*, 104042. [[CrossRef](#)]
69. Zhu, X.; Ao, X.; Qin, Z.; Chang, Y.; Liu, Y.; He, Q.; Li, J. Intelligent financial fraud detection practices in post-pandemic era. *Innov.* **2021**, *2*, 100176. [[CrossRef](#)] [[PubMed](#)]
70. Harner, M.M. Barriers to effective risk management. *Seton Hall L. Rev.* **2010**, *40*, 1323.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.