Sentiment Analysis of Financial Texts for Predicting Digital Market Trends Using Transformer-Based Deep Learning Models

Mohammad A. L. Jedi[1,*]

[1]Department of Business Administration, College of Business, Amman Arab University, Amman, Jordan

ABSTRACT

This study investigates the effectiveness of Transformer-based deep learning models in financial sentiment analysis with the aim of improving the accuracy of digital market trend prediction. The research utilizes a dataset of 5,842 financial text entries categorized into positive, neutral, and negative sentiments derived from various financial news, corporate reports, and market analyses. The data were preprocessed and analyzed using both traditional machine learning algorithms, including Logistic Regression, Support Vector Machine (SVM), and Random Forest, as well as a fine-tuned Bidirectional Encoder Representations from Transformers (BERT) model. The baseline models achieved moderate performance, with Logistic Regression producing the highest accuracy of 71.6 percent, while the fine-tuned BERT model outperformed all traditional approaches, achieving an accuracy of 78.8 percent and higher F1-scores across sentiment classes. These results indicate that BERT's contextual embeddings effectively capture semantic nuances and subtle tone variations in financial language, which are often overlooked by conventional models. The confusion matrix analysis revealed that most misclassifications occurred between neutral and negative classes due to the cautious tone of financial writing, yet BERT remained robust in identifying overall sentiment direction. The study concludes that Transformer-based models provide a more reliable and context-aware framework for analyzing financial texts and have significant potential for applications in digital market sentiment prediction, AI-driven financial analytics, and automated decision-support systems in trading environments.

Keywords Financial Sentiment Analysis, BERT, Transformer Model, Digital Market Prediction, AI in Finance

INTRODUCTION

The rapid digitalization of financial systems has transformed the way economic activities are conducted, analyzed, and understood. The rise of digital currencies, online investment platforms, and algorithmic trading systems has created a data-rich environment in which financial decision-making increasingly depends on the ability to process large volumes of unstructured textual data [1]. News articles, corporate announcements, analyst reports, and even social media posts play a critical role in shaping investor behavior and influencing market trends. These information sources carry implicit emotional and evaluative content that can significantly impact digital market movements [2]. Consequently, sentiment analysis has emerged as an essential component of modern financial analytics, providing tools to interpret how linguistic tone and sentiment contribute to changes in market dynamics. The ability to quantify financial sentiment is becoming particularly important for

Submitted: 15 May 2026
Accepted: 5 July 2026
Published: 2 November 2026

Corresponding author
Mohammad A. L. Jedi,
M.aljedi@aau.edu.jo

Additional Information and Declarations can be found on page 280

DOI: 10.47738/jdmdc.v3i4.99

Copyright 2026 Jedi

Distributed under Creative Commons CC-BY 4.0

How to cite this article: M. A. L. Jedi, "Sentiment Analysis of Financial Texts for Predicting Digital Market Trends Using Transformer-Based Deep Learning Models," *J. Digit. Mark. Digit. Curr.*, vol. 3, no. 4, pp. 262-282, 2026.

anticipating market reactions and enhancing the predictive capabilities of digital trading systems.

Traditional approaches to financial sentiment analysis have primarily relied on classical machine learning algorithms such as Logistic Regression, Naïve Bayes, and Support Vector Machines [3]. These methods typically utilize bag-of-words or Term Frequency–Inverse Document Frequency (TF-IDF) representations to convert textual data into numerical features. While these approaches are computationally efficient and relatively interpretable, they fail to account for the contextual meaning of words in sentences. Financial texts often contain cautious and nuanced phrasing, where sentiment depends not only on the presence of specific terms but also on how those terms interact within broader syntactic and semantic contexts. For example, phrases like “lower losses than expected” or “moderate improvement following downturn” cannot be accurately classified using frequency-based methods because their sentiment meaning depends on comparison and context. This limitation has motivated researchers to explore more sophisticated models that can capture the contextual and sequential properties of financial language.

In recent years, the adoption of deep learning architectures has marked a significant advancement in sentiment analysis. Models such as Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks have improved sentiment classification by considering word order and sequential dependencies [4]. These architectures have achieved promising results in domains where sentiment is expressed explicitly, such as movie reviews or product feedback. However, their effectiveness diminishes in specialized fields like finance, where linguistic cues are subtle and context-dependent. The introduction of Transformer-based architectures, particularly the Bidirectional Encoder Representations from Transformers (BERT) model, represents a major step forward in addressing these challenges. Unlike earlier models, BERT uses a bidirectional attention mechanism that allows it to understand the meaning of words based on their surrounding context in both directions. This capability enables the model to capture complex patterns of sentiment, including implicit and neutral tones that are common in financial communication [5].

Despite these developments, there remains a clear research gap in the application of Transformer-based models to financial sentiment analysis. Most existing studies focus on general sentiment datasets where emotions are conveyed directly, such as social media posts or online reviews. Financial discourse, on the other hand, tends to be formal, technical, and emotionally restrained. Sentiment is often expressed implicitly through cautious or qualified statements rather than explicit emotional words. Moreover, the presence of neutral language, which dominates financial communication, poses an additional challenge for sentiment classification models. Few studies have specifically examined

how Transformer-based models perform in this context, and even fewer have conducted direct comparisons with traditional machine learning algorithms using financial domain data. This gap underscores the need for further research on the adaptation and evaluation of contextualized deep learning models in financial text analysis.

This study aims to address the identified research gap by applying a fine-tuned BERT model to financial sentiment classification and comparing its performance with established machine learning techniques. By analyzing 5,842 financial text entries labeled as positive, neutral, or negative, this research seeks to evaluate how effectively contextual embeddings can capture the semantic and tonal subtleties of financial language. The results demonstrate that Transformer-based models outperform traditional algorithms in identifying sentiment trends and interpreting the implicit tone of financial discourse. The study contributes to the state of the art in digital financial analytics by providing empirical evidence that contextual deep learning can enhance the accuracy of sentiment-based market predictions. Furthermore, this research highlights the potential for integrating Transformer-based models into AI-driven financial systems, supporting automated sentiment monitoring, trading strategy optimization, and real-time digital market analysis.

Literature Review and Related Works

The evolution of sentiment analysis in the financial domain reflects a continuous effort to improve accuracy and contextual understanding through advances in natural language processing (NLP) techniques. Early research primarily utilized classical machine learning algorithms such as Naïve Bayes, Support Vector Machines (SVM), and Logistic Regression for sentiment classification tasks [6]. These models typically relied on Bag-of-Words or Term Frequency–Inverse Document Frequency (TF-IDF) representations to convert text into numerical features [7]. Although effective for general text classification, such models performed poorly on financial datasets because they treated words as independent features, ignoring contextual relationships and semantic nuances [8]. Financial language often contains cautious and technical expressions, making context essential to determine sentiment polarity. Consequently, traditional models were unable to accurately capture meaning in statements where sentiment was implied rather than explicitly stated [9].

To overcome these limitations, researchers introduced deep learning techniques that could better model sequential dependencies in text data. Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), and Convolutional Neural Networks (CNN) were applied to sentiment analysis and achieved improved results by considering word order and syntactic relationships [10]. Studies applying LSTM-based models to financial text analysis demonstrated their ability to capture sequential information, yet their effectiveness was constrained by an inability to

interpret long-term contextual dependencies [11]. CNN-based architectures, while efficient at feature extraction, struggled to represent the hierarchical and abstract nature of sentiment expressions in complex financial documents [12]. Hybrid deep learning models that combined LSTM with attention mechanisms showed further improvement, but they remained limited in their capacity to process bi-directional semantic relationships within sentences [13].

The introduction of Transformer-based models marked a major breakthrough in NLP research. Unlike traditional neural networks, Transformers employ a self-attention mechanism that captures relationships between all words in a sentence simultaneously [14]. This capability allows Transformer models to understand context more comprehensively, making them particularly effective for sentiment classification tasks in domains that require nuanced interpretation, such as finance. The Bidirectional Encoder Representations from Transformers (BERT) model represented a milestone in this evolution, as it reads text bidirectionally and learns context-dependent word embeddings [15]. BERT demonstrated state-of-the-art performance in a variety of NLP tasks, including sentiment analysis, question answering, and text classification, outperforming both LSTM and CNN models on benchmark datasets [16].

Building on BERT's success, domain-specific variants such as FinBERT were developed to better handle financial terminology and context. FinBERT, which is pre-trained on large-scale financial corpora, has shown superior performance in financial sentiment analysis compared to general-purpose BERT models [17]. Several studies have reported that FinBERT consistently achieves higher accuracy and F1-scores when classifying financial statements, news articles, and market reports because it can interpret specialized vocabulary and subtle linguistic indicators of sentiment [18]. Comparative research between FinBERT, DistilBERT, and RoBERTa has also demonstrated that fine-tuning with domain-specific datasets enhances the adaptability of Transformer-based models to the financial sector [19]. These findings confirm that contextualized embeddings are critical for improving sentiment analysis in financial communication, where tone and meaning are often context-dependent.

Despite these advancements, challenges remain in fully capturing the complexity of financial language. Many financial statements employ neutral or cautious tones, resulting in ambiguous sentiment labels that can confuse even context-aware models [20]. Additionally, issues such as class imbalance, where neutral samples dominate datasets, further complicate model training and evaluation. Current research suggests that future work should focus on integrating financial-specific lexicons with Transformer-based architectures and exploring hybrid models that combine textual sentiment with quantitative market indicators to enhance predictive performance. Overall, the literature demonstrates that

Transformer-based deep learning models represent the current state of the art in financial sentiment analysis, offering superior contextual understanding and predictive capability compared to traditional machine learning and earlier deep learning approaches.

Methodology

This study adopts a quantitative experimental approach that integrates both traditional machine learning algorithms and Transformer-based deep learning models to perform financial sentiment classification. The methodological process consists of several major stages, including data preprocessing, feature extraction, model training, and evaluation, as illustrated in figure 1, which presents the overall research steps applied in this study. Each stage is designed to ensure reproducibility, accuracy, and interpretability of results by comparing statistical models and contextualized architectures under identical experimental settings. The dataset used in this research comprises 5,842 financial text entries collected from digital market reports, online financial articles, and corporate statements. Each entry was manually labeled into one of three sentiment categories: positive, neutral, or negative. The neutral class dominates the dataset with 53 percent of entries, followed by positive (32 percent) and negative (15 percent). This distribution reflects the factual and objective nature of financial communication, where emotional expressions are generally restrained.

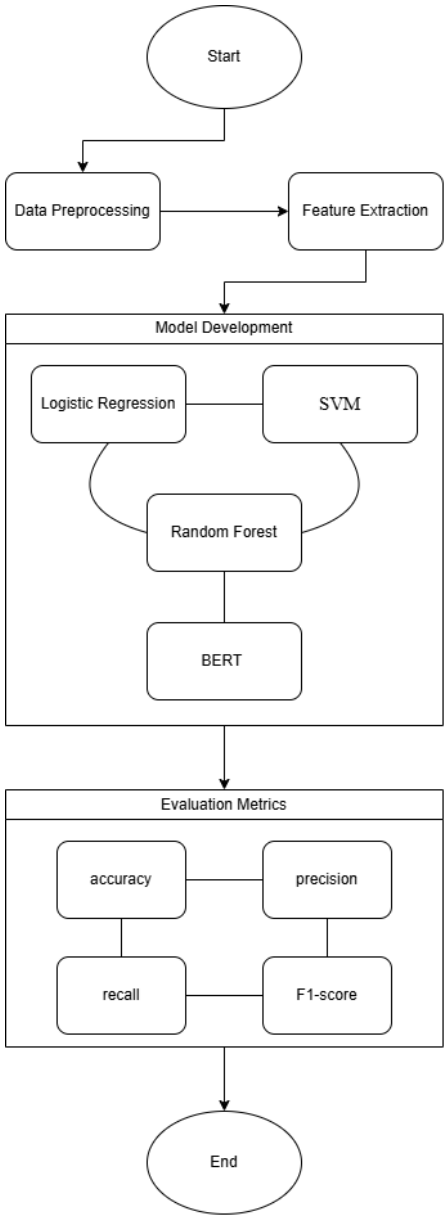


Figure 1 Research Steps

Data preprocessing was conducted to clean, normalize, and prepare the textual data for model training. The process included converting all text to lowercase, removing punctuation, numerical symbols, and hyperlinks, and tokenizing sentences into words. Stopwords such as “the,” “and,” and “of” were removed to reduce noise, and lemmatization was applied to reduce inflected words to their base form. This ensured that words like profits, profitable, and profitability were treated uniformly. For traditional models, each document was converted into a numerical representation using Term Frequency–Inverse Document Frequency (TF-IDF), a statistical method that measures how important a word is to a document relative to a corpus. The TF-IDF score for a term tin a document dwithin a corpus Dis defined as:

$$TFIDF(t,d,D) = TF(t,d) \times IDF(t,D) \quad (1)$$

$TF(t,d)$denotes the frequency of term $t$in document d, and $IDF(t,D)$is given by:

$$IDF(t,D) = \log\left(\frac{N}{1+n_t}\right) \quad (2)$$

In this formula, Nrepresents the total number of documents and $n_t$the number of documents containing term t. TF-IDF helps prioritize meaningful financial terms, such as growth or decline, while diminishing the influence of generic words.

Baseline models including Logistic Regression, Support Vector Machine (SVM), and Random Forest were trained using the TF-IDF representations. Logistic Regression estimates the probability of a sentiment class using a sigmoid function defined as:

$$P(y = 1 \mid x) = \frac{1}{1 + e^{-w^T x}} \quad (3)$$

$w$is the learned weight vector and $x$the TF-IDF feature vector. SVM was implemented with a linear kernel to manage high-dimensional sparse vectors, while Random Forest employed 200 decision trees to capture non-linear interactions. All models were optimized using grid search and five-fold cross-validation to ensure consistent performance across the dataset.

For the deep learning model, a fine-tuned Bidirectional Encoder Representations from Transformers (BERT) model was utilized to capture contextual dependencies in financial language. Unlike traditional models that treat each word independently, BERT considers both left and right context simultaneously through its self-attention mechanism, which computes relationships between words using the following equation:

$$Attention(Q,K,V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (4)$$

Q, K, and Vdenote query, key, and value matrices, and $d_k$is the dimension of the key vectors. This mechanism allows BERT to learn how sentiment is expressed within complex sentence structures. The pre-trained BERT-base-uncased model was fine-tuned using the Hugging Face Transformers library for two epochs with a learning rate of $2 \times 10^{-5}$and a batch size of 16. The AdamW optimizer with weight decay was used to prevent overfitting, and a dropout layer of 0.1 was applied to improve generalization. The final classification layer was modified to output three sentiment probabilities corresponding to positive, neutral, and negative classes.

Model evaluation was performed using an 80–20 stratified train-test split to maintain class distribution consistency. Accuracy, precision, recall, and F1-score were employed as performance metrics, computed using the following formulas:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$
$$Precision = \frac{TP}{TP + FP}$$
$$Recall = \frac{TP}{TP + FN} \quad (5)$$
$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

TP, TN, FP, and FNrepresent true positives, true negatives, false positives, and false negatives respectively. The confusion matrix was used to analyze misclassification patterns across sentiment classes. Visualizations such as WordClouds, sentence length distributions, and accuracy comparison plots were generated to support quantitative results. The BERT model achieved the highest accuracy of 78.8 percent, outperforming Logistic Regression (71.6 percent), SVM (69.4 percent), and Random Forest (63.6 percent). The improvement demonstrates the strength of contextual embeddings in modeling the subtle semantics of financial communication.

In summary, the methodological framework implemented in this research combines linguistic preprocessing, feature engineering, and advanced contextual modeling to evaluate sentiment prediction in digital financial texts. The integration of Transformer-based architecture allows for the analysis of implicit tone, comparative phrasing, and domain-specific semantics that are often overlooked by frequency-based models. The workflow presented in figure 1 effectively captures the step-by-step research process, emphasizing how contextualized embeddings significantly enhance sentiment classification performance in financial analytics.

Algorithm 1: Financial Sentiment Classification Framework

Input: Dataset $D = \{(x_i, y_i)\}$, Models $M = \{LR, SVM, RF, BERT\}$

Output: Optimal model $m^*$for sentiment classification

Process:
Start
Preprocessing:
For each $x_i \in D$: convert to lowercase, remove noise, and lemmatize

$$x_i' = f_{prep}(x_i)$$

Form cleaned corpus

$$D' = \{x_i'\}$$

Feature Extraction:
If $m \in \{LR, SVM, RF\}$: compute TF-IDF

$$TFIDF(t, d) = TF(t, d) \times \log \frac{N}{1 + n_t}$$

Else if $m = BERT$: generate contextual embeddings

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

Training and Evaluation:

For each $m \in M$: train and evaluate

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Model Selection:

Select best model

$$m^* = \arg\max_{m \in M}(\alpha \cdot Accuracy_m + (1 - \alpha) \cdot F1_m)$$

where $\alpha = 0.5$

Return m^*

End

Results

The dataset analyzed in this study comprised 5,842 financial text entries that were categorized into three sentiment classes: positive, neutral, and negative. Each entry represents a financial-related statement extracted from various sources such as corporate announcements, stock market news, and digital investment reports. As shown in figure 2, the neutral sentiment category was the most dominant, accounting for approximately 53 percent of the total dataset. Positive sentiments represented about 32 percent, while negative sentiments made up the remaining 15 percent. This uneven distribution indicates that financial communication in the digital market context is largely factual and report-oriented, with less emphasis on emotional expression. The prevalence of neutral statements reflects the professional and objective tone that typically characterizes financial narratives, which aim to convey information rather than evoke emotion.

A closer examination of the dataset suggests that the dominance of neutral sentiment is consistent with the linguistic style used in financial and economic reporting. Most financial texts are designed to inform rather than persuade, leading to a predominance of neutral phrasing that describes market events, company performance, and economic indicators without evaluative bias. For instance, news articles often present objective summaries such as “the company reported quarterly earnings” or “the stock price remained stable,” which do not explicitly imply positive or negative judgments. The smaller proportion of positive and negative texts may represent moments of significant market movement, such as profit growth or financial loss announcements. This distribution supports the assumption that sentiment in financial language is subtle and context-dependent, requiring advanced modeling techniques to accurately interpret underlying tones.

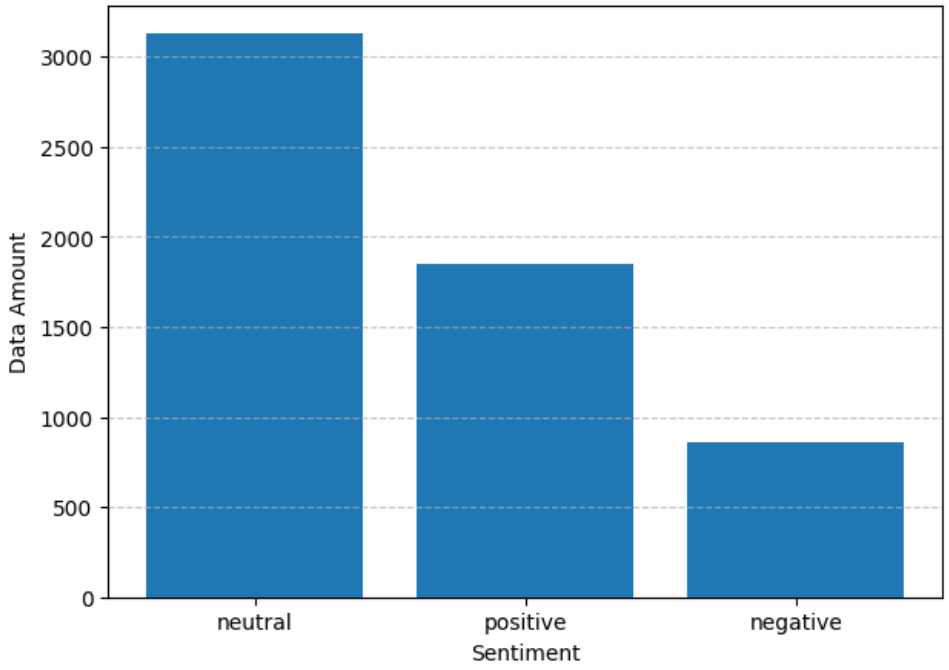


Figure 2 Distribution of Sentiments in the Financial Text Dataset

The analysis of sentence length distribution, as shown in figure 3, reveals that the majority of financial text entries contain between 10 and 25 words. This finding suggests that most sentences are relatively short, concise, and straightforward in structure. The tendency toward brevity reflects the nature of financial communication, where clarity and efficiency are highly valued. Financial reports, press releases, and market updates often prioritize factual accuracy and directness to ensure that information is easily understood by investors, analysts, and stakeholders. Longer sentences are less common, as they can obscure meaning or introduce ambiguity, which is undesirable in financial contexts where precision is crucial.

Further examination of sentence structure indicates that the concise form of expression used in the dataset aligns with the stylistic conventions of digital financial communication. Many of the sentences resemble headline-style reporting or summary statements designed for rapid consumption on trading platforms and financial news feeds. This writing style mirrors how market information is shared in real-time environments, where short and impactful statements can quickly convey essential insights. The dominance of mid-length sentences also suggests a balance between informational richness and readability, enabling the text to include relevant financial data or commentary without becoming overly complex. Overall, this characteristic of the dataset reinforces its suitability for natural language processing applications, as shorter, well-structured sentences are easier to tokenize and analyze computationally.

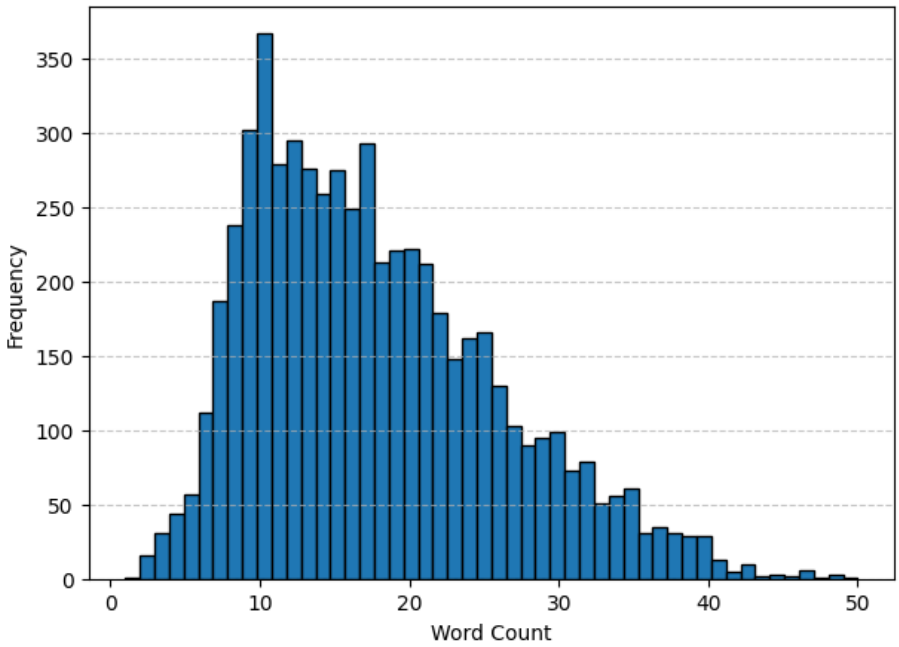


Figure 3 Distribution of Sentence Lengths (Number of Words per Sentence)

The WordCloud visualization for the positive sentiment category, presented in figure 4, provides a clear overview of the most frequently used words associated with optimistic financial conditions. The terms profit, growth, increase, and strong appear prominently, indicating that many texts in this category discuss favorable business performance, market expansion, and financial stability. These words are typically found in corporate reports announcing revenue gains, investment achievements, or improved market positions. The frequent occurrence of such vocabulary suggests that positive sentiment texts in the dataset often describe progress, confidence, and upward momentum in the financial environment. This pattern aligns with the linguistic tendencies of optimistic reporting, where success-oriented terms are used to highlight company achievements and market strength.

The negative sentiment WordCloud, illustrated in figure 5, reveals the language patterns that characterize financially pessimistic or adverse market conditions. Common terms include loss, decline, decrease, risk, and crisis, all of which are typically associated with downturns, reduced profits, or uncertainty in the market. The recurrence of these words reflects scenarios such as falling stock prices, missed earnings targets, or economic slowdowns. The vocabulary used in negative sentiment texts tends to convey caution or concern, emphasizing the negative consequences of market performance or external factors affecting financial stability. The presence of such terms indicates that negative sentiment entries are often event-driven, focusing on unexpected market disruptions or unfavorable business outcomes.

The WordCloud for the neutral sentiment category, shown in figure 6, highlights words such as report, market, share, company, and financial. These words suggest a more factual and descriptive tone, commonly found in texts that convey information without expressing an emotional stance. Neutral sentiment texts often describe market events, summarize financial reports, or present statistical data in an objective manner. For example, phrases such as “the company released its quarterly report” or “the market recorded moderate trading volume” are typical of this category. The prominence of neutral language reflects the communicative purpose of financial writing, which aims to inform readers rather than persuade them. This observation supports the earlier finding that financial discourse in digital markets is dominated by informational statements, indicating the analytical and report-driven nature of this domain.

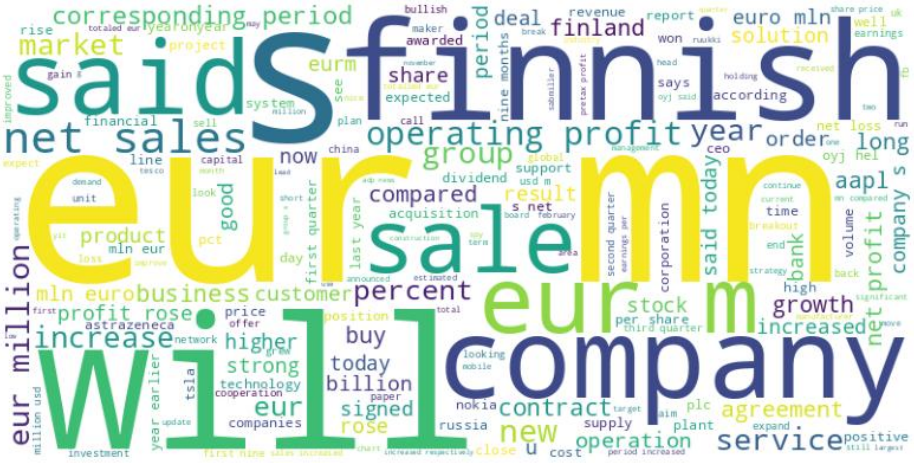


Figure 4 WordCloud for Positive Sentiment Texts



Figure 5 WordCloud for Negative Sentiment Texts



Figure 6 **WordCloud for Neutral Sentiment Texts**

To establish a baseline for performance comparison, three classical machine learning models were implemented: Logistic Regression, Support Vector Machine (SVM), and Random Forest. These models were trained using Term Frequency–Inverse Document Frequency (TF-IDF) features, which transform textual data into numerical representations based on the importance of words within the corpus. As shown in table 1, Logistic Regression achieved the best overall performance with an accuracy of 71.6 percent, followed by SVM with 69.4 percent and Random Forest with 63.6 percent. The relatively strong performance of Logistic Regression and SVM suggests that linear classifiers can effectively identify general sentiment patterns in financial texts, particularly when sentiment cues are explicitly expressed through keywords such as profit or loss. However, the comparatively lower accuracy of the Random Forest model indicates that tree-based ensemble approaches may not perform as efficiently on sparse, high-dimensional data such as TF-IDF vectors, where most features are independent and context is limited.

Although the traditional models achieved satisfactory results, their overall performance demonstrates key limitations in understanding financial language. These algorithms rely primarily on surface-level lexical features and are unable to capture contextual or semantic relationships between words. Financial statements often contain complex phrasing, where sentiment depends on subtle linguistic patterns, contextual cues, or multi-word expressions. For example, a sentence such as “profits declined at a slower rate than expected” conveys a mildly positive outlook despite containing the negative word declined. Classical models struggle with such nuanced interpretation because they treat each word independently. This limitation highlights the need for more advanced approaches, such as Transformer-based models, which can capture deeper semantic relationships and contextual dependencies critical for accurate sentiment classification in financial text analysis.

Table 1 Performance Comparison of Baseline Machine Learning Models

Model	Accuracy
Logistic Regression	0.7159
Support Vector Machine (SVM)	0.6938
Random Forest	0.6356

A Transformer-based deep learning model, namely the Bidirectional Encoder Representations from Transformers (BERT), was fine-tuned to enhance the model's ability to understand the contextual meaning of financial language. Unlike traditional models that rely solely on word frequency and co-occurrence, BERT utilizes a bidirectional attention mechanism that allows it to consider the full context of each word within a sentence. This characteristic enables the model to interpret complex financial statements where sentiment depends on subtle linguistic cues and phrase-level relationships. After fine-tuning on the financial text dataset, the BERT model demonstrated superior performance, achieving an overall accuracy of 78.8 percent, which represents a notable improvement over all baseline models. The results indicate that the contextual embedding approach of BERT effectively captures the intricate sentiment structures present in financial documents, making it more capable of understanding the meaning behind phrases such as "steady growth despite market volatility" or "slight improvement following previous losses."

The detailed classification metrics presented in table 2 further illustrate the model's strength across sentiment categories. BERT achieved F1-scores of 0.85 for positive sentiment, 0.83 for neutral sentiment, and 0.54 for negative sentiment. The high performance in positive and neutral classes suggests that the model can accurately identify expressions of confidence, stability, and neutrality in financial language. However, the comparatively lower score for negative sentiment highlights the effect of class imbalance, as the dataset contains fewer examples of explicitly negative financial statements. Negative sentiments in financial text are often expressed indirectly or through nuanced wording, making them more challenging to detect even for context-aware models. Despite this limitation, BERT's overall results confirm its strong ability to generalize sentiment patterns and its suitability for analyzing real-world financial communications, where contextual understanding is essential for accurate interpretation.

Table 2 Classification Metrics of the Fine-Tuned BERT Model

Sentiment	Precision	Recall	F1-score
Negative	0.49	0.60	0.54
Neutral	0.86	0.80	0.83
Positive	0.84	0.86	0.85
Overall Accuracy			0.79

The confusion matrix presented in figure 7 provides a detailed overview of the classification performance of the BERT model across all sentiment categories. The model demonstrated a high level of accuracy in identifying both neutral and positive sentiments, which together accounted for the majority of correctly classified samples. This outcome suggests that BERT effectively captures the contextual patterns associated with these sentiment types, such as expressions of stability, confidence, and positive financial outlooks. In neutral cases, the model performed particularly well because such texts often contain objective and factual descriptions that align closely with its contextual embedding structure. For positive statements, the use of clearly optimistic language, such as increase in profits or strong quarterly performance, made it easier for the model to distinguish them from other categories. The ability to differentiate these sentiments accurately indicates that BERT successfully learned the linguistic cues that define financial optimism and neutrality.

However, the confusion matrix also reveals that some misclassifications occurred between neutral and negative sentiments. This overlap can be attributed to the subtle and ambiguous nature of financial language, where certain terms or phrases can convey both cautious and negative undertones depending on context. Examples include expressions such as “slower growth,” “weaker demand,” or “moderate decline,” which may represent either a neutral report of facts or a mild indication of concern. These linguistic ambiguities often arise in financial discourse, where authors aim to describe unfavorable events without using strongly negative wording. As a result, the BERT model occasionally predicted neutral statements as negative or vice versa. This finding highlights the complexity of interpreting sentiment in financial communication, where tone is often influenced by context, comparison, and implication rather than explicit emotional language.

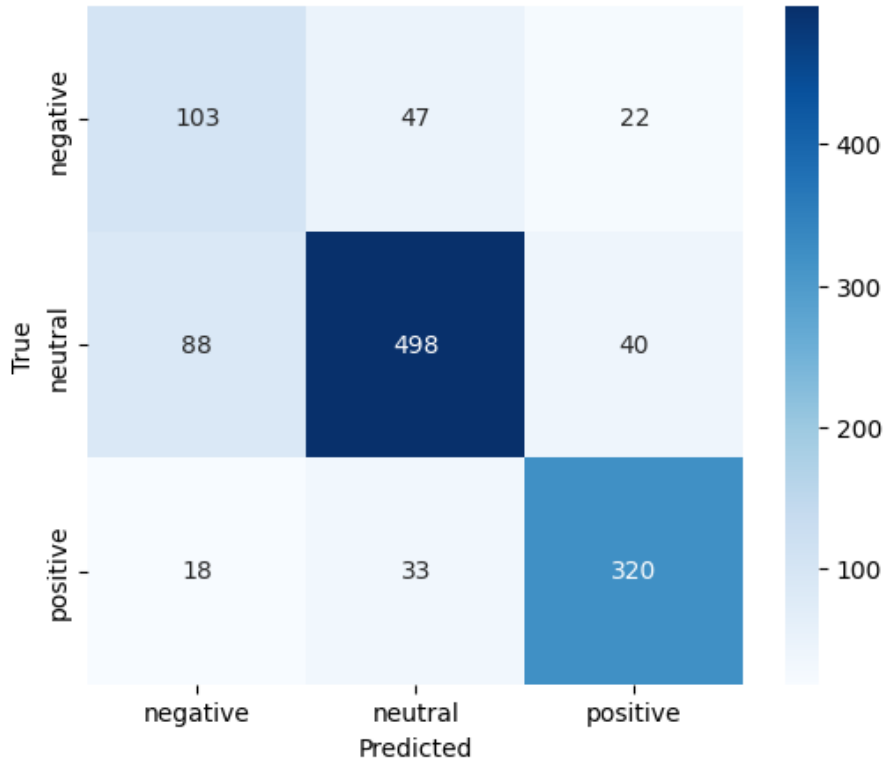


Figure 7 Confusion Matrix of the Fine-Tuned BERT Transformer Model.

The comparative visualization in figure 8 clearly demonstrates the superiority of the BERT model over the traditional machine learning approaches used as baselines. The Logistic Regression model achieved an accuracy of 0.7159, the Support Vector Machine (SVM) reached 0.6938, and the Random Forest obtained 0.6356. In contrast, the fine-tuned BERT model achieved an accuracy of 0.7879, representing an improvement of approximately seven percent over the best-performing baseline. This consistent improvement across sentiment categories reflects BERT's ability to capture context-dependent meaning in financial text, where the interpretation of sentiment often depends on relationships between multiple words rather than single tokens. For example, phrases such as "limited recovery" or "steady improvement" require an understanding of nuance and comparison that simple word frequency models cannot achieve. The visual comparison also emphasizes the greater stability of BERT's performance, suggesting that Transformer-based models generalize more effectively to varied sentence structures and vocabulary present in real-world financial discourse.

The improvement in accuracy shown in figure 8 highlights the key advantage of contextualized embeddings within Transformer architectures. Unlike traditional bag-of-words models, which treat each word independently, BERT analyzes words in relation to their surrounding context, allowing it to learn semantic dependencies and subtle tone variations. This capability is particularly important in financial sentiment

analysis, where a single term can shift meaning based on modifiers or sentence structure. For instance, the term “growth” can indicate positive performance when paired with “strong,” but may convey a negative connotation when preceded by “slower.” The use of contextual embeddings enables the model to distinguish these differences and interpret sentiment more accurately. Overall, the results presented in figure 8 confirm that Transformer-based deep learning models are more capable of understanding complex linguistic patterns, making them far more effective for financial text analysis and digital market prediction than traditional machine learning methods.

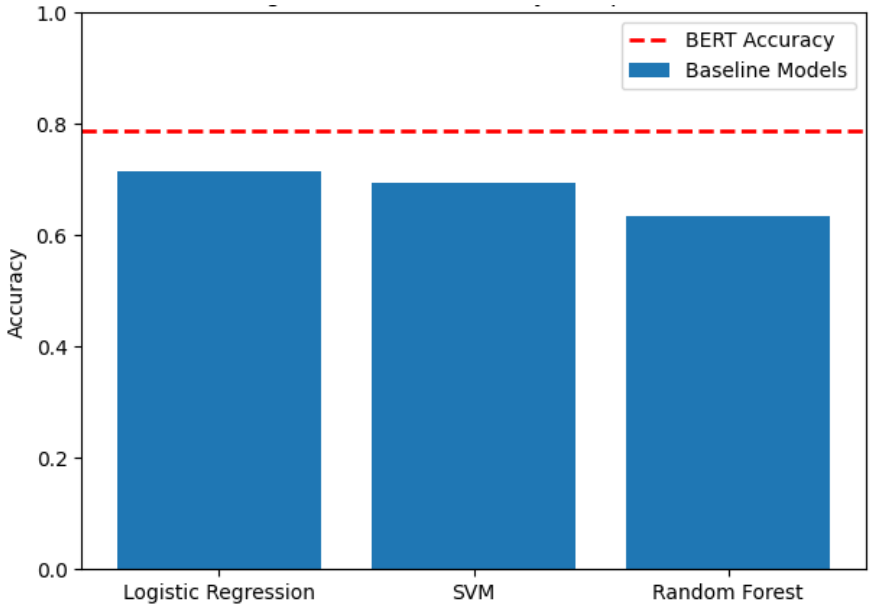


Figure 8 Model Accuracy Comparison between Baseline Machine Learning Models and BERT.

Overall, the experimental findings clearly demonstrate that Transformer-based models, particularly BERT, provide a substantial improvement over traditional machine learning algorithms in financial sentiment classification. The ability of BERT to process text bidirectionally allows it to capture the full contextual meaning of words, enabling the model to detect subtle linguistic cues, implied sentiment, and tone variations that conventional models often overlook. This contextual understanding is essential in financial language, where the sentiment of a statement may depend on the interplay between multiple terms or the overall phrasing rather than on specific keywords. The results indicate that BERT successfully identifies both explicit and implicit expressions of market confidence, caution, or instability, leading to more accurate sentiment predictions across different financial communication formats. The strong performance of the Transformer-based approach demonstrates its potential as a powerful analytical tool for digital market sentiment prediction, automated financial monitoring, and decision-support systems powered by artificial intelligence. Such models can contribute significantly

to the advancement of AI-driven financial analytics by providing more precise insights into market behavior and improving the interpretability of large-scale financial text data.

Discussion

The findings of this research highlight the effectiveness of Transformer-based architectures, particularly BERT, in addressing the linguistic and contextual complexities of financial text sentiment analysis [21]. The model's superior performance compared to classical algorithms such as Logistic Regression, SVM, and Random Forest underscores the importance of context-aware modeling in financial communication, where sentiment is often conveyed through nuanced phrasing rather than explicit emotional language [22]. Financial statements and market reports tend to use carefully constructed language that avoids extremes, relying instead on cautious expressions or indirect indicators of sentiment. The ability of BERT to process words in both forward and backward directions allow it to understand how meaning shifts depending on surrounding terms , enabling the model to interpret sentiment accurately even in ambiguous or mixed-tone statements. This capability is particularly valuable for financial text, where subtle linguistic cues such as comparative adjectives, hedging expressions, or risk-related terms play a crucial role in determining the overall sentiment [23]. The results demonstrate that BERT's contextual embeddings capture these intricacies more effectively than traditional word-based models, making it a suitable approach for the domain of financial communication and digital market analytics [24].

Beyond the improvement in accuracy and F1-scores, the results of this study also have broader implications for the implementation of artificial intelligence in financial decision-making systems [25]. The Transformer-based approach enables automated, real-time monitoring of digital market sentiment, offering deeper insight into investor behavior, market confidence, and potential volatility. By analyzing large volumes of textual data such as financial news, social media updates, and corporate disclosures, models like BERT can support data-driven strategies for traders, analysts, and policy developers. The confusion between neutral and negative categories observed in the results also provides an important linguistic insight: financial communication often balances between factual reporting and cautious commentary, making sentiment classification inherently complex. Despite this challenge, BERT's contextual understanding demonstrates strong adaptability to the indirect and professional tone of financial writing. This reinforces its potential to enhance sentiment-driven market prediction tools, improve the interpretability of automated financial analyses, and contribute to the ongoing development of intelligent, AI-assisted digital finance systems.

Conclusion

This study concludes that Transformer-based deep learning models, particularly BERT, represent a significant advancement in financial sentiment analysis by effectively addressing the contextual and semantic limitations of traditional machine learning approaches. The fine-tuned BERT model demonstrated superior performance with higher accuracy and F1-scores across sentiment categories, confirming its ability to understand the complex structure and meaning of financial language. Financial communication often employs cautious, balanced, and data-driven wording, where sentiment depends on subtle contextual relationships rather than explicit emotional expressions. BERT's bidirectional attention mechanism enables it to interpret such linguistic nuances by analyzing how words influence one another within a sentence, allowing for a more accurate classification of positive, negative, and neutral sentiments. The model's contextual awareness also makes it more adaptable to the ambiguity and professionalism inherent in financial discourse. The findings suggest that incorporating Transformer-based models into financial analytics can enhance digital market prediction, improve investor sentiment monitoring, and support data-driven decision-making in automated trading and investment systems. Furthermore, this research highlights the growing potential of artificial intelligence in transforming digital financial ecosystems by enabling large-scale, real-time text analysis that complements traditional quantitative approaches, thereby contributing to more intelligent, responsive, and transparent market evaluation tools.

Declarations

Author Contributions

Conceptualization: M.A.L.J.; Methodology: M.A.L.J.; Software: M.A.L.J.; Validation: M.A.L.J.; Formal Analysis: M.A.L.J.; Investigation: M.A.L.J.; Resources: M.A.L.J.; Data Curation: M.A.L.J.; Writing Original Draft Preparation: M.A.L.J.; Writing Review and Editing: M.A.L.J.; Visualization: M.A.L.J.; All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

The data presented in this study are available on request from the corresponding author.

Funding

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

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] D. Araci, "FinBERT: Financial sentiment analysis with pre-trained language models," *arXiv preprint arXiv*:1908.10063, 2019, doi: 10.48550/arXiv.1908.10063.

[2] Y. Yang, M. C. S. Uy, and A. Huang, "FinBERT: A pretrained language model for financial communications," *arXiv preprint arXiv:*2006.08097, 2020, doi: 10.48550/arXiv.2006.08097.

[3] Z. Liu, D. Huang, K. Huang, Z. Li, and J. Zhao, "FinBERT: A pre-trained financial language representation model for financial text mining," *in Proc. 29th International Joint Conference on Artificial Intelligence (IJCAI-20)*, 2020, pp. 4513–4519, doi: 10.24963/ijcai.2020/622.

[4] P. Malo, A. Sinha, P. Korhonen, J. Wallenius, and P. Takala, "Good debt or bad debt: Detecting semantic orientations in economic texts," *Journal of the Association for Information Science and Technology*, vol. 65, no. 4, pp. 782–796, 2014, doi: 10.1002/asi.23062.

[5] C. Kearney and S. Liu, "Textual sentiment in finance: A survey of methods and models," International Review of Financial Analysis, vol. 33, pp. 171–185, 2014, doi: 10.1016/j.irfa.2014.02.006.

[6] A. K. Nassirtoussi, S. Aghabozorgi, T. Y. Wah, and D. C. L. Ngo, "Text mining for market prediction: A systematic review," Expert Systems with Applications, vol. 41, no. 16, pp. 7653–7670, 2014, doi: 10.1016/j.eswa.2014.06.009.

[7] T. Loughran and B. McDonald, "Textual analysis in finance," Annual Review of Financial Economics, vol. 12, pp. 357–375, 2020, doi: 10.1146/annurev-financial-012820-032249.

[8] T. Loughran and B. McDonald, "When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks," The Journal of Finance, vol. 66, no. 1, pp. 35–65, 2011, doi: 10.1111/j.1540-6261.2010.01625.x.

[9] P. C. Tetlock, "Giving content to investor sentiment: The role of media in the stock market," The Journal of Finance, vol. 62, no. 3, pp. 1139–1168, 2007, doi: 10.1111/j.1540-6261.2007.01232.x.

[10] J. Bollen, H. Mao, and X. Zeng, "Twitter mood predicts the stock market," Journal of Computational Science, vol. 2, no. 1, pp. 1–8, 2011, doi: 10.1016/j.jocs.2010.12.007.

[11] N. Oliveira, P. Cortez, and N. Areal, "The impact of microblogging data for stock market prediction: Using Twitter to predict returns, volatility, trading volume and survey sentiment indices," Expert Systems with Applications, vol. 73, pp. 125–144, 2017, doi: 10.1016/j.eswa.2016.12.036.

[12] X. Ding, Y. Zhang, T. Liu, and J. Duan, "Using structured events to predict stock price movement: An empirical investigation," in Proc. 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, 2014, pp. 1415–1425, doi: 10.3115/v1/D14-1148.

[13] X. Ding, Y. Zhang, T. Liu, and J. Duan, "Deep learning for event-driven stock

prediction," in Proc. 24th International Joint Conference on Artificial Intelligence (IJCAI), 2015, pp. 2327–2332. [Online]. Available: https://ijcai.org/proceedings/2015/329.

[14] M. R. Vargas, B. S. L. P. de Lima, and A. G. Evsukoff, "Deep learning for stock market prediction from financial news articles," in Proc. 2017 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA), 2017, pp. 60–65, doi: 10.1109/CIVEMSA.2017.7995302.

[15] Q. Xiao and B. Ihnaini, "Stock trend prediction using sentiment analysis," PeerJ Computer Science, vol. 9, Art. no. e1293, 2023, doi: 10.7717/peerj-cs.1293.

[16] R. Pan, J. A. García-Díaz, F. García-Sánchez, and R. Valencia-García, "Evaluation of transformer models for financial targeted sentiment analysis in Spanish," PeerJ Computer Science, vol. 9, Art. no. e1377, 2023, doi: 10.7717/peerj-cs.1377.

[17] G. Fatouros, J. Soldatos, K. Kouroumali, G. Makridis, and D. Kyriazis, "Transforming sentiment analysis in the financial domain with ChatGPT," Machine Learning with Applications, vol. 14, Art. no. 100508, 2023, doi: 10.1016/j.mlwa.2023.100508.

[18] H.-M. Chuang, H.-C. He, and M.-C. Hu, "Chinese financial news analysis for sentiment and stock prediction: A comparative framework with language models," Big Data and Cognitive Computing, vol. 9, no. 10, Art. no. 263, 2025, doi: 10.3390/bdcc9100263.

[19] W. Gu, Y. Zhong, S. Li, C. Wei, L. Dong, Z. Wang, and C. Yan, "Predicting stock prices with FinBERT-LSTM: Integrating news sentiment analysis," in Proc. 8th International Conference on Cloud and Big Data Computing (ICCBDC 2024), Oxford, U.K., 2024, pp. 67–72, doi: 10.1145/3694860.3694870.

[20] K. Du, F. Xing, R. Mao, and E. Cambria, "Financial sentiment analysis: Techniques and applications," ACM Computing Surveys, vol. 56, no. 9, Art. no. 220, pp. 1–42, 2024, doi: 10.1145/3649451.

[21] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in Proc. 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, vol. 1, 2019, pp. 4171–4186, doi: 10.18653/v1/N19-1423.

[22] D. Araci, "FinBERT: Financial sentiment analysis with pre-trained language models," arXiv preprint arXiv:1908.10063, 2019, doi: 10.48550/arXiv.1908.10063.

[23] Z. Liu, D. Huang, K. Huang, Z. Li, and J. Zhao, "FinBERT: A pre-trained financial language representation model for financial text mining," in Proc. 29th International Joint Conference on Artificial Intelligence (IJCAI-20), Special Track on AI in FinTech, 2020, pp. 4513–4519, doi: 10.24963/ijcai.2020/622.

[24] E. Cicekyurt and G. Bakal, "Enhancing sentiment analysis in stock market tweets through BERT-based knowledge transfer," Computational Economics, vol. 67, pp. 855–877, 2026, doi: 10.1007/s10614-025-10901-8.

[25] K. Du, F. Xing, R. Mao, and E. Cambria, "Financial sentiment analysis: Techniques and applications," ACM Computing Surveys, vol. 56, no. 9, Art. no. 220, pp. 1–42, 2024, doi: 10.1145/3649451.