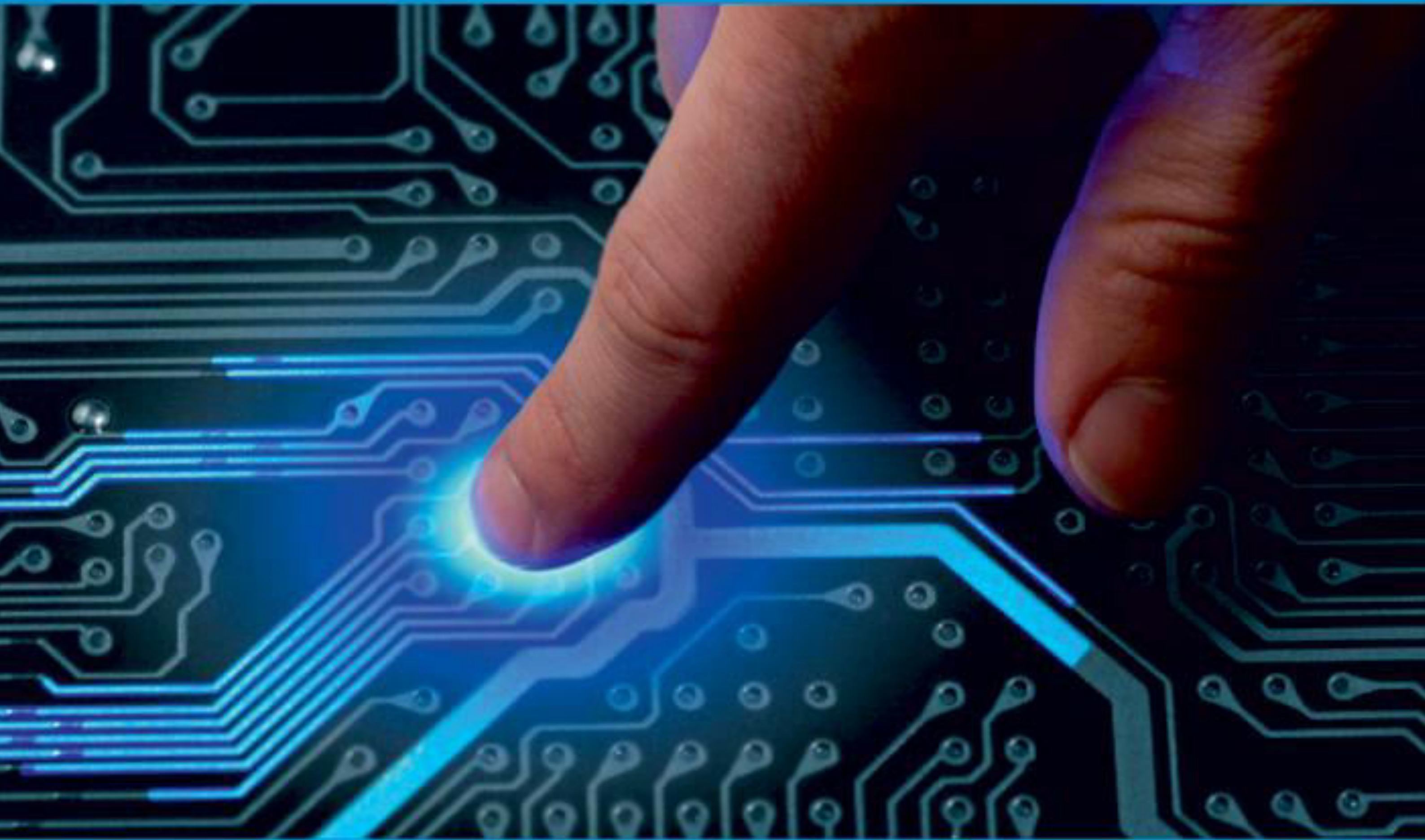
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 11, November 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.625

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Sentimental Effects of World Stock Market Investors During Bear Market

Dhiraj Mudvari

Department of Management, Dr. K.N.Modi University, Newai, Rajasthan, India

ABSTRACT: The sentimental impact of international stock market investors during a bear market has been assessed, focusing on the way emotions like fear, panic, and optimism influence market conditions. Since traditional financial analysis is mostly numbers-based, the addition of sentiment analysis using advanced NLP methods, such as BERT (Bidirectional Encoder Representations from Transformers), provides a greater understanding of investor behavior. BERT's two-way contextualization allows sentiment classification by reading financial texts, such as news and social media, to detect the psychological drivers of market movements. The proposed framework is fine-tuning BERT on a labeled financial dataset with the goal of classifying sentiments into positive, negative, or neutral categories, focusing on emotions that drive market volatility. The results indicated that fear and panic in bear markets tend to cause more volatility and colossal sell-offs, whereas optimism marks the beginning of some recovery or induces selective buying. This research work identifies how investor sentiment is not merely a reflection of the market's status but also drives the market into certain behaviors, hence the notion of feedback loops. Much value was offered to analysts, traders, and policymakers alike with its data-driven approach toward comprehension and response to investor sentiment regarding financial downturns.

KEYWORDS: Bear Market, Bidirectional Encoder Representations from Transformers, Investors, Sentimental analysis, Stock Market

I. INTRODUCTION

The behavior of stock market investors is significantly influenced by the sentiments they feel about that market, especially in conditions where financial distress has a huge impact, like a bear market [1]. Sentiments, particularly fear, panic, and optimism, significantly dictate what happens in the markets as this often amplifies volatility and sets large price movements [2]. During a bear market, sustained stock price declines alongside widespread investor pessimism characterize a setting for examining the forces that determine the psyche of this behavior [3]. Knowledge about the psychological impacts of these investors should be established in developing risk reduction strategies to ensure better decision making. Investor sentiment is a complex phenomenon which depends on various factors including economic conditions, performance of corporations, geopolitical events, and market rumors [4]. Negative sentiments such as fear and panic dominate the bear markets resulting in herd behavior, major sell-offs, and dramatic falls in the indices of the market [5]. On the flip side, optimism, when even slight and transitory, could propel temporary recoveries or trigger selective buying activity that fosters periods of relative calm within the storm. Emotional responses to any given set of market happenings not only represent, but also constitute, huge driving forces behind market activity and can feed back, further reinforcing volatility. NLP techniques developed have helped researchers quantify investor sentiment based on textual information from news, analyst reports, and social media sites [6].

One of the highly effective tools developed for use in this area is called Bidirectional Encoder Representations from Transformers, or BERT. Unlike traditional models, BERT captures word contexts in a sentence by analyzing their words preceding and succeeding to each word [7]. This makes BERT more efficient in identifying the nuanced language used in financial texts [8]. With fine-tuning on domain-specific datasets like financial news or stock market-related tweets, it can achieve an accuracy of high classification of investor sentiments into positive, negative, or neutral. The current study uses BERT-based NLP techniques to analyze and classify investor emotions regarding sentimental effects on world stock market investors during bear markets. This is in an effort to study the extent to which sentiments such as fear, panic, and optimism can influence market behavior in times of downturns. This research aims to discover the extent to which investor emotions drive market dynamics through leverage and use of a sentiment analysis dataset and BERT's capabilities in contextual understanding. It thus will be important to improve strategies for better response to better

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

psychology concerning bear markets without allowing an amplification of risks followed by stabilization of markets. The major key contributions of the research are,

- This study utilizes BERT for sentiment analysis to classify and analyze investor emotions during bear markets.
- The methodology integrates financial datasets, including news articles and social media, to capture real-time sentiment trends, offering insights into how emotions like fear, panic, and optimism influence market volatility.
- By fine-tuning BERT for financial sentiment analysis, the study improves the accuracy and contextual understanding of sentiment classification compared to traditional models.
- The research highlights the role of sentiment-driven feedback loops in bear markets, revealing how investor emotions not only reflect market conditions but also actively shape stock market dynamics and volatility.

The remaining sections are aligned as follows: related work in section II, section III into the methodology for sentimental analysis. The results in section IV model's performance are given, along with a comparison to prove its effectiveness. The conclusion in section V presents the outcome of the experiments.

II. RELATED WORKS

Kumar [9] focus on the sentiment analysis for stock price prediction with the data taken from Twitter. Tweets collected are pre-processed and then fed into BERT and Word2Vec models in order to train the sentiment models used in the study. Twitter sentiment analysis will help the investors to avoid the certain level of stock market investments' risk in case if they are going to make some decisions concerning stock market share. This article uses the dataset, obtained mainly through the Twitter API, to train BERT model for stock prediction. This study continues the line of research using social media data and NLP to predict stock market trends and provides investors, including newcomers, with a practical tool for more cautious investment decisions based on public sentiment. Fatouros [10] aims to understand the capabilities of such a LLM, more specifically ChatGPT 3.5, when applied to the foreign exchange financial sentiment analysis. The authors use a zero-shot prompting method utilizing forex-related news headlines and then measure ChatGPT's effectiveness against FinBERT which is an extant model for analyzing financial sentiment. The outcome reveals that ChatGPT delivers about 35% higher accuracy than FinBERT on sentiment classification and is about 36% more efficient in terms of correlation with stock market returns. This study establishes the importance of timely engineering particularly in the case of zero-shot application and explains how newer models like the ChatGPT could offer tremendous improvement to the financial sentiment analysis.

Yadav [11] proposes a BERT-based sentiment analysis model that is combined with supervised machine learning algorithms to derive improved results for stock trend prediction. This study proposes extensive model of stock market predictions by incorporating sentiment features from financial news articles and the conventional financial indicators. The model collects net sentiment data on a daily basis and shows that they are highly statistically significant in forecasting stock exchange movement. This study advances in this research on stock prediction, as it demonstrates the importance of sentiment data, and how to apply it in conjunction with the more established financial data in dynamic market. Costola [12] aims to determine how the COVID-19 news affects the expected value of stocks through studying the early part of the pandemic when investors received a large inflow of new and diverse information. The authors use a financial market-adapted BERT model for sentiment extraction on 203,886 of articles from the MarketWatch, NYTimes, and Reuters. In the final analysis, this study found a statistically significant correlation between sentiment scores and the S&P 500 index. Furthermore, the study finds that sentiment components and news categories of NYTimes co-move differently with market returns, implying the need to analyze sentiment for market expectation.

Hasan [13] explores the impact of sentiment on stock and cryptocurrency daily price movements, this paper presents an approach to building a computerized sentiment analysis system for aggregating web-based news. When the authors use BERT together with Natural Language Toolkit (NLTK), they are able to get a high performance of 95.84% for text summarization. This study is focus on the impact of sentiment derived from social networks and news portals on market conditions in the mostly unstable stocks and cryptocurrencies. Through sentiment summarization, the system assists investors to understand the market sentiment and make fast decisions especially in fast moving market.

© 2024 IJIRCCE | Volume 12, Issue 11, November 2024| DOI: 10.15680/IJIRCCE.2024.1211121

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

The above studies highlighted the increased importance of using sentiment analyses for stock market prediction, including BERT, Chat GPT and FinBERT. Twitter messages, news articles and financial texts are used to determine investors' perceptions and the effectiveness of these in explaining market action are shown. These models improve stock forecast when sentiment analysis is combined with the basic financial data thus enabling investors to make wiser choices. The existing approach highlights the significance of early-engineering along with text summarization procedures for enhancing the efficiency of sentiment analysis, within the stocks and cryptocurrency domains.

III. PROPOSED BERT MODEL FOR SENTIMENT ANALYSIS

The study proposes the advanced NLP methods, particularly BERT, which performs Bidirectional Encoder Representations from Transformers, in capturing investor sentiment and its impact in the dynamics of the bear market. The approach begins at the collection and preprocessing stages of a financial sentiment dataset. The dataset is supposed to be cleaned by removing all of the irrelevant information for data remain of high quality. Next, BERT is fine-tuned using the pre-processed financial data. The model is trained in classifying sentiments into three categories - positive, negative, and neutral using labelled data, giving special attention to emotions that include fear, panic, and optimism. The model captures the context of the words in discussion about finances. The output from the model is analyzed to examine, investor sentiment's influence is sought on market behaviour, with particular attention given to sentiment-driven feedback loops that amplify market volatility during bear markets. The overall workflow of a proposed model is depicted in fig.1.

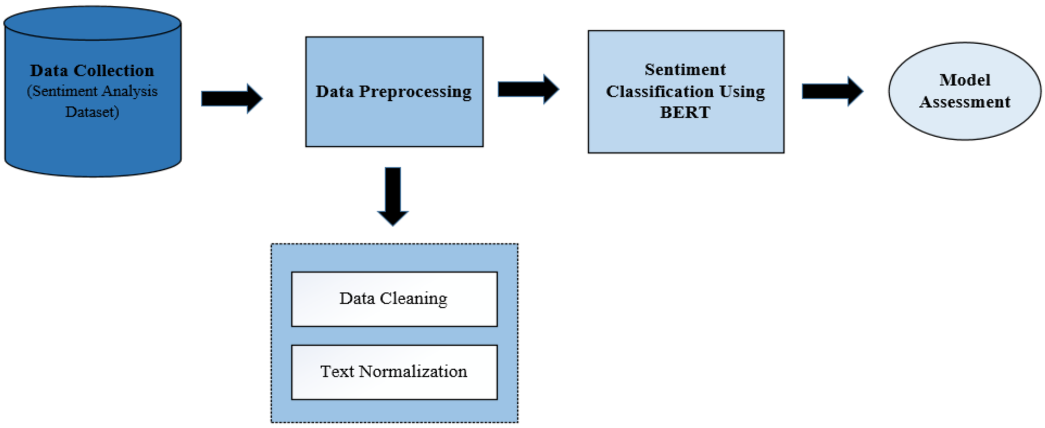


FIG 1 Workflow of the proposed model for sentimental analysis

A. Data collection

The Sentiment Analysis Dataset is collected from Kaggle [14]. This is a dataset with sentiment polarity for tweets, where the sentiment assigned to any particular tweet includes 0 – which is negative, 2 – which is neutral, and 4 which is positive. The data was then categorized automatically on the basis of emoticons, positive emoticons for positive polarity and negative emoticons for negative polarity. Besides the sentiment label, the tweets and their metadata are included in the dataset in terms of user's age, the country of the user, the population, and the land area of the country, but the emoticons in the tweets are deleted.

B. Data Pre-processing

1. Data Cleaning

Data cleaning is a step removes columns that are not needed such as tweet ID, query and user handle but keeps essential fields - text, sentiment, and the time the tweet was posted. Text cleaning remove special characters, punctuation, etc., or any non-alphanumeric symbol (@, #, &, etc.) that doesn't add to the sentiment. URLs, which are pretty popular in tweets, are removed using regular expressions in order to remove any pattern that looks like a web address.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

2. Text Normalization

Text normalization has several major stages to prepare the data analysis. First, all text is converted to lowercase to remove extra features that are not necessary to simplify the text for the feed forward neural network. Subsequently, tokenization is used to divide the tweet text into smaller parts of words called tokens for further analysis. Some of the groups of data processed before analysis include Stop words such as, the, and, it, and is among others since such words are meaningless when determining text sentiment. At last, lemmatization is applied to convert the all the different forms of a word to one form, for instance, replacing the word "running" with "run" to ensure that the various form of a word is treated as one during analysis. These steps put together leash the text into one normal set of standards of analysis and improves the quality of sentiment assessment.

C. Sentiment Classification with BERT

BERT stand for Bidirectional Encoder Representations from Transformers [15] is a cutting-edge language model for natural language understanding (NLU) tasks that helps the machine to understand phrases in right contextual sense. It uses transfer learning in a way that they first pre-train high-level model and then leverage the pre-trained model for downstream related tasks, like sentiment classification. It helps machines in understanding the contextual of words by so employing transfer learning – making use of a pre-trained model and then applying it to a different task, such as the sentiment analysis of investors in bear market. BERT model is built on transformer architecture where understand textual data by capturing both the left and right contexts of a word can be obtained. BERT while being pre-trained on two basic tasks that inevitably introduce it to context-level understanding of the language. The architecture of BERT is given in fig.2. The first of these tasks is called Masked Language Modeling (MLM) where complete sentences are entered to the model while certain words are masked. After the input data is preprocessed, BERT is subsequently employed on the prediction of such masked words given the context of the other words in the particular sentence. This has the advantage of making BERT understand how words relate to each other within the natural language since it learns the relations at the words level. The second problem, Next Sentence Prediction (NSP), is in fact aimed at training BERT in the relationships between the sentences. In this task, one gets two sentences at a time, and the BERT model will consider one of them and figure out if the next one makes sense or not. This capability is particularly important for the model to understand the relationships between the different sentences of a text, and is very useful for tasks that need to identify sentences relationships. In order to evaluate the sentiments of investors during bear trends, BERT needs to be fine-tuned on a financial sentiment data. BERT is critically involved in financial sentiment analysis because it is capable of processing textual information while recognizing the effects individual words in different settings. It starts with the cordial separation of the input sentences into tokens to create smaller sub-languages. Each of these tokens is then sent through BERT's Transformer encoder that encodes the entire sentence at once. This processing allows the model to capture the context of every word through the use of previous and next words in the sentence. The result of this process is contextualized vectors for each token, measure of the meaning of the word in a certain context. This layer assesses the overall sentiment as being Positive, Negative, and Neutral based on the contextualized embeddings procured by the process. This approach allows BERT to work with investors' sentiment at scale and determine fear, panic, or optimism within the textual data. Such analysis is very useful during the periods of bear market indicators especially where feelings are known to dictate the market.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

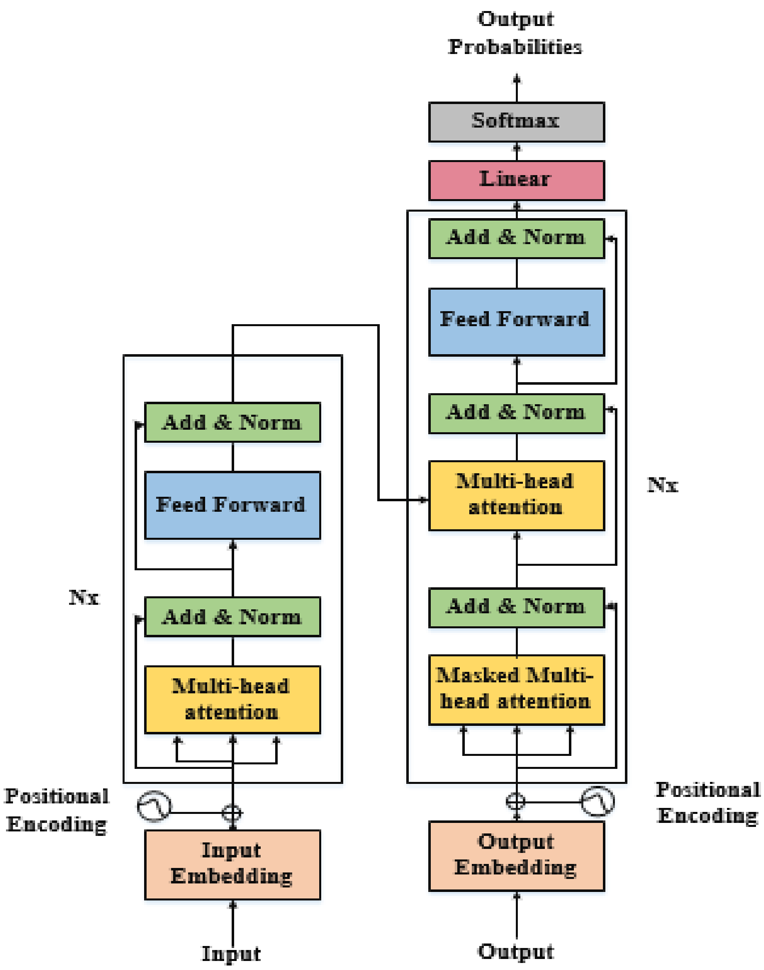


Fig.2 Architecture of BERT

Incorporating BERT's fine-grained text classification it becomes possible to monitor sentiment shifts, gauge market response, and provide actionable insights into the psychological drivers behind market behavior during downturns. One of the major strengths of BERT is supports bidirectional contextualization. BERT can use preceding as well as following words for generating the context of a word. Another important benefit of BERT is its ability to utilize transfer learning, where a pre-trained model is fine-tuned on domain-specific financial data to improve its accuracy in sentiment classification. By architecture, BERT processes an entire sentence simultaneously, thereby capturing long-range dependencies and complex relationships between words that, therefore allow for better performance than traditional models. These abilities make BERT highly efficient for analysis in the stock market, especially under bearish markets. This ability to classify investor sentiment at scale can bring insightful understanding of trends within the stock market for analysts, traders, and policymakers alike since they can better understand the movement of markets caused by the emotions of investors.

IV. RESULT AND DISCUSSION

This study demonstrates the performance of the BERT in the sentiment classification in stock market prediction is established with accuracy, precision, recall, and F1- score. Machine learning based sentiment predictor using BERT is found to perform better than the traditional machine learning based sentiment predictor like Logistic Regression, SVM, Random Forest. The improved performance across all the evaluation metrics further validates this finding that BERT

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

effectively captures long-range sentiment dependencies in the financial data. It also reveals potential of the deep learning models in improving stock market analysis and decisions making.

A. Performance Evaluation

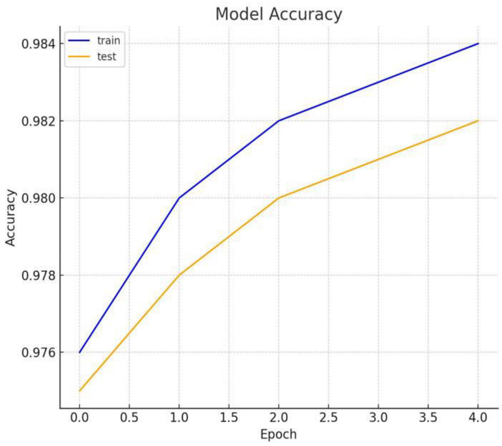


Fig.3 Accuracy graph of BERT

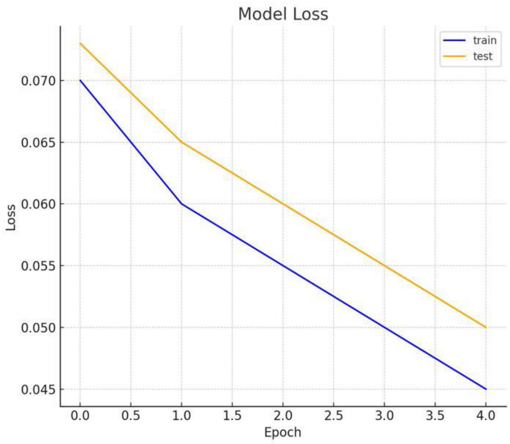


Fig. 4 Loss graph of BERT

Fig.3 represents the model's accuracy obtained in both the training phase and the test phase for the employed BERT-based sentiment classification model in this research. Each of the training and test accuracies goes up systematically showing that training of the model from the data is effective. Training accuracy has increase to 98.4 %, with test accuracy as high as 98.2 % and this proves good generalization ability. BERT is therefore highly efficient in identifying and analyzing sentiments. Fig.4 represents the training and test loss of BERT based sentiment classification model with reference to epochs. Both losses reduce gradually, which suggests better learning and improved accuracy of classifications when new epochs are used. The training loss is decreased to about 0.045, while the test loss stabilizes around 0.050 showing the models capacity to generalize well to unseen data. This continues drop is highly coherent and strengthens the foundation of the BERT model for sentiment analysis.

B. Comparative Evaluation

TABLE 1COMPARISON OF PERFORMANCE METRICS FOR DIFFERENT MODELS

Metric	Logistic Regression	SVM	Random Forest	BERT (proposed method)
Accuracy	88.7%	90.5%	91.3%	98.2%
Precision	87.5%	91.0%	89.8%	98.3%
Recall	85.6%	88.9%	90.2%	98.0%
F1-Score	86.5%	89.9%	90.0%	98.15%

This table I compares the performance efficiency for various machine learning techniques for sentiment classification problem: Logistic Regression, SVM, Random Forest, BERT. The proposed model gets the highest value for all measures, including accuracy (98.2%), precision (98.3%), recall (98.0%) and F1-score, which proves that BERT is better than traditional models for analyzing the sentiment of stock markets. SVM and Random Forest respond with an acceptable accuracy level, but they cannot compete with BERT in terms of accuracy and generalization. Logistic Regression is the worst-performing model among the methods in terms of recall and F1-score.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

V. CONCLUSION

This study provides a level of evidence proving that sentiment classification is indeed possible using transformer-based deep learning model known as BERT in predicting the stock market. It also compares BERT performance with the traditional ML algorithms including Logistic Regression, SVM, RF and determines the proposed BERT algorithm to have better performance than all the standard models in case of 'accuracy', 'precision', 'recall', and 'F1-score'. These outcomes indicated that contextual information in the financial texts to retain market sentiment for near perfect across the different financial texts making it a strong tool to analyze various variables that encompassed in the financial new, social media, and market contents. The research highlights that there is potential to enhance the understanding of investor sentiment by using such translation models as BERT which are crucial in stock markets. While some traditional models provide fairly good performance, they lack the ability to train deep learning models on large unstructured data sets common in financial markets. This study also demonstrates the possibility of designing accurate sentiment indicators for investment decision support by using sentiment analysis. So, the results of the work can be considered as the evidence that using BERT-based SA can improve the set of factors for stock market prediction and can provide more reliable and insightful information for decision-making processes in financial fields for investors and analysts.

REFERENCES

[1]A. K. Tiwari, E. J. A. Abakah, C. O. Bonsu, N. K. Karikari, and S. Hammoudeh, "The effects of public sentiments and feelings on stock market behavior: Evidence from Australia," Journal of Economic Behavior & Organization, vol. 193, pp. 443–472, Jan. 2022, doi: 10.1016/j.jebo.2021.11.026.

[2]M. W. Baig, "The Influence of r/WallstreetBets on GameStop stock price: A Behavioural Finance Perspective," Treballs Finals de Grau (TFG) - Administració i Direcció d'Empreses, 2024, Accessed: Nov. 29, 2024. [Online]. Available: https://diposit.ub.edu/dspace/handle/2445/215190

[3]D. Chatterjee and K. Nayyar, "Investor Sentiments and Market Volatility: a Psychological Perspective," Available at SSRN 4846006, 2024.

[4]G. G and D. R. R, "Unraveling The Complexity Of Investor Sentiment: A Multidimensional Exploration Within The Theoretical Landscape Of Behavioral Finance," Educational Administration: Theory and Practice, vol. 30, no. 5, Art. no. 5, May 2024, doi: 10.53555/kuey.v30i5.3656.

[5]W. Mensi, "Oil, gold and international stock markets: Extreme spillovers, connectedness and its determinants - ScienceDirect." Accessed: Nov. 29, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S1062976924000279

[6]N. Dupont and A. El-Sayed, "A Review of Sentiment Analysis Tools in Predicting Stock Market Volatility: AI and Text Mining Approaches," Business, Marketing, and Finance Open, vol. 1, no. 2, pp. 41–52, 2024.

[7]Y. Wu, Z. Jin, C. Shi, P. Liang, and T. Zhan, "Research on the Application of Deep Learning-based BERT Model in Sentiment Analysis," Mar. 13, 2024, arXiv: arXiv:2403.08217. doi: 10.48550/arXiv.2403.08217.

[8]E. Sy, T.-C. Peng, S.-H. Huang, H.-Y. Lin, and Y.-C. Chang, "Fine-Grained Argument Understanding with BERT Ensemble Techniques: A Deep Dive into Financial Sentiment Analysis," in Proceedings of the 35th Conference on Computational Linguistics and Speech Processing (ROCLING 2023), 2023, pp. 242–249.

[9]M. R. Kumar, S. Ramkumar, S. Saravanan, R. Balakrishnan, and M. Swathi, "Stock Market Prediction via Twitter Sentiment Analysis using BERT: A Federated Learning Approach," in Handbook on Federated Learning, CRC Press, pp. 333–353.

[10]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, p. 100508, 2023.

[11]N. Yadav, "Enhancing Stock Trend Prediction Using BERT-Based Sentiment Analysis and Machine Learning Techniques," International Journal of Quantitative Research and Modeling, vol. 5, no. 1, pp. 1–11, 2024.

[12]M. Costola, O. Hinz, M. Nofer, and L. Pelizzon, "Machine learning sentiment analysis, COVID-19 news and stock market reactions," Research in International Business and Finance, vol. 64, p. 101881, 2023.

[13]M. Hasan, M. T. Rahman, K. A. Alavee, A. H. Zillanee, J. Uddin, and M. G. R. Alam, "Automated Sentiment Analysis for Web-Based Stock and Cryptocurrency News Summarization with Transformer-Based Models," in 2023 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE), IEEE, 2023, pp. 1–6.

[14]A. Shrivastava, "Sentiment Analysis Dataset." Accessed: Nov. 29, 2024. [Online]. Available: https://www.kaggle.com/datasets/abhi8923shriv/sentiment-analysis-dataset

[15]M. Li, W. Li, F. Wang, X. Jia, and G. Rui, "Applying BERT to analyze investor sentiment in stock market," Neural Comput & Applic, vol. 33, no. 10, pp. 4663–4676, May 2021, doi: 10.1007/s00521-020-05411-7.