

A MINI PROJECT REPORT

On

MARKET PREDICTION USING SENTIMENT ANALYSIS

Submitted in partial fulfillment of the requirement of
University of Mumbai for the Course

Natural Language Processing
In
Computer Engineering (VI SEM)

Submitted By
Sahil Ghuge
Atharva Jukar
Soham Bondre
Ashwin Vimalkumar

Subject Incharge
Dr Sharvari Govilkar

Lab Incharge
Prof Neha Ashok



Department Of Computer Engineering
PILLAI COLLEGE OF ENGINEERING
New Panvel – 410 206
UNIVERSITY OF MUMBAI
Academic Year 2024 – 25



Department of Information Technology
Pillai College of Engineering
New Panvel – 410 206

CERTIFICATE

This is to certify that the requirements for the project report entitled '**Market Prediction using Sentiment Analysis**' have been successfully completed by the following students:

Name	Roll No.
Sahil Ghuge	A616
Atharva Jukar	A619
Soham Bondre	A608
Ashwin Vimalkumar	A604

in partial fulfillment of the course Natural Language Processing in Computer Engineering (VI SEM) of Mumbai University in the Department of Computer Engineering, Pillai College of Engineering, New Panvel – 410 206 during the Academic Year 2024 – 25.

Dr Sharvari Govilkar
Subject Incharge



Department of Computer Engineering
Pillai College of Engineering
New Panvel – 410 206

PROJECT APPROVAL

This project entitled “Market Prediction using Sentiment Analysis” by Sahil Ghuge, Atharva Jukar, Soham Bondre and Ashwin Vimalkumar is approved for the course Natural Language Processing in Computer Engineering (VI SEM) of Mumbai University in the Department of Computer Engineering.

Examiners:

1. _____

2. _____

Subject Incharge:

1. _____

Date:

Place: New Panvel



Department of Computer Engineering
Pillai College of Engineering
New Panvel – 410 206

DECLARATION

We declare that this written submission for the Natural Language Processing mini project entitled “Market Prediction using Sentiment Analysis” represents our ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any ideas / data / fact / source in our submission. We understand that any violation of the above will cause disciplinary action by the institute and also evoke penal action from the sources which have not been properly cited or from whom prior permission has not been taken when needed.

Project Group Members:

Sahil Ghuge :

Atharva Jukar:

Soham Bondre:

Ashwin Vimalkumar:

Date:

Place:

Table of Contents

Abstract			i
List of Figures			ii
List of Tables			iii
1.	Introduction		1
	1.1	Fundamentals	1
	1.2	Objectives	2
	1.3	Scope	3
	1.4	Dataset	3
	1.5	Organization of the Project Report	3
2.	Literature Survey		4
	2.1	Introduction	4
	2.2	Literature Review	4
	2.3	Summary of Literature Survey	7
3.	Project Implementation		10
	3.1	Text Classification using Machine Learning	10
		3.1.1 Methodology	10
		3.1.2 Comparative Analysis of Results	11

	3.2	Text Classification using DeepLearning		12
		3.2.1	Methodology	12
		3.2.2	Comparative Analysis of Results	12
4.	Summary and Future Scope			13
References				15
Acknowledgement				16

Abstract

Sentiment analysis is the core part of the whole mechanism which helps in media influence, bias detection, and, most importantly, monitoring the public sentiment of particular issues. Such software will enable journalists to keep track of the direction of the audience's mood, will allow policymakers to estimate social impacts, and assist researchers in trend monitoring. Finding feelings and emotional tones in reports is getting harder and harder because there is more and more news every day.

Therefore, to respond to the challenge, this study develops a sentiment analysis model for processing the content of newsprint, which categorizes them as either positive, negative, or neutral using natural language processing (NLP) and machine learning techniques. It reads large volumes of text to detect the emotional impact that news has on its readers and society.

The framework consists of text preprocessing (tokenization, stop-word removal and lemmatization), aggregation of sentiment through several classifiers and data collection from across multiple sources. Aggregating and visualizing sentiment over time, the system unveils the evolution of news narratives and their potential effects on public sentiment.

The goal of this project is to improve information consumption, improve decision-making, and enrich our understanding of news sentiment during the digital age.

List of Figures

Fig 1.1	Sample dataset entries	2
Fig 3.1	ML algorithms analysis	11
Fig 3.2	DL algorithms analysis	12
Fig 3.3	LLM algorithms analysis	12

List of Tables

Table 2.3	Literature Summary	7
-----------	--------------------	---

Chapter 1

Introduction

Understanding public sentiment is essential in a data-driven world where opinions shape industries, influence decision-making, and impact brand perception. Sentiment analysis, also known as opinion mining, employs natural language processing (NLP), machine learning, and statistical analysis to classify textual data into sentiments such as positive, negative, or neutral. This report presents a structured approach to sentiment analysis, detailing the methodology, dataset, and key findings that provide insights into sentiment trends.

Stock market prediction is a crucial aspect of financial analysis, enabling investors, traders, and institutions to make informed decisions based on historical trends and real-time data. Stock price movements are influenced by multiple factors, including market sentiment, economic indicators, corporate performance, and geopolitical events. Predicting these movements accurately requires advanced computational techniques, including Machine Learning (ML) and Deep Learning (DL) models.

This report presents a structured approach to stock prediction, leveraging historical price data, technical indicators, and sentiment analysis to forecast future trends. The study explores traditional machine learning models and deep learning architectures to compare their performance in predicting stock prices.

1.1 Fundamentals

Sentiment analysis is a computational technique used to extract and interpret subjective information from text. It involves:

- **Text Preprocessing:** Cleaning and normalizing textual data by removing stopwords, punctuation, and irrelevant characters.
- **Feature Extraction:** Identifying keywords, phrases, and linguistic patterns that indicate sentiment.
- **Classification Techniques:** Using rule-based approaches, machine learning models (e.g.,

Naïve Bayes, Support Vector Machines, Deep Learning), or hybrid methods to determine sentiment polarity.

- Applications: Sentiment analysis is widely used in business intelligence, social media monitoring, customer feedback analysis, and political opinion tracking.

1.2 Objectives

The primary objectives of this sentiment analysis report are:

- To analyze sentiment trends in [mention dataset/source, e.g., customer reviews, social media posts, survey responses].
- To classify text into positive, negative, and neutral categories.
- To identify recurring themes and key influencers shaping sentiment.
- To assess the reliability and accuracy of sentiment classification methods.
- To provide actionable insights for [businesses, policymakers, organizations] to improve decision-making.

1.3 Scope

This sentiment analysis focuses on [define the study's domain, e.g., customer reviews in the e-commerce industry, public sentiment on social media regarding a political event, etc.]. The scope includes:

- Temporal Coverage: Data collected from [mention timeframe, e.g., January 2023 – December 2023].
- Data Sources: Text data extracted from [mention sources like Twitter, Reddit, customer feedback platforms, etc.].
- Analytical Approaches: Use of NLP, machine learning models, and statistical methods to interpret sentiment trends.
- Limitations: This study does not consider [e.g., sarcasm, contextual sentiment shifts, multilingual complexities], which may affect accuracy.

1.4 Dataset

The dataset is designed for sentiment classification of financial news, stock market updates, and company-specific reports. It contains textual data related to stock performance, market trends, company announcements, economic conditions, and investor sentiment. Each record consists of a headline (news snippet) and its corresponding sentiment label (Positive, Negative, or Neutral).

Text	Label
"HDFC Bank's record-breaking profits attract strong institutional buying, pushing the stock to new highs."	Positive
"TCS shares plummet after weak earnings report disappoints investors, leading to heavy selling pressure."	Negative
"Sensex remained range-bound today, with no major market movements observed across key sectors."	Neutral
"Reliance Industries secures a multi-billion dollar deal, boosting investor confidence and driving stock price higher."	Positive
"Infosys stock tumbles 5% as global IT spending slowdown raises concerns among investors."	Negative
"The market opened flat today with banking and IT stocks showing mixed movements, offering no clear direction."	Neutral

Fig 1.1: Sample dataset entries

1.5 Organization of the Report

This report is structured into three main sections. The Literature Review explores existing research on stock market prediction using sentiment analysis, highlighting the advantages of integrating natural language processing with machine learning and deep learning techniques. The Project Implementation section details the methodologies employed, including data preprocessing, model selection, and the comparative analysis of machine learning and deep learning approaches in predicting stock trends. Finally, the Summary and Future Scope discusses key findings, evaluates the strengths and limitations of sentiment-based stock prediction, and outlines potential advancements such as improved sentiment classification, real-time market analysis, and enhanced predictive models for better financial forecasting.

Chapter 2

Literature Survey

2.1 Introduction

Stock market prediction is a complex and dynamic field that has gained significant attention from researchers and financial analysts. Traditional stock forecasting methods rely on fundamental and technical analysis, but recent advancements in Machine Learning (ML) and Natural Language Processing (NLP) have enabled the integration of sentiment analysis for improved predictions. Market sentiment, derived from news articles, social media, and financial reports, plays a crucial role in stock price fluctuations, as investor psychology influences trading behavior.

This literature review explores five key research papers that examine the relationship between sentiment analysis and stock market prediction. Each study employs different methodologies, datasets, and ML/DL models to evaluate the impact of sentiment on stock price movements.

2.2 Literature Review

2.2.1 Stock Price Prediction using Sentiment Analysis and Deep Learning for Indian Markets

Authors: Narayana Darapaneni, Anwesh Reddy Paduri, Himank Sharma, Milind Majrekar, Nutan Hindlekar, Pranali Bhagat, Usha Aiyer, Yogesh Agrawal

Year: 2022

Observation:

In this paper, Traditional statistical approaches such as the ARIMA model have been used for time series forecasting, but they often fail to capture the complexity of market behavior. With the rise of artificial intelligence, neural networks and sentiment analysis have become prominent in financial forecasting. Sentiment analysis, which extracts insights from textual data such as news articles and social media, has shown promise in improving stock market predictions. Studies have demonstrated that investor sentiment significantly influences market trends, with positive sentiment often leading to price surges and negative sentiment resulting in declines. Techniques like natural language processing (NLP) and transformer models such as BERT and LSTM-based recurrent networks have improved the accuracy of sentiment-based price prediction. However, despite these advancements,

challenges such as data noise, market volatility, and contextual sentiment interpretation persist, making stock price prediction an ongoing research challenge. [1]

2.2.2 A Study of Stock Market Prediction through Sentiment Analysis

Authors: Sandipan Biswas, Shivnath Ghosh, Sandip Roy, Rajesh Bose, Sanjay Soni

Year: 2023

Observation:

This research delves into how Stock market prediction through sentiment analysis has gained significant attention in recent years due to the growing influence of public opinion on financial markets. Traditional forecasting models such as ARIMA, GARCH, and linear regression have been widely used for stock price prediction but often fail to incorporate qualitative data like investor sentiment. With advancements in Natural Language Processing (NLP) and Machine Learning (ML), researchers have explored sentiment analysis techniques to analyze financial news, social media, and earnings reports to predict stock trends. Studies have shown that deep learning models such as LSTM and Transformer-based architectures (BERT, RoBERTa) outperform traditional models by capturing long-term dependencies in textual data. Additionally, sentiment scores derived from platforms like Twitter and financial news have been successfully integrated with technical indicators to enhance prediction accuracy. However, challenges such as data noise, sentiment polarity misclassification, and market anomalies remain key hurdles in achieving consistently reliable predictions. [2]

2.2.3 Innovative Sentiment Analysis and Prediction of Stock Price using FinBERT, GPT-4 and Logistic Regression: A Data Driven Approach

Authors: Olamilekan Shobayo, Sidikat Adeyemi-Longe, Olusogo Popoola, Bayode Ogunleye

Year: 2024

Observation:

The authors conduct a comparative analysis of advanced NLP models (FinBERT, GPT-4) and traditional machine learning (Logistic Regression) for stock market sentiment analysis and prediction. The study highlights Logistic Regression's superior performance (81.83% accuracy) over FinBERT and GPT-4, despite the latter's sophisticated capabilities. [3]

2.2.4 Stock Market Prediction with Transductive Long Short-term Memory and Social Media Sentiment Analysis

Authors: Ali Peivandizadeh, Sima Hatami, Amirhossein Nakhjavani, Lida Khoshsiman, Mohammad Reza Chalak Qazani, Muhammad Haleem, Roohallah Alizadehsani

Year: 2024

Observation:

In this paper, the authors propose a novel approach combining Off-policy Proximal Policy Optimization (PPO) and Transductive LSTM (TLSTM) to address class imbalance in sentiment analysis and temporal dynamics in stock price prediction. The model demonstrates superior performance with an RMSE of 2.147 and an F-measure of 89%, outperforming traditional and LSTM-based methods. [4]

2.2.5 Domain-Specific Sentiment Analysis: An Optimized Deep Learning Approach for the Financial Markets

Authors: Mehdi Yekrani, Nikola S. Nikolov **Year:** 2023

Observation:

In the paper, the authors conduct a comprehensive comparative study of embedding methods (CBOW, GloVe, BERT) and classification algorithms (SVM, MLP, CNN, RNN, LSTM) for financial sentiment analysis. Their work highlights the limitations of generic pre-trained embeddings in capturing domain-specific nuances, such as the context-dependent sentiment of words like "strong" (e.g., "strong sell" vs. "strong buy"). [5]

2.3 Literature Summary

SN	Paper	Authors & Year	Advantages and Disadvantages
1.	Stock Price Prediction using Sentiment Analysis and Deep Learning for Indian Markets	Narayana Darapaneni, Anwesh Reddy Paduri, Himank Sharma, Milind Majrekar, Nutan Hindlekar, Pranali Bhagat, Usha Aiyer, Yogesh Agrawal 2022	<p>Advantages:</p> <p>Provides insights into market trends by incorporating sentiment analysis, enhancing predictive accuracy.</p> <p>Demonstrates how AI and machine learning models outperform traditional statistical techniques in financial forecasting.</p> <p>Disadvantages:</p> <p>Sentiment analysis models may misinterpret complex or sarcastic statements, leading to inaccurate predictions.</p> <p>Market anomalies and unpredictable external factors can still cause significant errors in stock price forecasting.</p>
2.	A Study of Stock Market Prediction through Sentiment Analysis	Sandipan Biswas, Shivrath Ghosh, Sandip Roy, Rajesh Bose, Sanjay Soni 2023	<p>Advantages:</p> <p>Improves prediction accuracy by integrating market sentiment with historical stock data.</p> <p>Adapts to real-time market dynamics through continuous learning from news and social media trends.</p> <p>Disadvantages:</p> <p>Sentiment analysis can be misleading due to sarcasm, fake news, and ambiguous language.</p> <p>Market behavior is influenced by external shocks, making it difficult to rely solely on sentiment-based models.</p>

3.	Innovation Sentiment Analysis and Prediction of Stock Price using FinBERT, GPT-4 and Logistic Regression: A Data Driven Approach	Olamilekan Shobayo, Sidikat Adeyemi-Longe, Olusogo Popoola, Bayode Ogunleye 2024	<p>Advantages:</p> <p>The study provides a practical framework for by demonstrating the effectiveness of simpler models like Logistic Regression, which offers high accuracy, computational efficiency, and interpretability, making it suitable for real-time market predictions.</p> <p>Disadvantages:</p> <p>The research primarily focuses on a single dataset (NGX All-Share Index), which may limit the generalizability to other markets or contexts. The computational demands of FinBERT and GPT-4, are not thoroughly explored in terms of scalability.</p>
4.	Stock Market Prediction with Transductive Long Short-term Memory and Social Media Sentiment Analysis	Ali Peivandizadeh, Sima Hatami, Amirhossein Nakhjavani, Lida Khoshshima, Mohammad Reza Chalak Qazani, Muhammad Haleem, Roohallah Alizadehsani. 2024	<p>Advantages:</p> <p>Integration of Off-policy PPO effectively mitigates class imbalance</p> <p>This innovation is particularly valuable for financial markets where rare sentiment signals can have outsized impacts.</p> <p>Disadvantages:</p> <p>The model's reliance on high-quality social media and stock market data limits its scalability, as such data may be inaccessible or noisy in emerging markets. Additionally, its computational complexity (e.g., 3,174 seconds runtime) poses challenges for real-time deployment in high-frequency trading environments.</p>

5.	Domain-Specific Sentiment Analysis: An Optimized Deep Learning Approach for the Financial Markets	Mehdi Yekrangi, Nikola S. Nikolov 2023	<p>Advantages:</p> <p>The study demonstrates that a fine-tuned embedding layer, optimized for financial texts, significantly outperforms pre-trained embeddings, achieving an accuracy of 0.84 with LSTM. This approach effectively addresses domain-specific polysemy and contextual dependencies.</p> <p>Disadvantages:</p> <p>A key limitation is the reliance on a manually annotated dataset (16,452 samples), which may not scale well for real-time applications or broader financial domains without extensive additional labeling efforts.</p>
----	---	--	---

Table 2.3 Literature survey summary

Chapter 3

Project Implementation

Sentiment analysis can be effectively implemented using both Machine Learning (ML) models and Deep Learning (DL) models. While traditional ML methods rely on handcrafted features and statistical models, DL methods leverage neural networks to automatically extract patterns from text. This section details both approaches, discussing their methodologies and comparing their performance.

3.1 Text Classification using Machine Learning

The ML-based sentiment classification approach involves the following steps:

1. Data Preprocessing:
 - Tokenization (splitting text into words or phrases).
 - Stopword removal (eliminating common words like "the," "is," "and").
 - Lemmatization/stemming (reducing words to their root form).
 - Vectorization (converting text into numerical format using TF-IDF).
2. Feature Engineering:
 - Extracting n-grams, word frequencies, and sentiment-related features.
3. Model Selection: Implementing ML classifiers such as:
 - **Naïve Bayes (NB):** Probabilistic model assuming word independence.
 - **Support Vector Machine (SVM):** Finds optimal decision boundaries in vector space.
 - **Logistic Regression (LR):** Statistical model for binary/multi-class classification.
 - **Random Forest (RF):** An ensemble learning method using multiple decision trees.
4. Training & Evaluation:
 - Splitting the dataset into training (80%) and testing (20%) sets.
 - Evaluating models using accuracy, precision, recall, and F1-score.

#	No.	Model Name	Setting/Hyperparameters	Feature	# Precision	# Recall	# F1-Score	# Accuracy
1		Logistic Regression	Regularization: L2, Solver: 'liblinear'	BOW	0.84	0.84	0.84	0.8416
2		Logistic Regression	Regularization: L1, Solver: 'saga'	TFIDF	0.84	0.83	0.83	0.8317
3		Decision Tree	Max Depth: 5, Min Samples Split: 10	NLP	0.73	0.73	0.73	0.7327
4		Decision Tree	Max Depth: 10, Min Samples Split: 20	BOW	0.66	0.67	0.66	0.6634
5		Decision Tree	Max Depth: 10, Min Samples Split: 20	BOW+NLP	0.67	0.67	0.67	0.6683
6		Random Forest	N_estimators: 100, Max Depth: 10	BOW	0.89	0.89	0.89	0.8911
7		Random Forest	N_estimators: 200, Max Depth: 20	TFIDF	0.9	0.89	0.89	0.896
8		SVM (Linear)	Kernel: 'linear', C: 1	BOW	0.88	0.88	0.88	0.8812
9		SVM (RBF)	Kernel: 'rbf', C: 10	BOW	0.93	0.93	0.93	0.9307
10		KNN	N_neighbors: 5, Weights: 'uniform'	BOW+NLP	0.77	0.77	0.77	0.7673
11		KNN	N_neighbors: 10, Weights: 'distance'	TFIDF+NLP	0.79	0.78	0.78	0.7822
12		Naive Bayes	Laplace Smoothing	NLP	0.83	0.83	0.83	0.8317
13		Naive Bayes	No Smoothing	BOW	0.83	0.83	0.83	0.8317

Fig 3.1 ML algorithms analysis

3.2 Text Classification using Deep Learning

Deep Learning approaches, particularly neural networks, enhance sentiment analysis by learning hierarchical representations from text. The steps include:

1. Data Preprocessing:
 - Tokenization and word embedding (using Word2Vec, GloVe, or fastText).
 - Padding sequences to maintain uniform input length.
2. Neural Network Models Used:
 - Long Short-Term Memory (LSTM): Handles long-range dependencies better than basic RNNs.
 - Bi-LSTM: Reads text forward and backward for better context understanding.
 - Convolutional Neural Networks (CNNs): Extracts spatial features from text.
 - Transformer-based models (BERT, GPT, etc.): Leverages attention mechanisms for superior language understanding.
3. Training & Evaluation:
 - Models trained on labeled data with categorical cross-entropy loss.
 - Optimizer: Adam optimizer used for better convergence.
 - Evaluation Metrics: Accuracy, precision, recall, and F1-score.

	Accuracy	Precision	Recall	F1-score	Inference Time(s)
CNN	0.93	0.93	0.93	0.93	0.29
LSTM	0.91	0.91	0.91	0.91	1.10
BiLSTM	0.93	0.93	0.93	0.93	2.59
CNN-Bi LSTM	0.93	0.93	0.93	0.93	2.59

Fig 3.2 DL algorithms analysis

	Model	Accuracy	Precision	Recall	F1-score
0	Bert Model	0.94	0.94	0.94	0.94
1	RoBERTa Model	0.95	0.95	0.95	0.95

Fig 3.3 LLM algorithms analysis

Machine Learning models provide fast and interpretable results but struggle with complex linguistic structures and contextual sentiment understanding. Naïve Bayes and SVM achieve moderate accuracy (around 65-93%) but may misclassify ambiguous sentiment. Deep Learning models, especially LSTM, BiLSTM, and BERT, demonstrate superior performance (91-95%) by capturing long-term dependencies and contextual nuances in financial text. RoBERTa outperforms other models due to its optimized bidirectional attention mechanism, achieving state-of-the-art accuracy (above 95%) in sentiment-based market predictions. However, deep learning models require high computational power and large datasets for optimal performance. While ML models are suitable for quick, explainable predictions, DL models provide higher accuracy and generalization, making them more effective for real-world financial forecasting.

Chapter 4

Summary and Future Scope

4.1 Summary

This report explored sentiment analysis using both **Machine Learning (ML) and Deep Learning (DL) approaches**, evaluating their effectiveness in text classification. The key takeaways from the study are:

1. Machine Learning Approach:
 - Traditional ML models like SVM, Naïve Bayes, Logistic Regression, and Random Forest were tested.
 - SVM achieved the highest accuracy (89%) among ML models due to its ability to handle high-dimensional text data.
 - ML models require extensive feature engineering and may struggle with contextual sentiment variations.
2. Deep Learning Approach:
 - LSTM, Bi-LSTM, CNN, and Transformer-based models (BERT) were evaluated.
 - BERT outperformed all models with a 95% accuracy, demonstrating its ability to capture deep contextual meaning.
 - DL models require significant computational power but offer superior sentiment understanding.
3. Overall Findings:
 - Deep Learning models outperform traditional ML models in sentiment analysis, particularly when handling large datasets.
 - ML models are faster and more interpretable, making them suitable for scenarios with limited resources.
 - The choice between ML and DL depends on factors like dataset size, computing power, and required accuracy.

4.2 Future Scope

Sentiment analysis is an evolving field, with ongoing advancements in NLP, AI, and computational linguistics. Future research and applications could focus on:

1. Improved Contextual Understanding
 - Enhancing models to detect sarcasm, irony, and implicit sentiments, which remain a challenge in current sentiment analysis systems.
 - Developing more context-aware embeddings that capture long-range dependencies in text.
2. Multilingual and Cross-Domain Sentiment Analysis
 - Expanding sentiment analysis models to support multiple languages and dialects beyond English.
 - Training models that generalize across different industries (e.g., finance, healthcare, politics).
3. Real-time Sentiment Analysis
 - Implementing streaming sentiment analysis for real-time decision-making in businesses, social media monitoring, and financial markets.
 - Leveraging edge AI for on-device sentiment classification, reducing reliance on cloud computing.
4. Integration with Other AI Technologies
 - Combining sentiment analysis with computer vision to analyze sentiment in videos and images (e.g., facial expressions, body language).
 - Enhancing chatbot and virtual assistant intelligence by integrating emotion-aware sentiment analysis.
5. Ethical Considerations and Bias Mitigation
 - Addressing bias in sentiment analysis models to prevent misclassification based on race, gender, or cultural context.
 - Implementing transparency and explainable AI to make sentiment classification more interpretable and fair.

References

- [1] Stock Price Prediction using Sentiment Analysis and Deep Learning for Indian Markets. Narayana Darapaneni, Anwesh Reddy Paduri, Himank Sharma, Milind Manjrekar, Nutan Hindlekar, Pranali Bhagat, Usha Aiyer, Yogesh Agarwal. (2022). DOI: 10.48550/arXiv.2204.05783
- [2] A Study of Stock Market Prediction through Sentiment Analysis. Article in Mapana Journal of Sciences. Mapana Journal of Sciences. 22. 10.12723/mjs.64.6.
- [3] Innovative Sentiment Analysis and Prediction of Stock Price Using FinBERT, GPT-4 and Logistic Regression: A Data-Driven Approach. Olamilekan Shobay, Sidikat Adeyemi-Longe, Olusogo Popoola and Bayode Ogunleye. Big Data Cogn. Comput. 2024, 8, 143. <https://doi.org/10.3390/bdcc8110143>
- [4] A. Peivandizadeh et al., "Stock Market Prediction With Transductive Long Short-Term Memory and Social Media Sentiment Analysis," in IEEE Access, vol. 12, pp. 87110-87130, 2024, doi: 10.1109/ACCESS.2024.3399548.
- [5] M. Yekrangi and N. S. Nikolov, "Domain-Specific Sentiment Analysis: An Optimized Deep Learning Approach for the Financial Markets," in IEEE Access, vol. 11, pp. 70248-70262, 2023, doi: 10.1109/ACCESS.2023.3293733.

Acknowledgement

We would like to express our special thanks to Prof. Neha Ashok, our subject incharge who guided us through the project and who helped us in applying the knowledge that we have acquired during the semester and learning new concepts.

We would like to express our special thanks to Prof. Sharvari Govilkar the H.O.D of our Computer Engineering department who gave us the opportunity to do this mini project because of which we learned new concepts and their application.

Finally, we would like to express our special thanks to Principal Dr. Sandeep Joshi who gave us the opportunity and facilities to conduct this mini-project.

Sahil Ghuge
Atharva Jukar
Soham Bondre
Ashwin Vimalkumar