

**CHAITANYA BHARATHI INSTITUTE OF TECHNOLOGY**  
Gandipet- 500 075, Hyderabad  
**DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND DATA SCIENCE**  
**MINOR PROJECT-II (DATA SCIENCE), 2023-2024**  
**CLASS: B.E VI Semester, AI&DS-2**

**TITLE:** Strategic Stock Market Position Optimization

ROLL NO.	NAME	SIGNATURE
160121771076	Chinam Chandana	
160121771088	Pasunuri Kathyayini	
160121771095	Ambekar Tejas	

**ABSTRACT:**

In the dynamic realm of stock market trading, optimizing positions to maximize gains and minimize losses is paramount for traders seeking success. Historical data analysis of option chains for prominent indices like Nifty, BankNifty, and FinNifty offers valuable insights into market trends and potential opportunities. This project aims to leverage such data to develop a strategic position optimization framework. Stock market traders often rely on retrospective data to inform their trading strategies. However, manually analyzing vast amounts of option chain data spanning several years can be daunting and time-consuming. This project seeks to automate this process by developing a data-driven approach to optimize trading positions. By analyzing historical trends and patterns in Nifty, BankNifty, and FinNifty option chains, the system will recommend strategic positions that increase gains while minimizing potential losses.

Currently, traders may employ manual or semi-automated methods to analyze historical data and optimize their positions. However, these approaches are often limited in their scope and efficiency. By introducing a comprehensive algorithmic solution, this project aims to streamline the optimization process and provide traders with actionable insights in real-time. The proposed solution will involve collecting and preprocessing historical option chain data for Nifty, BankNifty, and FinNifty indices. Machine learning algorithms will then be deployed to analyze this data and identify patterns indicative of potential market movements. Using these insights, the system will generate optimized trading positions that align with the trader's risk tolerance and investment objectives. Additionally, the platform will offer features such as backtesting to validate the effectiveness of proposed strategies and fine-tune parameters for optimal performance. The development of this project will predominantly utilize R for its robust capabilities in data analysis and machine learning. R packages such as tidyverse, caret, and xgboost will facilitate data manipulation, modeling, and prediction tasks. For data visualization, ggplot2 will be employed to create insightful plots and visualizations. Machine Learning Model for prediction and optimization of prices can be built using Python and other required libraries.

# DETAILS OF THE DATASET

## Names of Datasets (Stocks) Considered for Analysis:

- **BANKNIFTY** : an index comprised of the most liquid and large capitalised Indian banking stocks.

### List of banks included:

1. AxisBank
2. PNB
3. Bandhan Bank
4. Kotak Bank
5. IDFC First Bank
6. Federal Bank
7. HDFC Bank
8. IndusIND Bank
9. State Bank of India
10. Bank of Baroda
11. ICICI Bank
12. AU Bank

- **NIFTY50**: an index that represents the performance of the top 50 companies listed on the National Stock Exchange (NSE) of India.

### List of companies included:

1. Asian Paints Ltd
2. Britannia Industries Ltd
3. Cipla Ltd
4. Eicher Motors Ltd
5. Nestle India Ltd
6. Grasim Industries Ltd
7. Hero MotoCorp Ltd
8. Hindalco Industries Ltd
9. Hindustan Unilever Ltd
10. ITC Ltd
11. Larsen & Toubro Ltd
12. Mahindra & Mahindra Ltd
13. Reliance Industries Ltd
14. Tata Consumer Products Ltd

15. Tata Motors Ltd
16. Tata Steel Ltd
17. Wipro Ltd
18. Apollo Hospitals Enterprise Ltd
19. Dr Reddys Laboratories Ltd
20. Titan Company Ltd
21. State Bank of India
22. Shriram Finance Ltd
23. Bharat Petroleum Corporation Ltd
24. Kotak Mahindra Bank Ltd
25. Infosys Ltd
26. Bajaj Finance Ltd
27. Adani Enterprises Ltd
28. Sun Pharmaceuticals Industries Ltd
29. JSW Steel Ltd
30. HDFC Bank Ltd
31. Tata Consultancy Services Ltd
32. ICICI Bank Ltd
33. Power Grid Corporation of India Ltd
34. Maruti Suzuki India Ltd
35. IndusInd Bank Ltd
36. Axis Bank Ltd
37. HCL Technologies Ltd
38. Oil & Natural Gas Corpn Ltd
39. NTPC Ltd
40. Coal India Ltd
41. Bharti Airtel Ltd
42. Tech Mahindra Ltd
43. LTIMindtree Ltd
44. Divis Laboratories Ltd
45. Adani Ports & Special Economic Zone Ltd
46. HDFC Life Insurance Company Ltd
47. SBI Life Insurance Company Ltd
48. UltraTech Cement Ltd
49. Bajaj Auto Ltd
50. Bajaj Finserv Ltd

- **FINNIFTY:** an index that includes the stock values of various companies that are part of the Indian financial sector.

**Dataset parameters:**

- **Time Period:** Sep 17, 2007 - Present day (Updated daily)
- **Considered:** Historical Prices
- **Frequency:** Daily

**Columns:**

1. **Date:** Given in “month dd, yyyy” format (eg: Mar 28, 2024)
2. **Open:** The opening price of the stock for the corresponding date, given upto 2 decimal points
3. **Close:** The closing price of the stock for the corresponding date, given upto 2 decimal points
4. **High:** The highest price reached by the stock on the corresponding date, given upto 2 decimal points
5. **Low:** The lowest price reached by the stock on the corresponding date, given upto 2 decimal points
6. **Adj. Close:** the closing price after adjustments for all applicable splits and dividend distributions.
7. **Volumes:** number of options, contracts bought or sold on a given trading day

**Note:** All prices are given in INR

**Number of rows:** Updating daily, around 4083 as of now.

**Source:** <https://finance.yahoo.com/>

# LITERATURE SURVEY

In recent years, advancements in computational techniques have significantly impacted stock market analysis and prediction. Four research papers focusing on forecasting the movement of the Nifty50 index were surveyed to understand the diverse methodologies employed in this domain.

The first paper, "Stock Market Analysis and Prediction for Nifty50 using LSTM Deep Learning Approach," proposed a Long Short-Term Memory (LSTM) algorithm leveraging ten years of historical data to forecast stock price movement. Achieving an accuracy of 83.88%, this study challenges the traditional belief that stock markets are inherently unpredictable.

The second paper, "A systematic deep learning approach to forecast Nifty50 index trend," introduced a deep learning model integrating sentiment analysis to predict Nifty50 stock trends. This approach achieved a remarkable prediction accuracy of over 96% by combining sentiment analysis with market positions, providing investors with valuable insights amidst market volatility.

In the third paper, "PSO-tuned support vector classifier based Nifty50 index movement prediction using market positions and sentiment scores," an effective model was proposed that combined social media sentiment analysis with market positions to forecast daily returns of the Nifty50 index. Through careful hyperparameter tuning and supervised machine learning techniques, the model achieved over 99% prediction accuracy.

Finally, the fourth paper, "Analyzing the Effectiveness of Machine Learning Models in Nifty50 Next Day Prediction: A Comparative Analysis," conducted a comparative analysis of five machine learning algorithms in predicting NIFTY50 stock prices. Support Vector Regression (SVR) emerged as the most accurate model, closely followed by Lasso Regression and Random Forest, highlighting the efficacy of machine learning in financial forecasting.

These studies collectively illustrate the diverse range of methodologies and approaches employed in stock market prediction, emphasizing the growing importance of advanced computational techniques in enhancing prediction accuracy and informing investment decisions.

AUTHOR AND YEAR	DATASET	TITLE AND ABSTRACT	FUTURE SCOPE	LINK
Pushpendra Singh Sisodia, Anish Gupta, Yogesh Kumar, Gaurav Kumar Ameta (2022)	10-year historical data from the National Stock Exchange (NSE) for 10 random NIFTY 50 equities for the period 10 December 2011 to 10 December 2021	<b>Stock Market Analysis and Prediction for Nifty50 using LSTM Deep Learning Approach</b> This project proposes a Deep Learning-based LSTM Algorithm for predicting stock price movement using ten years of historical data from India's NIFTY 50 index. With an accuracy of 83.88%, the model demonstrates promising results in forecasting stock price movements, challenging the notion that precise predictions are impossible in stock markets.	This study introduces an LSTM model for predicting stock prices of ten Nifty50 equities over a decade, yielding highly accurate results. Metrics such as RMSE, MSE, MAE, and MAPE indicate minimal prediction errors across the ten stocks. Notably, SBI stock achieved the highest accuracy of 83.88%. Future work aims to enhance forecast accuracy further.	<a href="https://ieeexplore.ieee.org/document/9754148">https://ieeexplore.ieee.org/document/9754148</a>
Deepshi Garg, Rakesh Kumar Pandey, Prakash Tiwari (2023)	Yahoo Finance and four social media platforms, including Twitter, StockTwits, Facebook, and YouTube, for four years from 2018 to 2021.	<b>A systematic deep learning approach to forecast Nifty50 index trend</b> A deep learning method integrating sentiment analysis was developed to predict Nifty50 stock trends amidst high volatility. By combining sentiment scores with market positions, a novel approach achieved over 96% prediction accuracy using LSTM and densely connected neural networks, empowering investors in making informed stock predictions.	Further enhancements could make it more accessible, inclusive, and intuitive for diverse individuals, analysts, and policymakers. The proposed model is robust and dependable to forecast NIFTY50 returns. In future work, we intend to investigate the performance of the proposed model to examine the returns of the sector-specific market indices.	<a href="https://ieeexplore.ieee.org/document/10444538">https://ieeexplore.ieee.org/document/10444538</a>
Deepshi Garg, Rakesh Kumar Pandey, Prakash Tiwari	Yahoo Finance and four social media sites, including Twitter, StockTwits,	<b>PSO-tuned support vector classifier based Nifty50 index movement prediction using market positions and sentiment scores</b>	The study intends to investigate the proposed model's performance to examine the sector-specific market returns.	<a href="https://ieeexplore.ieee.org/document/10306670">https://ieeexplore.ieee.org/document/10306670</a>

(2023)	Facebook, and Youtube, for four years, from 2018 to 2021.	This study proposes an effective model combining social media sentiment analysis with market positions to forecast Nifty50 daily returns, achieving over 99% prediction accuracy through carefully tuned hyperparameters and supervised machine learning techniques.		
Dhruv Sharma, Amisha, Pradeepta Kumar Sarangi, Ashok Kumar Sahu (2023)	NIFTY50 index data for the period of March 2022 to March 2023.	<b>Analyzing the Effectiveness of Machine Learning Models in Nifty50 Next Day Prediction: A Comparative Analysis</b> This research compares the performance of five machine learning algorithms in predicting NIFTY50 stock prices using data from March 2022 to March 2023. SVR stands out as the most accurate model, followed closely by Lasso Regression and Random Forest, highlighting the efficacy of machine learning in financial forecasting.	Future research should be done to make these models more and more reliable and accurate. So, these models can be used for making investments-based decisions	<a href="https://ieeexplore.ieee.org/document/10182943">https://ieeexplore.ieee.org/document/10182943</a>

**Date of Submission:**

**Name & Signature of  
Project Coordinator**

**Name & Signature of  
Mentor**

**Name & Signature of Class  
Coordinator**