Stock Market Prediction Using ML

# Submitted for

**Statistical Machine Learning CSET211**

Submitted by:

# (<E23CSEU2161>) TANYA MISHRA (<E23CSEU2165>) SAYANTIKA CHOWDHURY (<E23CSEU2190>) ISHITA SHRIVASTAVA

Submitted to

# DR. PRASHANT KAPIL

**July-Dec 2024**

# SCHOOL OF COMPUTER SCIENCE AND ENGINEERING



**INDEX**

|  |  |  |
| --- | --- | --- |
| Sr.No | Content | Page  No |
| 1 | ABSTRACT | 3 |
| 2 | INTRODUCTION | 3 |
| 3 | RELATED WORK | 3 |
| 4 | METHODOLOGY | 4 |
| 5 | SOFTWARE REQUIRED | 5 |
| 6 | EXPERIMENTAL RESULTS | 5 |
| 7 | CONCLUSION | 5 |

1. **ABSTRACT**

This project focuses on the predictive analysis of stock market trends using machine learning techniques. The stock market is characterized by its volatility and complexity, making accurate predictions challenging yet essential for investors. This study surveys various machine learning algorithms and methodologies employed in stock price forecasting, including Random Forest, Artificial Neural Networks (ANNs), and Long Short-Term Memory (LSTM) networks. By analyzing historical stock prices and sentiment data, we aim to develop a robust predictive model that enhances investment decision-making. The results demonstrate the effectiveness of machine learning in capturing market trends and improving prediction accuracy.

https://github.com/Sayantika2327/stock-market-prediction-ml.git

1. **INTRODUCTION**

The stock market is a dynamic environment influenced by numerous factors, including economic indicators, market sentiment, and historical price movements. Predicting stock prices is crucial for traders and investors to make informed decisions. Traditional methods of analysis often fall short in capturing the complexities of market behavior.

This project leverages machine learning algorithms to analyze historical data and sentiment analysis from news articles, aiming to create predictive models that outperform conventional statistical methods. The goal is to provide insights into market trends and enhance the accuracy of stock price predictions.

1. **RELATED WORK**

The literature survey highlights several significant studies in the field of stock market prediction using machine learning by Sayantika Chowdhury and etda:

The study focuses on predicting stock prices using historical and sentiment data, highlighting Random Forest’s superior performance over logistic regression. It emphasizes the importance of data preprocessing and sentiment polarity in enhancing predictions and reviews various machine learning techniques, proposing hybrid models to further improve accuracy in stock market forecasting. These studies establish a foundation for our project, showcasing the effectiveness of machine learning in capturing market trends and improving prediction accuracy.

1. **METHODOLOGY**

The methodology involves several key steps:

* 1. **Data Collection**: Historical stock prices and sentiment data from news articles are collected. The datasets are preprocessed to derive sentiment polarity scores.
  2. **Data Preprocessing**: The data undergoes cleaning, normalization, and transformation to ensure it is suitable for analysis.
  3. **Model Implementation**: Various machine learning algorithms, including Random Forest, ANNs, and LSTM, are implemented to analyze the data and generate predictions.
  4. **Model Evaluation**: The models are evaluated using performance metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), and accuracy to determine their effectiveness in predicting stock prices.
  5. **Comparison**: The performance of different models is compared to identify the most effective approach for stock price prediction.

1. **SOFTWARE REQUIRED**

The following software tools and libraries are required for the project:

* Python Programming Language
* Libraries:
  + Pandas (for data manipulation)
  + NumPy (for numerical calculations)
  + Scikit-learn (for machine learning algorithms)
  + TensorFlow/Keras (for deep learning models)
  + Matplotlib/Seaborn (for data visualization)

1. **EXPERIMENTAL RESULTS**

The experimental results demonstrate the efficacy of the implemented models. The Random Forest model showed superior performance in predicting stock prices, achieving lower MAE and MSE compared to logistic regression. The LSTM model also exhibited high accuracy, particularly in capturing temporal dependencies in stock price movements. The results indicate that machine learning models can significantly enhance prediction accuracy and provide valuable insights into market trends.

1. **CONCLUSION**

This project demonstrates the effective use of machine learning, particularly Random Forest and LSTM, in predicting stock market trends using historical data and sentiment scores. Future work could integrate high-frequency trading data and economic indicators, develop hybrid models for improved accuracy, and explore reinforcement learning for real-time trading strategies.