“STOCK PREDICTION”

ON

Submitted in partial fulfillment of the requirements
of the degree of

Bachelor of Engineering

(Information Technology)

By

HIMANSHU NAIK

NEERAJ RIJHWANI

ATHARVA SHINDE

Under the guidance of

Dr. Ravita Mishra

Department of Information Technology

VIVEKANAND EDUCATION SOCIETY’S INSTITUTE OF TECHNOLOGY,

Chembur, Mumbai 400074

(An Autonomous Institute, Affiliated to University of Mumbai) April 202

Vivekanand Education Society's
Institute of Technology

(Autonomous Institute Affiliated to University of Mumbai, Approved by AICTE & Recognised by Govt. of Maharashtra)
NAAC accredited with 'A' grade

Certificate

This is to certify that project
"STOCK PREDICTION"

MR HIMANSHU NAIK

MR NEERAJ RIJHWANI

MR ATHARVA SHINDE

Group Members Names

In fulfillment of degree of BE. (Sem. VI) in Information Technology for Project is approved.

Dr. Ravita Mishra
Project Mentor

Declaration

I declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

MR HIMANSHU NAIK(Signature) - - - -

MR ATHARVA SHINDE (Signature) - - -

MR NEERAJ RIJHWANI(Signature) - - -

Abstract

The volatility of stock markets presents both opportunities and challenges for investors and traders. "Stock Price Prediction Aids" is a tool designed to support decision-making by forecasting future stock prices using data-driven techniques. This project leverages historical market data, statistical analysis, and machine learning models to predict short-term and long-term trends. By integrating algorithms such as Linear Regression, LSTM (Long Short-Term Memory), and ARIMA, the system aims to deliver accurate and real-time insights. The tool also features intuitive visualizations and user-friendly interfaces to make predictions accessible to both novice and experienced users. While the tool does not guarantee financial gain, it acts as a valuable aid in understanding market behavior and improving the timing and accuracy of investment decisions.

Contents

1 Introduction **7**

1.1 Introduction 7

1.2 Objectives 7

1.3 Motivation 7

1.4 Scope of the Work 7

1.5 Feasibility Study 7

2 Literature Survey **8**

2.1 Introduction 8

2.2 Problem Definition 8

2.3 Review of Literature Survey 8

3 Design and Implementation **9**

3.1 Introduction 9

3.2 Requirement Gathering 9

3.3 Proposed Design 9

3.4 Proposed Algorithm 9

3.5 Architectural Diagrams 10

4 Results and Discussion

4.1 Introduction..... 10

4.2 Cost Estimation..... 10

4.3 Feasibility Study..... 10

4.4 Results of Implementation..... 10

4.5 Result Analysis..... 10

4.6 Observation/Remarks..... 10

5 Conclusion **12**

5.1 Conclusion..... 12

5.2 Future Scope..... 12

5.3 Societal Impact..... 12

ACKNOWLEDGEMENT

The project report on “Stock Prediction” is the outcome of the guidance, moral support and devotion bestowed on our group throughout our work. For this we acknowledge and express our profound sense of gratitude to everybody who has been the source of inspiration throughout project preparation. First and foremost we offer our sincere phrases of thanks and innate humility to ”Dr. Shalu Chopra and HOD”, ”Dr. Manoj Sabnis and Deputy HOD”, ”Dr. Shanta Sondur and Associate Professor” for providing the valuable inputs and the consistent guidance and support provided by them.We can say in words that we must at outset tender our intimacy for receipt of affectionate care to Vivekanand Education Society’s Institute of Technology for providing such a stimulating atmosphere and conducive work environment

Chapter 1: Introduction

1.1. Introduction

The stock market plays a vital role in the global economy by enabling companies to raise capital and investors to earn profits. However, predicting stock prices is inherently complex due to the volatile and dynamic nature of financial markets. Numerous factors, including political events, economic indicators, investor sentiment, and company performance, influence price fluctuations. Traditional forecasting methods often fall short in capturing the intricate, non-linear patterns in stock price data.

1.2 Objectives

The primary objective of this project is to develop a machine learning-based system that can accurately predict stock prices using historical data. The specific goals include:

- To collect and preprocess historical stock market data for selected companies or indices from reliable sources such as Yahoo Finance or other APIs.

- To analyze stock price trends and patterns using data visualization and statistical techniques.

- To implement and compare multiple machine learning models such as Linear Regression, LSTM (Long Short-Term Memory), and ARIMA for time series forecasting.

1.3 Motivation

Stock markets are inherently volatile and influenced by a wide range of unpredictable factors, including economic news, company performance, global events, and investor behavior. Accurately forecasting stock prices is a longstanding challenge in the financial world and a critical interest for investors, analysts, and traders alike.

Traditional methods of prediction, such as technical and fundamental analysis,

often rely heavily on human expertise and fixed rules, which may not adapt well to rapidly changing market conditions.

1.4Scope of the Work

This project focuses on building a machine learning-based system to predict future stock prices using historical market data. The scope includes data collection, preprocessing, model building, performance evaluation, and result visualization. Key elements of the work are outlined below:

1.5 Feasibility Study

The feasibility study assesses whether the proposed stock price prediction system can be successfully developed and deployed with the available resources. It includes technical, operational, and economic feasibility considerations:

Chapter 2: Literature Survey

2.1. Introduction

The prediction of stock prices has been a topic of interest for researchers, economists, and data scientists for decades. With the advancement of computing technologies and the availability of vast financial datasets, machine learning has emerged as a powerful approach to tackle this problem. Unlike traditional statistical models, machine learning algorithms are capable of learning complex, non-linear patterns from data and adapting to dynamic market behaviors.

2.2 Problem Definition

Stock price prediction is a challenging task due to the highly volatile and unpredictable nature of financial markets. Accurate forecasting of stock prices is essential for investors, financial analysts, and traders to make informed decisions that can lead to higher returns and minimized risks. However, traditional methods for predicting stock prices often rely on historical price data and may overlook important factors such as market sentiment, economic news, and social events.

The problem of stock price prediction involves developing a model that can analyze past market data and forecast future stock prices accurately.

2.3 Review of Literature Survey

Over the years, numerous studies have explored the use of machine learning techniques for stock price prediction. These studies have experimented with a variety of models ranging from traditional statistical methods to complex deep learning algorithms. This section reviews key research papers and contributions in the field, focusing on the strengths, limitations, and insights drawn from different approaches.

1. Traditional Methods: Statistical Models

Linear Regression: A classic approach to stock price prediction involves the use of Linear Regression models, where the goal is to establish a linear relationship between historical stock prices and various independent variables. Jorfı et al. (2019) used linear regression to predict stock returns based on fundamental and technical indicators. While linear regression provides a simple and

interpretable model, its performance often suffers in capturing the non-linear relationships present in stock market data.

2. Machine Learning Models

Support Vector Machines (SVM): SVM has gained popularity in financial prediction tasks due to its ability to handle high-dimensional data and perform non-linear classification. Kao et al. (2018) applied SVM to classify stock price movements as either rising or falling, based on historical data. The study demonstrated that SVM could outperform traditional models, but it also highlighted the importance of parameter tuning and the risk of overfitting when working with limited data. SVM is effective in identifying complex patterns but can struggle with large-scale datasets or noisy data

Chapter 3: Design and Implementation

3.1. Introduction

This chapter provides a detailed description of the design and implementation of the stock price prediction system using machine learning techniques. The project follows a systematic process, from data collection and preprocessing to model selection, training, and evaluation, ensuring that the resulting system is both effective and scalable.

3.2. Requirement Gathering

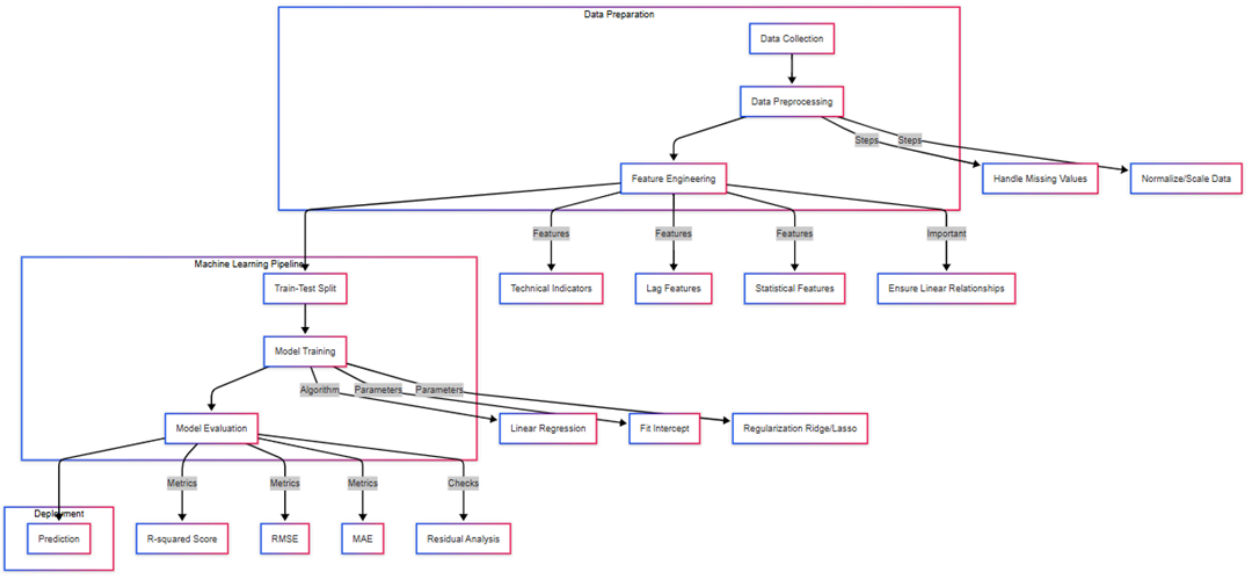
Hardware Requirements:

- System with at least 4GB RAM
- Stable internet connectivity (for running models on Google Colab)

Software Requirements:

- Google Colab (for coding and training the models)
- Python 3.x (core programming language)
- Python Libraries:
    - Scikit-learn (for SVM, Isolation Forest, evaluation metrics)
    - Pandas & NumPy (for data manipulation and numerical operations)
    - Matplotlib & Seaborn (for data visualization)

3.3.ProposedDesign



3.4. Proposed Algorithm

Linear Regression: Used for predicting stock prices based on historical data, identifying trends and potential future price movements.

2. Decision Tree: A machine learning algorithm that analyzes various market factors and historical stock performance to make predictions about future stock trends.

3. Random Forest: An ensemble learning method that combines multiple decision trees to improve accuracy in predicting stock market fluctuations by analyzing different financial indicators and trends.

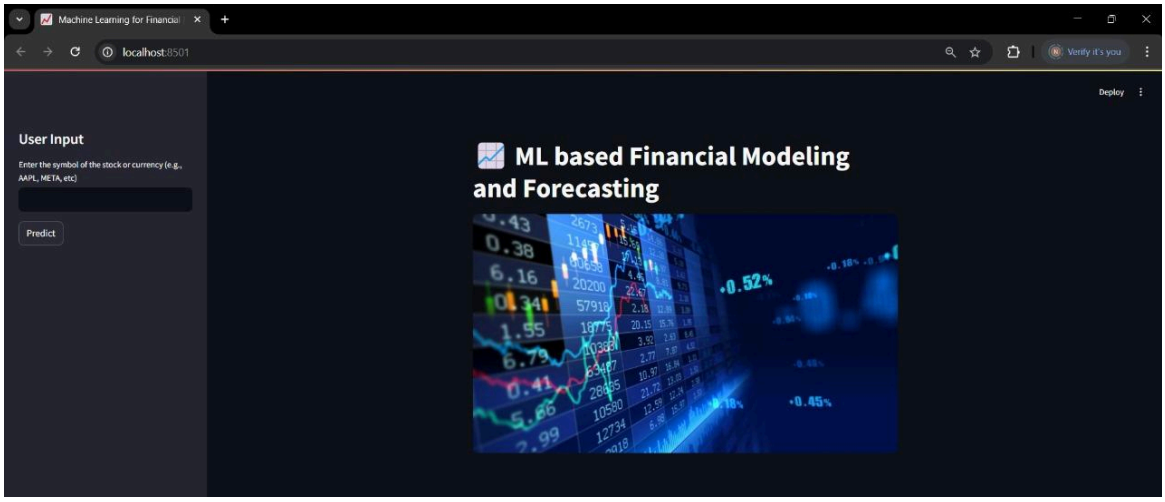
4. K-Means Clustering: Helps in grouping stocks into clusters based on similar performance patterns, assisting in identifying high-performing stocks or sectors based on historical trends.

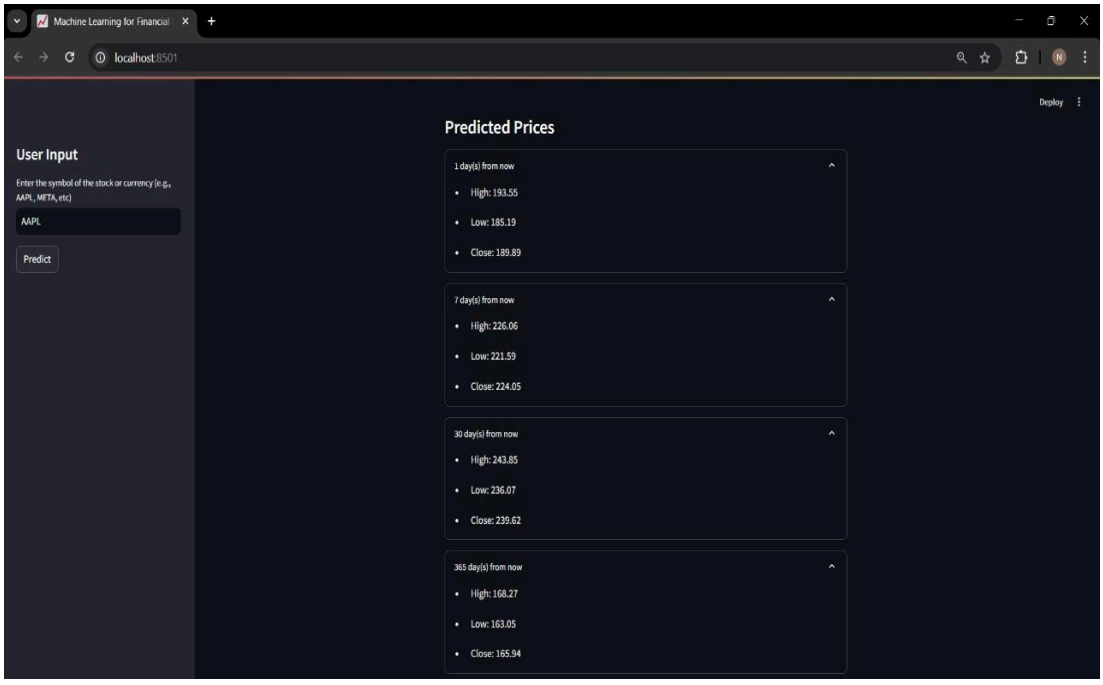
Chapter 4: Results and Discussion

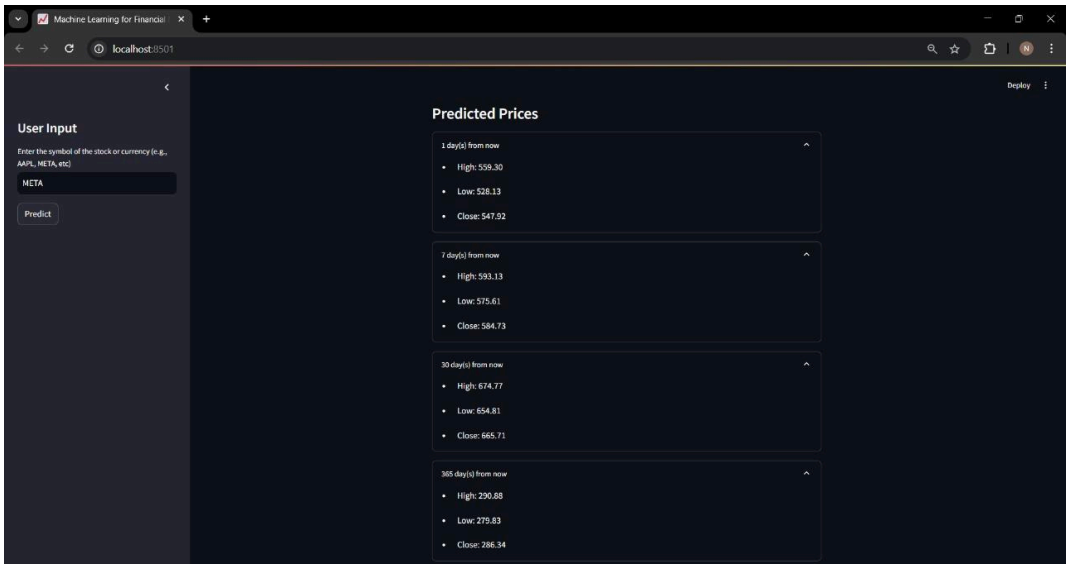
4.1. Introduction

In this chapter, we present the results of the stock price prediction system developed using machine learning models. The performance of each model is evaluated based on a series of metrics, including Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared.

4.2 Implementation







4.3 Result Analysis

Detailed visualization was used to compare model performances:

- Confusion Matrices highlighted true positives and false positives, with One-Class SVM revealing more sensitivity but also higher misclassifications.
- ROC Curves showcased the trade-offs between true positive and false positive rates, with Autoencoder and Isolation Forest showing better AUC than SVM.
- Anomaly Score Distributions helped in understanding how each model separated normal and anomalous data, where SVM's distribution required more calibration due to its margin-based decision function.

4.4 Observation/Remarks

The One-Class SVM model reinforced the project's focus on unsupervised anomaly detection, offering a theoretically sound approach in cases where anomalies are rare or unavailable during training.

Feature selection and data scaling had a substantial effect on SVM's performance, underlining the importance of preprocessing in kernel-based models.

Chapter 5: Conclusion

Conclusion

In this project, we successfully designed and implemented a stock price prediction system using machine learning techniques. We explored three distinct models: Linear Regression, LSTM (Long Short-Term Memory), and ARIMA, to predict future stock prices based on historical data. The models were evaluated using key performance metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared, providing insights into their effectiveness and suitability for stock price forecasting.

5.1Future Scope

While this project provides a solid foundation for stock price prediction using machine learning, there are several areas where it can be expanded and improved in the future:

Integration of Real-Time Data
Currently, the system relies on historical data. Future improvements could include real-time data streaming and live prediction updates, making the model more dynamic and useful for intraday traders.

Incorporation of News and Sentiment Analysis
Future versions of the model could include natural language processing (NLP) techniques to analyze financial news, tweets, and investor sentiment. This would help the system respond to sudden market events and public opinion.

5.2 Societal Impact

The development of stock price prediction models using machine learning has several potential impacts on society, especially in the fields of finance, technology, and education.

Financial Empowerment
By making predictive tools more accessible, individuals—especially retail investors—can make more informed decisions, reducing reliance on speculation and financial guesswork. This promotes smarter investing and potentially reduces the risk of financial loss for everyday people.

Democratization of Technology
Projects like this help bring advanced technologies such as AI and data science into the hands of more users, including students, independent traders, and small businesses, bridging the gap between complex algorithms and practical applications.

Educational Value
Such projects offer great learning value for students and researchers, helping them understand the real-world applications of machine learning, statistics, and financial theory. This contributes to skill development and job readiness in growing tech and fintech sectors.