**MINI PROJECT REPORT**Title: STOCK PRICE PREDICTION  
  
Submitted by: A.S. DHANYA AASRITHA  
Roll Number: 2024021839  
Course: MCA, 2nd Year  
Guided by: Pavan Kumar Ilapanda  
Date: June 2025

**Table of Contents**

1. Abstract  
2. Introduction  
3. Problem Statement  
4. Objectives  
5. Literature Survey  
6. Tools & Technologies Used  
7. Data Collection  
8. Data Preprocessing  
9. Model Implementation  
10. Evaluation Metrics  
11. Results  
12. Limitations  
13. Future Work  
14. Conclusion  
15. References

* Abstract

Stock markets are volatile, and predicting their prices is challenging. This project uses machine learning (linear regression) to predict the next-day closing price of a stock using historical stock data. The data is collected using Yahoo Finance API (finance library), and the model is trained using Sci-Kit.

* **Introduction**

Stock price prediction is a key area in financial data analysis. Investors use predictions to make informed decisions. In this project, we use historical data to forecast future stock prices using a regression model.

* **Problem Statement**

To design a machine learning model that can predict the **next day's closing price** of a selected stock based on its previous closing prices.

* **Objectives**
* Collect historical stock price data.
* Preprocess and clean the data.
* Apply a regression model (Linear Regression).
* Evaluate model performance using Mean Squared Error.
* Visualise predicted vs actual stock prices.
* **Literature Survey**

Various methods have been proposed for stock price prediction, including:

* Technical analysis
* Fundamental analysis
* Machine learning models like SVM, LSTM, ARIMA  
  Linear regression remains one of the simplest yet effective models for baseline predictions.
* **Tools & Technologies Used**
* **Python**: Programming language
* **Jupyter Notebook / Google Colab**: Development platform
* **Libraries**:

yfinance for fetching data

pandas and numpy for data handling

matplotlib for visualization

scikit-learn for ML models

* **Data Collection**

Data is collected using Yahoo Finance via the yfinance library. For example, we used:

import yfinance as yf

data = yf.download('AAPL', start='2020-01-01', end='2024-01-01')

* **Data Preprocessing**

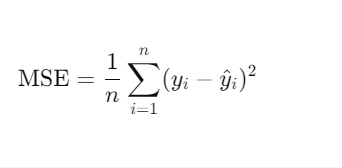
1. Only the ‘Close’ column is used for prediction.
2. The target is created by shifting the ‘Close’ price by one day.
3. Data is then split into features (X) and target (y).

* **Model Implementation**

We used the **Linear Regression** model from sklearn. The data is split into train and test sets (80:20). The model is trained and then predictions are made

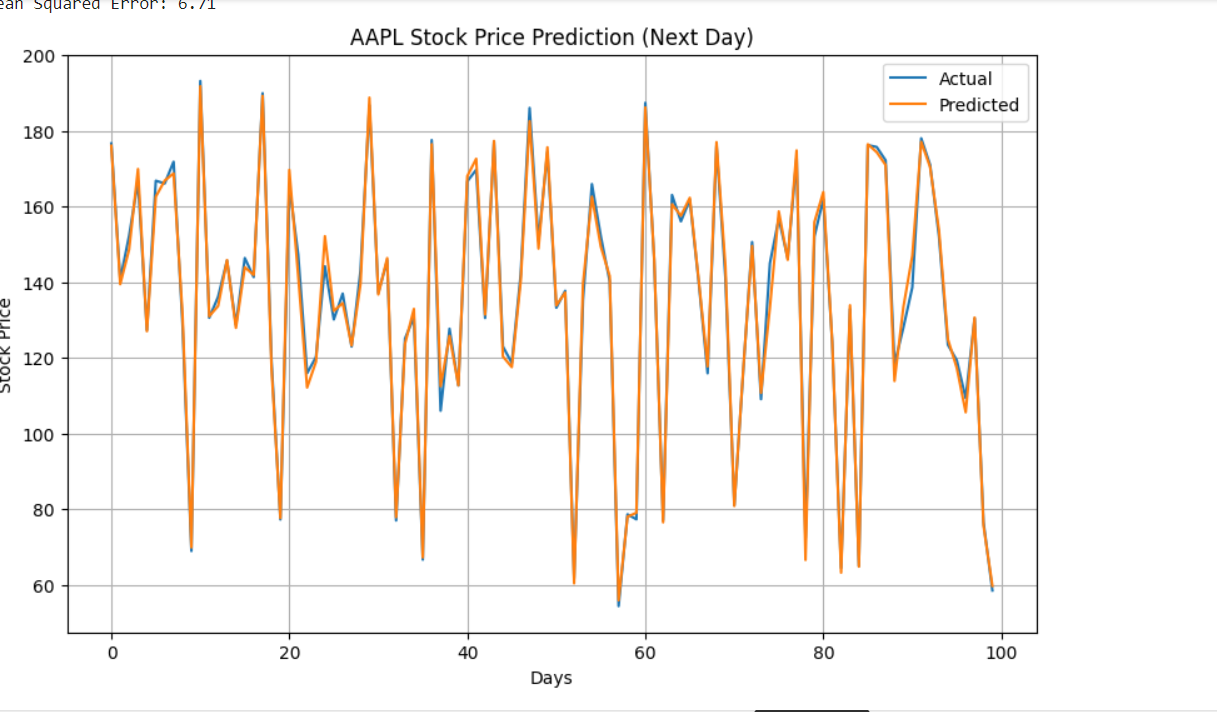
* **Evaluation Metrics**

We used **Mean Squared Error (MSE)** to evaluate model performance.



**. Results**

The model achieved a low MSE, showing that the predictions are reasonably close to actual prices. Below is the graph of predicted vs actual prices:



* **Limitations**

1. Only one feature (‘Close’ price) was used.
2. It predicts only one day ahead.
3. Linear Regression is too simple for complex financial patterns.

**Future Work**

1. Use more advanced models like LSTM, ARIMA.
2. Include more features like Volume, Open, High, Low.
3. Predict prices for multiple days.
4. **Conclusion**
5. This project demonstrates how machine learning can be applied to financial data for price forecasting. Though basic, the model can serve as a foundation for more complex predictive systems
6. **References**
7. Yahoo Finance
8. scikit-learn Documentation
9. yfinance GitHub
10. Hands-On Machine Learning with Scikit-Learn by Aurélien Géron