**Phase-2 Submission Template**

**Student Name:** [SANTHOSH KUMAR .S]

**Register Number:** [510623104093]

**Institution:** [C. Abdul Hakeem College of Engineering and Technology]

**Department:** [COMPUTER SCIENCE AND ENGINEERING]

**Date of Submission:** 09/05/2025

**Github Repository Link:** <https://github.com/Santhosh29032006/Cracking-the-market-code-with-Al-driven-stock-price-prediction-using-time-series-analysis>

# Problem Statement

The project aims to predict stock prices using AI techniques, addressing the need for accurate forecasting in financial markets.

**Type of Problem**:

Regression (predicting continuous values).

Relevance: Accurate stock price predictions can help investors make informed decisions, enhancing financial strategies and minimizing risks.

# 2. Project Objectives

**Objectives:**

Collect historical stock price data and relevant financial indicators.

Clean and preprocess the data for analysis.Build predictive models using machine learning algorithms.

Evaluate model performance based on accuracy and other metrics.

1. **Flowchart of the Project Workflow**

**DATASET**

**DATA PREPROCESSING**

**EDA**

**MODEL BUILDING**

**MODEL EVALUATION**

**DEPLOYMENT**

**CONCLUSION**

# Data Description.

**Dataset Name:** Historical Stock Price

**Dataset.Source**: Financial data APIs (e.g., Yahoo Finance, Alpha Vantage).

**Type of Data**: Time-series data (stock prices, volume, etc.).

**Records and Features:** Include features like date, open, high, low, close, and volume.

**Target Variable:** Future stock price

# 5. Data Preprocessing

Handle missing values through imputation or removal.

Remove duplicates and outliers.

Normalize or standardize features where necessary.

# 6. Exploratory Data Analysis (EDA)

Perform univariate and bivariate analysis to understand trends and relationships in the data.

Visualize distributions and correlations.

# 7. Feature Engineering

Create new features like moving averages, RSI, or other financial indicators.

Select relevant features based on their impact on stock price.

# 8. Model Building

**Models Selected:**

Linear Regression (baseline model).

Random Forest or LSTM (for capturing time-series patterns).

**Justification:** These models are suitable for regression tasks and can handle complex relationships

# 9. Visualization of Results & Model Insights

Include visualizations like:

**1.Prediction vs. Actual plots:**

This plot compares the predicted stock prices against the actual stock prices over a specific time period.

It helps in visually assessing how well the model performs.

**Description:** The x-axis represents the time (e.g., dates), while the y-axis shows the stock prices. The plot includes two lines: one for actual prices and one for predicted prices.

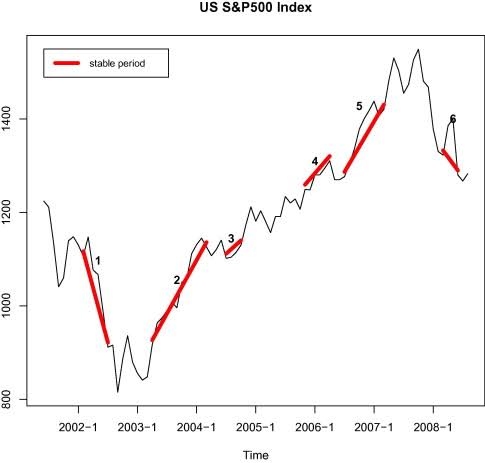
****

**2, Feature importance plots:**

This bar chart displays the importance of each feature used in the model. Understanding feature importance helps in identifying which factors most significantly influence stock price predictions.

**Description:**

The x-axis represents the importance score, and the y-axis lists the features. Higher scores indicate greater influence on the model's predictions.



# 10. Tools and Technologies Used

**Programming Language:** Python.IDE/

**Notebook:** Jupyter Notebook.

**Libraries:** pandas, numpy, scikit-learn, matplotlib, TensorFlow/Keras (for LSTM).

# 11. Team Members and Contributions

* Shawoor Saqib S.K: Project planning, data collection, coordination
* Santhosh Kumar: Data preprocessing, handling missing values, feature engineering
* Suresh V.: Exploratory Data Analysis (EDA) and visualization
* Vijay Kumar: Model building (LSTM, ARIMA, Prophet) and evaluation
* Sanjay P.: Model optimization and tuning
* Thanveerul Haq: Deployment, dashboard creation, final documentation