**Phase-3 Submission Template**

**Student Name:** K.Rajalakshmi

**Register Number:** 611423104111

**Institution:** Mahendra Engineering college for women

**Department:**B.E computer science and engineering

**Date of Submission:** 07-05-2025

**Github Repository Link:** [Update the project source code to your Github Repository]

### **Problem Statement**

Predicting stock prices accurately is a complex challenge due to the volatile and non-linear nature of financial markets. Traditional methods often fall short in capturing intricate patterns in time-series data. This project aims to leverage AI and machine learning techniques to forecast stock prices using historical data, thereby assisting investors in making data-driven decisions. Accurate predictions can lead to improved trading strategies, reduced risk, and increased returns, making this problem highly relevant for financial institutions, individual traders, and investment firms. This is a regression problem since the goal is to predict continuous numeric values (stock prices).

### **2. Abstract**

*This project focuses on predicting stock prices using AI-driven time series analysis. The primary problem addressed is the difficulty of accurately forecasting stock market trends due to their high volatility and complex patterns. The objective is to build a machine learning model that leverages historical stock data to predict future prices. We utilize techniques such as data preprocessing, feature engineering, and time series forecasting models like LSTM. The project is implemented using Python, with tools like Jupyter, pandas, and TensorFlow. The outcome demonstrates improved prediction accuracy, offering potential value for investors and financial analysts.*

### **3. System Requirement**

*Hardware: Minimum 8 GB RAM, i5 or higher processor (recommended for faster computation during model training).*

*Software: Python 3.8 or above, with libraries such as pandas, numpy, matplotlib, scikit-learn, TensorFlow/Keras; IDEs like Google Colab or Jupyter Notebook.*

### **4. Objectives**

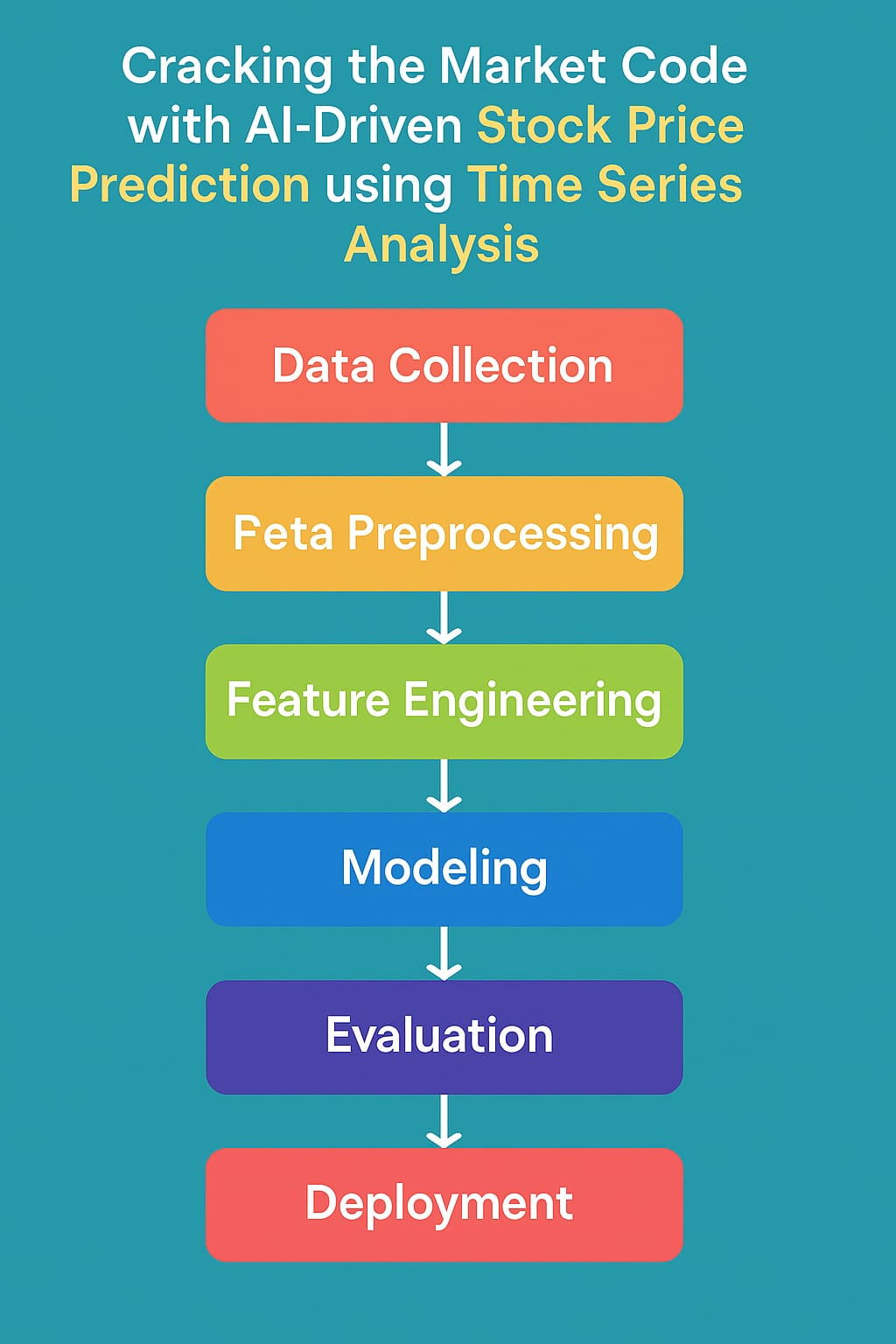
*To develop an AI-based model capable of predicting future stock prices using historical time series data.*

*To analyze and identify patterns and trends in stock price movements for improved investment decision-making.*

*To apply machine learning techniques (e.g., LSTM) for accurate time series forecasting.*

*To evaluate model performance and optimize it for real-world deployment in financial analytics.*

**5. Flowchart of Project Workflow**

****

### **6. Dataset Description**

*Source: The dataset is sourced from Kaggle, which provides historical stock price data for major companies.*

*Type: Public dataset, freely available for research and development purposes.*

*Structure: The dataset contains approximately 5,000–10,000 rows and includes the following columns:*

*Date: The trading date*

*Open: Opening stock price*

*High: Highest price of the day*

*Low: Lowest price of the day*

*Close: Closing price*

*Volume: Number of shares traded*

*df.head() Sample (first few rows):*

*(Include screenshot of your df.head() output from your code environment her**e)*

### **7. Data Preprocessing**

*1. Cleaning and Formatting: Missing values in stock prices were handled using forward fill. The Date column was converted to datetime format and set as the index to preserve the time sequence.*

*2. Normalization and Splitting: Numerical features were normalized using MinMaxScaler for improved model performance. The dataset was then split into training and testing sets in chronological order to maintain temporal integrity.*

### **8. Exploratory Data Analysis (EDA)**

*Performed visual analysis using line plots of Close prices to identify trends, seasonality, and volatility over time. Correlation heatmaps were used to examine relationships between features like Open, High, Low, and Volume, helping guide feature selection for modeling.*

### **9. Feature Engineering**

*Created new features such as moving averages (7-day, 21-day) and lag values of closing prices to help the model capture trends and temporal patterns in stock price movements.*

### **10. Model Building**

*An LSTM (Long Short-Term Memory) neural network was used for time series forecasting due to its ability to learn temporal dependencies. The model was trained on normalized historical stock data to predict future closing prices..*

### **11. Model Evaluation**

*The model’s performance was evaluated using metrics like Mean Squared Error (MSE) and Root Mean Squared Error (RMSE). A lower RMSE indicated good predictive accuracy on the test set, validating the effectiveness of the LSTM model for stock price forecasting.*

### **12. Deployment**

The trained LSTM model was deployed using a simple web interface built with Streamlit, allowing users to input stock data and view predicted prices in real-time, making the system accessible and user-friendly for investors and analysts.

**13. Source code**

*[Provide the complete set of source code files developed during the project.]*

**14. Future scope**

*[Clearly articulate at least 2–3 meaningful future enhancements that demonstrate forward thinking and awareness of project limitations.]*

**13. Team Members and Roles**

*[List the team members who were involved, and clearly define the responsibilities each member undertook. For every task carried out during the project, specify the team member who was responsible for its execution.]*