PROJECT TITLE: STOCK PRICE PREDICTION

Phase 1: Problem Definition and Design thinking

Problem Definition:

The problem is to build a predictive model that forecasts stock prices based on historical market data. The goal is to create a tool that assists investors in making well-informed decisions and optimizing their investment strategies. This project involves data collection, data preprocessing, feature engineering, model selection, training, and evaluation.

Design Thinking Approach:

1. Data collection:

- **Objective:** Gather historical market data to create a comprehensive dataset for stock price forecasting.

- **Actions:**
  - ➔ Identify and access sources of historical stock market data, such as financial APIs or data providers.
  - ➔ Retrieve data for selected stocks or indices, including features like date, open price, close price, volume, and other relevant indicator.
  - ➔ Set up a data collection pipeline to periodically update the dataset for real-time analysis.

2. Data Preprocessing:

- **Objective:** Prepare the collected data for analysis by addressing issues like missing values and outliers.

- **Actions:**
  - ➔ Clean the data by handling missing values through imputation or removal.
  - ➔ Identify and handle outliers that may affect the model's performance.
  - ➔ Convert categorical features into numerical representations using techniques like one-hot encoding or label encoding.

3. Feature Engineering:

- **Objective:** Create informative features that capture relevant patterns and trends in the data.

- **Actions:**
  - ➔ Generate lag features to capture historical price movements.
  - ➔ Calculate moving averages and technical indicators (e.g., RSI, MACD) to provide additional context.
  - ➔ Incorporate external factors, such as economic indicators or news sentiment, as features.

4. Model Selection:

- **Objective:**Choose an appropriate forecasting model that aligns with the characteristics of the data.

- **Actions:**
  - ➔ Evaluate various forecasting models such as ARIMA, LSTM, or Prophet.
  - ➔ Consider model assumptions, data stationarity, and computational requirements when making the selection.
  - ➔ Optimize model hyperparameters to improve accuracy and robustness.

5. Model Training:

- **Objective:** Train the selected model on the preprocessed data to learn from historical patterns.

- **Actions:**
  - ➔ Split the dataset into training, validation, and test sets to train and evaluate the model.
  - ➔ Implement a time series cross-validation strategy to account for temporal dependencies.
  - ➔ Train the model using appropriate algorithms and techniques for time series forecasting.

6. Evalution:

- **Objective:** Assess the performance of the predictive model to determine its accuracy and reliability.

- **Actions:**
  - ➔ Calculate forecasting metrics like Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE).
  - ➔ Visualize the model's predictions against actual stock prices using time series plots.
  - ➔ Compare the model's performance against baseline models or benchmarks.

Conclusion:

This document outlines the problem statement and design thinking approach for Stock price prediction data. All the above components contributes to the development of a robust prediction model for stock price forecasting.The objectives and actions will help to create a tool that assists investors in making informed decisions and optimizing their investment strategies.Continuous monitoring and adaptation to changing market conditions are also crucial for long-term success.