

STOCK MARKET ANALYSIS

1. Objectives

- **Goal:** Predict the closing price of a company's stock for a specific future date.
- **Scope:** Determine the range of dates and the accuracy requirements for the forecast.

2. Data Collection

- **Historical Data:** Gather historical stock price data, including closing prices, for the company in question. This data might also include open, high, low prices, volume, and other relevant metrics.
- **Sources:** Utilize financial databases and APIs such as Yahoo Finance, Here we selected NVDA company stock price data from Yahoo Finance

3. Data Preparation

- **Cleaning:** Handle missing values, remove duplicates, and filter out irrelevant data.
- **Feature Engineering:** Create additional features like moving averages, trading volume trends, and technical indicators.

4. Exploratory Data Analysis (EDA)

- **Data cleaning-**
 - **Checking Invalid Records:** Identify and rectify any invalid entries in the dataset.
 - **Missing Value Detection and Imputation:** Detect missing values and apply appropriate imputation methods to fill gaps.
 - **Duplicated Records:** Find and remove duplicate records to ensure data integrity.
 - **Handling Outliers:** Detect and handle outliers that may skew the analysis and model performance.
- **Numerical Data visualizations-**
 - **Pie Chart:** Visualize the distribution of categorical features such as book genres.
 - **Bar Graph:** Display the frequency of top authors or the number of books published each year.
 - **Histogram:** Show the distribution of numerical features such as book ratings.

- **Scatterplot:** Explore relationships between two numerical features, such as ratings and number of pages.
- **Heatmap:** Display correlations between different features in the dataset.
- **Boxplot:** Outlier detection using boxplot

5. Model Selection

Choose a forecasting model based on the complexity and requirements of your project:

- **Time Series Models:**
 - **ARIMA (AutoRegressive Integrated Moving Average):** Useful for univariate time series data.
 - **SARIMA (Seasonal ARIMA):** Extends ARIMA to account for seasonality.
 - **Exponential Smoothing:** Includes models like Holt-Winters for capturing trend and seasonality.
- **Machine Learning Models:**
 - **Linear Regression:** Simple model for predicting prices based on features.
 - **Decision Trees/Random Forests:** For handling non-linear relationships.
 - **Support Vector Machines (SVM):** Useful for regression tasks with complex relationships.
- **Deep Learning Models:**
 - **Recurrent Neural Networks (RNNs):** Including Long Short-Term Memory (LSTM) networks, which are effective for sequential data.
 - **Gated Recurrent Units (GRUs):** Similar to LSTMs but with fewer parameters.

6. Model Training

- **Split Data:** Divide data into training and testing sets. Consider using a rolling window approach for time series data.
- **Hyperparameter Tuning:** Optimize model parameters to improve performance.
- **Cross-Validation:** Use techniques like k-fold cross-validation to assess model performance.

7. Model Evaluation

- **Metrics:** Evaluate models using metrics like Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared.

8. Deployment

- **Streamlit App:** Develop a Streamlit application to deploy the Stock Market Analysis System.
- **User Interface:** Develop an interface (web or desktop) where users can input a date and receive a forecasted stock price.