CS661: Project Proposal Report

Interactive Stock Market Performance Analysis

GROUP-3

1. Introduction

The stock market is a dynamic and complex financial system, generating vast amounts of time-series data daily. Analyzing this data can provide valuable insights into market trends, individual stock performance, and the intricate interplay of economic factors and market sentiment. Our project aims to develop an interactive web-based data visualization system that offers insights into historical stock market performance, specifically focusing on the influence of macroeconomic conditions, market sentiment, and major crisis events. Using the Stock Market Dataset from Kaggle, we will analyze and visualize various statistics such as historical price movements, trading volumes, and the impact of external economic and sentiment drivers on stock behavior. The system will offer interactive charts and dashboards to help investors, financial analysts, and enthusiasts gain meaningful insights from historical stock data, particularly understanding how specific economic events and sentiments correlate with stock performance.

2. Data Source

The primary data source for this project will be the **Stock Market Dataset from Kaggle** (https://www.kaggle.com/datasets/ziya07/stock-market-dataset). This comprehensive dataset contains a wide range of historical stock prices and related information for each trading day. Key features in the dataset include:

- **Date:** Represents the trading day (YYYY-MM-DD format).
- **Open, High, Low, Close:** Simulated daily stock prices.
- **Volume:** Trading volume for each day.
- **Technical Indicators:** RSI, MACD, Bollinger Bands (Upper and Lower), and other calculated indicators.
- **Sentiment Score:** Captures sentiment derived from financial news and social media.
- **Macroeconomic Data:** Inflation rate and GDP growth, which influence market trends.
- **Target Variable:** Buy/Sell Signal (0/1) providing labels for stock trading decisions.

We will combine this data with information from **fred** api and **yahoo finance** api that will help combining the following

- **Macroeconomic Indicators:** Inflation rate and GDP growth, which influence market trends.
- **Sector Label:** Which sector our current stock belongs to.

This combined with the dataset is particularly suitable for:

- Analyzing the direct and indirect influence of macroeconomic conditions on stock performance.
- Investigating the correlation between market sentiment and stock price movements.
- Conducting event studies to understand stock reactions during periods of crisis.
- Visualizing relationships between various financial metrics and external economic factors.

3. Specific Tasks

For this project, we will perform the following main tasks:

- **Data Aggregation and Preprocessing:** Extract and compile historical stock data, macroeconomic data, and sentiment scores into a structured, time-series format, handling any missing values or inconsistencies to ensure data readiness for analysis.
- **Data Visualization Development:** Design and implement interactive visual representations that effectively highlight relationships between stock metrics, macroeconomic indicators, and sentiment scores.
- **Economic Insights Generation:** Apply analytical methods to uncover patterns related to the impact of inflation, GDP growth, and crisis events on stock behavior.
- **User Interaction Implementation:** Integrate interactive features allowing users to filter data by date ranges, examine specific economic conditions, and explore correlations between different variables.

4. Visualization Tasks and Key Questions to Answer

Our project will primarily focus on addressing the following five core questions, utilizing various visualization techniques to provide clear and insightful answers:

- **1. What is the observed relationship between Macroeconomic Data (like Inflation rate and GDP growth) and the overall performance of the stock?**
  - **Visualization Approach:** Time-series line charts comparing stock price trends (e.g., closing price or percentage change) against Inflation rate and GDP growth over extended periods. Correlation matrices or scatter plots to show the statistical relationship between these variables.
- **2. Can we identify specific macroeconomic conditions (e.g., periods of high inflation or low GDP growth) that coincide with notable changes in the stock's price or Sentiment Score?**
  - **Visualization Approach:** Highlighted regions on time-series charts (for price, sentiment, and macro data) corresponding to specific high/low Inflation rate or GDP growth periods. Event markers on charts indicating these conditions.
- **3. How did the stock's performance (price, volume, volatility) react during major identified crisis periods, such as the Covid-19 pandemic or significant geopolitical events (e.g., a specific war onset)?**
  - **Visualization Approach:** Annotated time-series charts showing stock prices, Volume, and Bollinger Bands (as a proxy for volatility) during and immediately following crisis periods. Comparative charts showing pre-crisis vs. post-crisis performance metrics.
- **4. Did the Sentiment Score and Macroeconomic Data show distinct patterns during these crisis periods, and how did they align with the stock's performance and Buy/Sell Signals?**
  - **Visualization Approach:** Multi-panel time-series plots displaying stock price, Sentiment Score, relevant Macroeconomic Data (e.g., GDP growth), and overlaid Buy/Sell Signals during crisis periods. This allows for visual inspection of their co-movement.

- **5. How does the Sentiment Score, derived from financial news and social media, correlate with the daily price movements of the stock?**
  - **Visualization Approach:** Scatter plots showing Sentiment Score versus daily stock price change. Line charts comparing the Sentiment Score trend directly with the stock's closing price or daily returns over time.

5. Overall Solution

Our proposed solution will be a web-based interactive dashboard that allows users to explore stock market statistics through various visual analytics techniques, with a strong emphasis on economic and sentiment-driven insights. The system will feature integrated charts and graphs, enabling users to examine historical stock data in depth and understand the complex relationships between market performance, macroeconomic conditions, and public sentiment. Key functionalities will include:

- **Interactive Filtering:** Users can filter data by specific date ranges to analyze different economic cycles or crisis periods.
- **Comparative Analysis:** Visual comparison of stock metrics against Macroeconomic Data and Sentiment Score.
- **Event-Driven Insights:** Clear visualization of stock reactions during significant economic and geopolitical events.
- **Signal Contextualization:** Display of Buy/Sell Signals in the context of prevailing economic and sentiment conditions.

The dashboard will be designed for intuitiveness and clarity, aiming to effectively present data-driven stories that offer a compelling narrative on financial market trends, the impact of external economic forces, and the role of sentiment in stock performance.

6. Tech Stack

- **Backend/Data Processing:** Pandas for data manipulation and analysis.
- **Frontend:** JavaScript React.js for building the interactive user interface; D3.js, Plotly, or Altair for sophisticated and interactive data visualizations.

7. Team Members & Responsibilities

Each member will contribute to different aspects of the project, leveraging their strengths, with a focus on integrating economic insights into the data analysis and visualization.

- **Assigned Roles:**
  - **Economic Analysis:**
    - 230625 Manya Dixit
    - 230848 Ravpreet Singh Ahluwalia

    Responsibilities: Focusing on interpreting economic trends, analyzing the impact of macroeconomic variables and sentiment on stock performance, and deriving insights for the visualization tasks.
  - **Data Processing:**
    - 230338 Daksh Shettar
    - 230644 Md Waqar Moid

    Responsibilities: Leading data cleaning, transformation, and aggregation.Data aggregation, and developing APIs for data retrieval by the frontend.

- **Visualization Development:**
  - 221087 Srisha Singh
  - 220248 Astitva Roy
  - 230975 Shivang Sonker

Responsibilities: Designing and implementing interactive charts and dashboards using selected visualization libraries (D3.js, Plotly, Altair). Ensuring visualizations effectively answer the key project questions.

- **Frontend Development:**
  - 230586 Kulshreshth Chikara
  - 221208 Vishap Raj

Responsibilities: Building the user interface, integrating visualizations from the visualization team, and ensuring a smooth user experience.