Team name: Finance Bull🐂

Name & USC-ID:

- Chu-Huan Huang, 1934208430
- Chung-Sheng Chang, 1676324151
- Jiaying Wang, 3018182967

Domain:

- Stock trading
- Finance
- Economics

Data collection

Stock Market Data

- **Data sources**:
  - Yahoo finance api (https://github.com/ranaroussi/yfinance)
  - Trading View (https://www.tradingview.com/)
- **Description**: Includes prices, trading volume, market capitalization, and other metrics of publicly traded companies.
- **Benefits**:
  - Helps identify trends and patterns for investment decisions.
  - Enables the creation of technical indicators for trading strategies.
  - Provides historical data for backtesting trading models.
  - Facilitates portfolio performance analysis and diversification.

Finance Data

- **Data sources**:
  - Company Websites (Nvidia: https://nvidianews.nvidia.com/news/nvidia-announces-financial-results-for-third-quarter-fiscal-2025)
- **Description**: Includes company-specific financial metrics such as revenue, profit, expenses, cash flow, and balance sheets.
- **Benefits**:
  - Offers insights into a company's financial health and performance.
  - Assists in valuation calculations (e.g., P/E ratio, discounted cash flow models).
  - Enables risk assessment and creditworthiness analysis.
  - Supports informed investment decisions based on fundamental analysis.

Economics Data

- **Data sources**:
    - Yahoo finance api (https://github.com/ranaroussi/yfinance)
    - FRED api (https://frb.readthedocs.io/en/latest/)
- **Description**: Covers macroeconomic indicators like GDP, unemployment rates, inflation, and interest rates.
- **Benefits**:
    - Provides a broader understanding of market conditions and cycles.
    - Aids in forecasting long-term economic trends and their impact on industries.
    - Guides central bank policy tracking and its influence on markets.
    - Helps investors align strategies with macroeconomic environments.

News Data

- **Data sources**:
    - Bloomberg (https://www.bloomberg.com/)
    - X api (https://docs.x.com/x-api/introduction)
- **Description**: Includes articles, press releases, and breaking news related to markets, industries, or companies.
- **Benefits**:
    - Provides real-time updates on events impacting markets (e.g., earnings reports, geopolitical events).
    - Helps investors react quickly to major announcements.
    - Facilitates the identification of emerging market trends and risks.
    - Offers contextual information to supplement data-driven models.

Sentiment Analysis Data

- **Data sources**::
    - Reddit Stock related Communities (https://www.reddit.com/r/StockMarket/)
    - X api (https://docs.x.com/x-api/introduction)
- **Description**: Extracts opinions, emotions, or sentiments from social media, news, forums, and other text-based sources.
- **Benefits**:
    - Gauges public and market sentiment toward a company, sector, or economy.
    - Identifies potential buy/sell signals from retail investor behavior.
    - Offers insights into market psychology during volatile periods.
    - Complements quantitative models with qualitative analysis.

Script Description

1. **Fetch Historical Stock Prices**

- Retrieves Nvidia's historical stock price data for a specified date range using the **Yahoo Finance API**.
- Saves the data to a CSV file in the `data/` directory for further analysis.

2. **Scrape Nvidia News**

- Scrapes financial news articles from Nvidia's official website.
- Saves the HTML content locally to `data/` for further reference.

3. **Extract Data from PDF**

- Extracts textual data from Nvidia-related PDFs (e.g., investor presentations) using the **pdfplumber** library.
- Stores the extracted data in a CSV file, including page numbers and text content.

4. **Collect Reddit Community Data**

- Utilizes the Reddit API to fetch the top 5 "hot" posts from the **Nvidia-related subreddit** (`r/NVDA_Stock`).
- Extracts information such as:
  - Post titles
  - URLs
  - Scores
  - Number of comments
  - Timestamps
- Saves the data in both JSON and CSV formats in the `data/` directory.

5. **Data Storage**

- All processed data is saved in the `data/` directory to maintain organization and accessibility.

Limitations of Chatbots in Analyzing Real-World Stock Markets

1. Lack of Real-Time Decision-Making

- **Reason**: Chatbots typically rely on pre-fetched or static data sources like APIs and do not perform real-time analysis at the speed required for high-frequency trading or dynamic decision-making.
- **Impact**: In fast-moving markets, even a slight delay in data retrieval and processing can lead to outdated insights or missed opportunities.
- **Solution**: Integration with real-time data streams and advanced predictive models would enhance the chatbot’s capability.

2. Limited Contextual Understanding of Macroeconomic Factors

- **Reason**: Chatbots often lack the ability to interpret and correlate macroeconomic indicators, global news, and geopolitical events with their potential impacts on stock markets. For example:
  - Understanding how interest rate changes affect stock sectors.
  - Correlating global conflicts or pandemics with market movements.

- **Impact**: Without this contextual awareness, chatbot predictions or insights may be overly simplistic or inaccurate.
- **Solution**: Integrating Natural Language Processing (NLP) models to analyze and summarize global news, combined with economic data processing, could improve the chatbot's contextual understanding.

Name & USC-ID:

Chung-Sheng Chang, 1676-3241-51