**Project Proposal: Stock Investment Advisor (SIA)**

**Goal**

**To create an interactive stock investment advisor that provides users with company information through a Q&A format.**

**Features include:**

* **Access to financial statements/reports of tech companies**
* **Key points from financial statements/reports**
* **Summarization of the latest related news with sentiment analysis on the news**
* **An organized report detailing company performance**
* **Insights on whether the company’s stock is growing or declining**

**Scope**

**The project will focus on the following elements:**

* **Coverage is limited to tech 50 companies**
* **Time Period for Report Summarization – 2023 & 2024 Year till Date**
* **Time Period for News Summarization - 12-months till query date**
* **APIs and News Sources – Maximum 2 each**

**Team**

**Chaeeon Lim:**

* **News Information Extraction**
* **Data Collection: Parsing financial reports, SEC filings, and earnings call transcripts**
* **Event Extraction: Identifying key metrics, events, and insights from text data**
* **Named Entity Recognition: Extracting named entities like company names, people, locations, etc.**
* **Document Classification**
* **Topic-specific Feature Engineering: Sorting news articles by topic (e.g. mergers, earnings, leadership changes)**
* **Text categorization: Classifying company filings by type**
* **Relevance scoring: Identifying relevant vs. irrelevant information**
* **Dialogue Assistant**
* **Financial summary Generation: Summarizing key financial metrics and performance**
* **Recommendation Generation: Generating portfolio reviews and market updates**
* **Natural Language Generation: Producing customized client communications**

**Uddesh Santosh Kumar Singh:**

* **Report Information Extraction**
* **Named Entity Recognition: Identifying the various metrics and entities mentioned in Financial Reports**
* **Information Retrieval: Find and retrieve structured information from the long reports**
* **Data Mining: Locate key words and metrics in unstructured data within the reports**
* **Sentiment Analysis**
* **Sentiment Classification: Classify text into predefined sentiment categories, such as positive, negative, or neutral**
* **Metric-Based Sentiment Understanding: Understand the emotions conveyed by the structured data. Example- Missed Target/ On Target/ Exceeded Target for net profits**
* **Dialogue Assistant**
* **Intent Recognition: Classify the user's intent, such as retrieving financial data, asking for stock performance, or requesting sentiment analysis of a news event**
* **Query Expansion and Paraphrasing: Rephrase user queries to improve the accuracy of results, and refine responses based on user input**
* **Response Generation: Generate coherent and informative responses to user queries, using data from financial reports, sentiment analysis, and other sources**

**Data sources**

* **Available Open-Source APIs for financial news**
* **Investing.com/Companies investor relations website**

**Examples**

* [**Sentiment Analysis,**](https://www.kaggle.com/code/indermohanbains/news-sentiment-analysis)[**public apis**](https://github.com/public-apis/public-apis?tab=readme-ov-file#news)
* **News APIs**
* [**Associated Press**](https://developer.ap.org/)**,** [**New York Times,**](https://developer.nytimes.com/) [**Bloomberg API**](https://www.bloomberg.com/professional/support/api-library/)
* **Financial APIs**
* [**News api,**](https://newsapi.org/pricing) [**Mediastack,**](https://mediastack.com/product) [**Marketaux**](https://www.marketaux.com/pricing)