A Major Project Synopsis on

NAVIGATING STOCKS SENTIMENTS

Submitted to Manipal University, Jaipur

Towards the partial fulfillment for the Award of the Degree of

MASTER OF COMPUTER APPLICATIONS

2023-2025

by

Yashvardhan Singh Nathawat

23FS20MCA00056

Under the guidance of

Dr. Pramod Kumar Soni

Department of Computer Applications

School of AIML, IoT & IS, CCE, DS and Computer Applications

Faculty of Science, Technology and Architecture

Manipal University Jaipur Jaipur,

Rajasthan

2025

Index :

Introduction :

The financial industry is undergoing a revolutionary transformation driven by advancements in technology, particularly through the integration of data analytics and artificial intelligence. In this evolving landscape, the ability to interpret vast amounts of unstructured data, such as news articles and social media posts, has become a crucial tool for investors. Understanding market sentiment — the collective mood of investors based on current events — has emerged as a key determinant of stock price movements and, ultimately, investment decisions.

The **"Navigating Stocks Sentiments"** project aims to provide a comprehensive solution for analyzing the impact of US news events on stock prices through real-time sentiment analysis. The project leverages cutting-edge Natural Language Processing (NLP) algorithms to process and understand financial news and social media content, extracting sentiment signals that directly influence market behavior. The goal is to enable investors to make more informed, data-driven decisions by providing them with real-time insights into market sentiment.

This project is divided into two primary phases. **Phase One** focuses on the ingestion of stock price data using the **Twelve Data API**, which provides both historical and real-time financial data. To facilitate seamless data flow, **Azure Logic Apps** and **Azure Event Hub** are utilized for automating data collection and ensuring real-time data streaming. This foundation ensures that stock prices are continuously updated and ready for analysis.

Phase Two of the project dives into sentiment analysis, utilizing the **MarketAux API** to collect real-time and historical news data. The platform employs a powerful combination of three state-of-the-art sentiment analysis tools: **finBERT**, **RoBERT**, and **ChatGPT**. **finBERT** specializes in financial text, **RoBERT** enhances understanding of broad linguistic patterns, and **ChatGPT** brings sophisticated language capabilities for complex sentiment interpretation. The output from these tools is used to generate a nuanced sentiment score that offers investors a deeper understanding of market trends.

By combining stock data with sentiment analysis, **Navigating Stocks Sentiments** provides investors with a unique lens to view the emotional undercurrent of the market, offering valuable insights that traditional financial models often overlook. This project not only aims to enhance

investment strategies but also contributes to the broader field of financial technology by demonstrating how AI and analytics can be integrated into real-time market analysis.

Objectives :

The **"Navigating Stocks Sentiments"** project aims to harness the power of advanced data analytics and Natural Language Processing (NLP) to provide investors with real-time insights into market sentiment and its impact on stock prices. The primary objectives of the project are as follows:

1. **Develop a Real-Time Sentiment Analysis Platform:**

   The core objective of this project is to build a robust platform that can analyze and interpret the emotional tone of news articles and social media content in real time. By capturing market sentiment based on current news events, the platform will empower investors to better understand how sentiment is influencing stock prices, thereby enabling more informed and data-driven decision-making.

2. **Automate Data Ingestion and Streamline Data Flow:**

   A critical component of the platform is the automation of data ingestion workflows. This objective focuses on integrating real-time stock price data from the **Twelve Data API** and news data from the **MarketAux API** into the system using **Azure Logic Apps** and **Azure Event Hub**. By automating the flow of data from external sources into the platform, the system ensures that investors have access to up-to-date and accurate information at all times.

3. **Integrate Advanced Sentiment Analysis Tools:**

   To extract meaningful insights from financial news and social media content, the project will integrate three state-of-the-art sentiment analysis tools: **finBERT**, **RoBERT**, and **ChatGPT**. Each of these tools brings unique strengths: **finBERT** is specialized in financial text, **RoBERT** excels in general language understanding, and **ChatGPT** offers sophisticated capabilities for interpreting complex language and sentiment. The goal is to generate highly accurate sentiment scores that reflect the market's mood and potential future trends.

4. **Provide Scalable and Efficient Data Processing:**

   Given the high volume of data involved in real-time market sentiment analysis, another key objective is to ensure that the platform is scalable and efficient. The project leverages **Microsoft Fabric** to process and analyze large datasets, ensuring that the platform can handle high-frequency data streams and scale effectively as data volumes grow. This scalability ensures that the platform remains fast, responsive, and capable of processing data from a wide range of sources.

5. **Enable Predictive Market Insights:**

   In addition to providing real-time sentiment analysis, the project seeks to lay the groundwork for predictive analytics. By correlating sentiment with stock price movements, the platform aims to develop predictive models that can forecast stock trends based on current sentiment. This predictive capability will allow investors to anticipate market shifts and develop proactive trading strategies, reducing uncertainty and risk in investment decisions.

6. **Enhance Investor Decision-Making:**

   Ultimately, the goal of the project is to empower investors by providing them with a more comprehensive view of the market. By combining real-time sentiment analysis with financial data, the platform will help investors gain a deeper understanding of the factors driving stock price movements. This holistic approach to market analysis will enable investors to make more informed, strategic decisions and improve their overall investment outcomes.
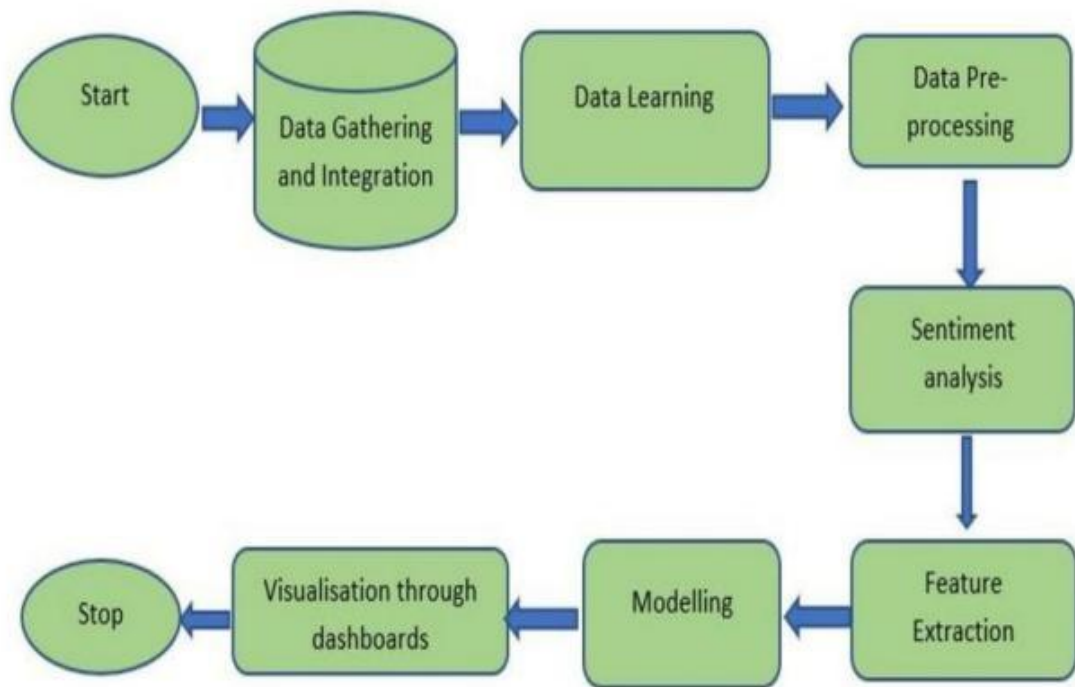
Fig: Workflow

Project Timeline:

The project is divided into two main phases, with each phase containing key milestones.

Phase 1: Data Ingestion (4 Weeks):

- **Week 1:** Integrate **Twelve Data API** for stock price retrieval and configure **Azure Logic Apps** and **Event Hub** for data ingestion workflows.

- **Week 2-3:** Automate workflows for real-time data fetching and integration into the platform.

- **Week 4:** Store ingested data in the **KQL database** for subsequent analysis.

Phase 2: Sentiment Analysis (6 Weeks):

- **Week 5-6:** Integrate **MarketAux API** for news data, and implement sentiment analysis using **finBERT**, **RoBERT**, and **ChatGPT**.

- **Week 7:** Develop user interfaces for displaying sentiment scores and insights from the analysis.

- **Week 8:** Conduct final testing, optimization, and debugging of the system.

Tools / Platform, Hardware, and Software Requirement Specifications:

Tools and Platforms:

- **Twelve Data API:** Provides access to financial data such as real-time stock prices.

- **MarketAux API:** Supplies real-time and historical market news.

- **Azure Logic Apps & Event Hub:** Automates data ingestion and manages real-time data streaming.

- **Microsoft Fabric:** Serves as the data processing and application development platform.

- **finBERT, RoBERT, ChatGPT:** NLP tools used for sentiment analysis of news and social media content.

Hardware Requirements:

- Standard workstations with a minimum of **16 GB RAM** and **500 GB SSD** storage for running sentiment analysis and processing data.

Software Requirements:

- **Python**: Programming language for data processing.

- **PySpark**: For handling large-scale data.

- **SQL Server Management Studio**: For querying the database.

- **Visual Studio Code**: For code development and integration.

References:

1. **Twelve Data API Documentation.** Accessed October 2024. https://twelvedata.com.

2. **MarketAux API Documentation.** Accessed October 2024. https://marketaux.com.

3. **Microsoft Fabric Documentation.** Accessed October 2024. https://microsoft.com/fabric.

4. **finBERT: A Pretrained Financial Sentiment Analysis Model.** (Li et al., 2020).

5. **RoBERT: A Pretrained Language Model for Financial Analysis.** (Zhang et al., 2021).

6. **OpenAI Documentation.** Accessed October 2024. https://openai.com