**Real-Time Stock Market Analysis Using Streaming Processing**.

* **Project Overview**

This project is all about creating a system that helps track and understand stock market data in real-time. Our goal is to build a tool that makes it easier to follow stock prices, analyze market trends, and get alerts when important changes happen. We also want to make the data easy to visualize so users can quickly grasp what’s going on in the market. The system is designed to give users useful insights that can help them make better decisions when it comes to trading or investing.

We focus on monitoring some key financial symbols like **AAPL** (Apple), **AMZN** (Amazon), **BINANCE**

(Bitcoin to US Dollar), **IC MARKETS:1** (forex index), and **MSFT** (Microsoft). These represent a mix of tech companies, cryptocurrency, and foreign exchange markets, giving users a wide view of different sectors. By keeping an eye on such diverse assets, the system provides a comprehensive snapshot of how various parts of the financial world are performing.

Beyond just stock data, we also factor in news. We know that news plays a big role in the way the stock market behaves. Whether it’s a company announcement, an economic report, or breaking world news, all these things can impact stock prices. To capture this, the system gathers market-related news and analyzes it for sentiment—essentially figuring out whether the news is positive, negative, or neutral. This helps give context to stock price movements, showing not only how prices are changing but also why they might be changing based on what’s happening in the world.

One of the most important aspects of the project is that it works in real-time. That means the system constantly pulls in fresh data and updates instantly. Users get live updates on stock prices and immediate alerts when there’s a significant shift, like a sudden price spike or drop. Having access to real-time data is crucial, especially for people who need to make fast decisions in the ever-changing stock market.

By combining stock data with news sentiment analysis, the system provides a more complete picture of what’s happening in the market. Users can see stock trends and understand the potential reasons behind them, based on news events. This added layer of analysis makes it easier to grasp the bigger story behind price movements.

In short, this project aims to give users a powerful tool for tracking and analyzing stock market data in real-time. It simplifies the process of following stock prices, while also offering insights from the latest news. With this system, users will be better equipped to make informed, data-driven decisions and stay up to date with what’s happening in the market.

**Architecture Diagram**

1. News Data Diagram

A diagram of a company's logo

Description automatically generated with medium confidence

1. Stock Data Diagram

A diagram of different logos

Description automatically generated

* **Data Sources**

**1. Stock Market Data:** Finnhub supplies real-time financial data, including stock prices, trade volumes, and market trends for symbols like **AAPL**, **AMZN**, **BINANCE**, **IC MARKETS:1**, and **MSFT**. This data allows us to track and analyze market movements in real time.

**2. Financial News Data:** Finnhub also delivers news related to the financial markets. We use this data to perform sentiment analysis, categorizing news as positive, negative, or neutral, which helps us understand how news impacts stock prices and market behavior.

* **Key Components**

Our project is built using several important components that work together to ensure real-time data collection, processing, analysis, and visualization. These components handle different aspects of the system, from fetching data to performing complex analyses and storing results for easy access.

**1. Data Collection**

The system collects real-time stock market and news data through the **Finnhub API**. This API provides both financial data, such as stock prices and trade volumes, and news articles, which are used for sentiment analysis.

**2. Data Processing and Analysis**

After data collection, the system processes stock prices and trends in real time to track market movements and generate alerts based on predefined conditions. It also processes news articles to perform sentiment analysis, categorizing them as positive, negative, or neutral to assess their impact on market trends.

**3. Data Storage**

The processed data, including real-time stock updates and sentiment analysis results, is stored for further use. This storage allows easy retrieval for visualizations and historical trend analysis.

**4. Visualization**

The final component is visualization, where the data is presented through charts. This allows users to see real-time stock movements, alerts, and the impact of news on the market, all in a clear format

Together, these key components enable the system to offer comprehensive, real-time insights into the financial market, making it easier for users to track and analyze stock performance and market sentiment.

* **Alert System**

The Alert System is an important part of our project that watches stock prices in real-time. Its main job is to notify users when a stock's price changes a lot, helping them make quick decisions.

Here’s how it works:

* **Keeping Track of Prices:** The system remembers the last price of each stock using a simple storage method. This makes it easy to access the most recent prices.
* **Checking for Price Changes:** Whenever new price information comes in; the system compares it to the last recorded price. It calculates how much the price has changed in percentage terms and checks if that change is more than 5%.
* **Creating Alerts:** If a stock's price change is greater than 5%, the system creates an alert. This alert includes details like the stock symbol, the old price, the new price, how much the price changed, and the time of the change.
* **Notifying Users:** The system then prints out alerts for any stocks that have had big price changes. This helps users see which stocks are moving quickly and decide if they want to act.
* **Challenges & Solutions**

During the project, we encountered a few challenges, particularly when it came to downloading and setting up the necessary tools and software. Some of these tools required specific configurations or had compatibility issues that slowed down our progress.

However, our team worked together effectively to address these challenges. We collaborated to troubleshoot the issues and share resources and solutions. This teamwork helped us quickly resolve any problems that arose.

* **Technologies Used**
* **Finnhub API:** Provides real-time stock market data and financial news for analysis.
* **Kafka:** Acts as a messaging system to efficiently stream data between components.
* **Flume:** Ingests news data into the system for further processing and sentiment analysis.
* **Spark:** Performs real-time data processing and analysis to identify trends and generate alerts.
* **Cassandra:** Stores large volumes of time-series data, ensuring fast access and scalability.
* **MongoDB:** Used for flexible storage of processed data and visualizations.
* **Python Scripts:** Automate data collection, processing, and alert generation, coordinating the entire workflow, Visualization.