DA331 - Big Data Analytics: Tools & Techniques

Project Proposal

Project Topic: **Real-Time Stock Market Analysis**

Team 23

Name: Sunanda K H

Roll Number : 210150018

Abstract:

The Real-Time Stock Market Analysis Project tries to showcase the application of technology in financial data analysis. This project utilizes real-time data in order to extract valuable insights into stock market trends.

The project revolves around using real-time stock market data, collected during market hours, to extract valuable insights on market trends. It includes the implementation of sentiment analysis techniques using data from Twitter feeds and financial news websites, which allows for more effective anticipation of market movements.

Sentiment analysis involves evaluating the emotional tone and sentiment expressed in financial news articles and social media discussions. This project aims to explore if this information can be used to enhance the precision of stock price predictions.

The core deliverable is an interactive dashboard. The dashboard provides an interface for accessing real-time data and sentiment analysis results, catering to a deeper understanding of the complexities of stock market analysis.

Introduction & Background:

The project revolves around the analysis of real-time stock market data, gathered during market hours, along with the sentiment analysis drawn from Twitter feeds and financial news websites. It offers an interactive dashboard to access this information.

Traditionally, financial analysis relied on historical data and mathematical models to make predictions. While this approach is often very valuable, it often lacks the immediate response to rapidly changing market conditions. Investors and traders need real-time information to make timely decisions in today's interconnected and fast-paced markets.

Real-time data analysis helps to make informed decisions based on the latest market developments.

News events, social media buzz, and public sentiment can significantly impact stock prices. By combining real-time data analysis with sentiment analysis, this project aims to aquire a better understanding of what drives stock prices.

Previous Investigations:

Prior research on real-time stock data analysis and sentiment prediction has laid a foundation for understanding the potential impact of data-driven insights in the financial sector.

Many studies have emphasized the significance of real-time data analysis in stock market decision-making. Various data analysis techniques, including time series analysis and machine learning algorithms, have been developed to extract meaningful patterns and trends from real-time stock data.

https://ieeexplore.ieee.org/document/9526506

https://scholarworks.sjsu.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1324&context=etd_projects

The above-mentioned papers and various other websites have focused on predicting stock prices using ML algorithms suck as RNN, LSTM (Long-Short term memory), Linear Regression, Random Forest and K- nearest neighbours.

While many projects focus solely on historical data, this project prioritizes real-time data. We collect and analyse data as it happens during market hours. This also aims to incorporate big data concepts to ensure scalability and performance, and also handle lage volumes of data efficiently.

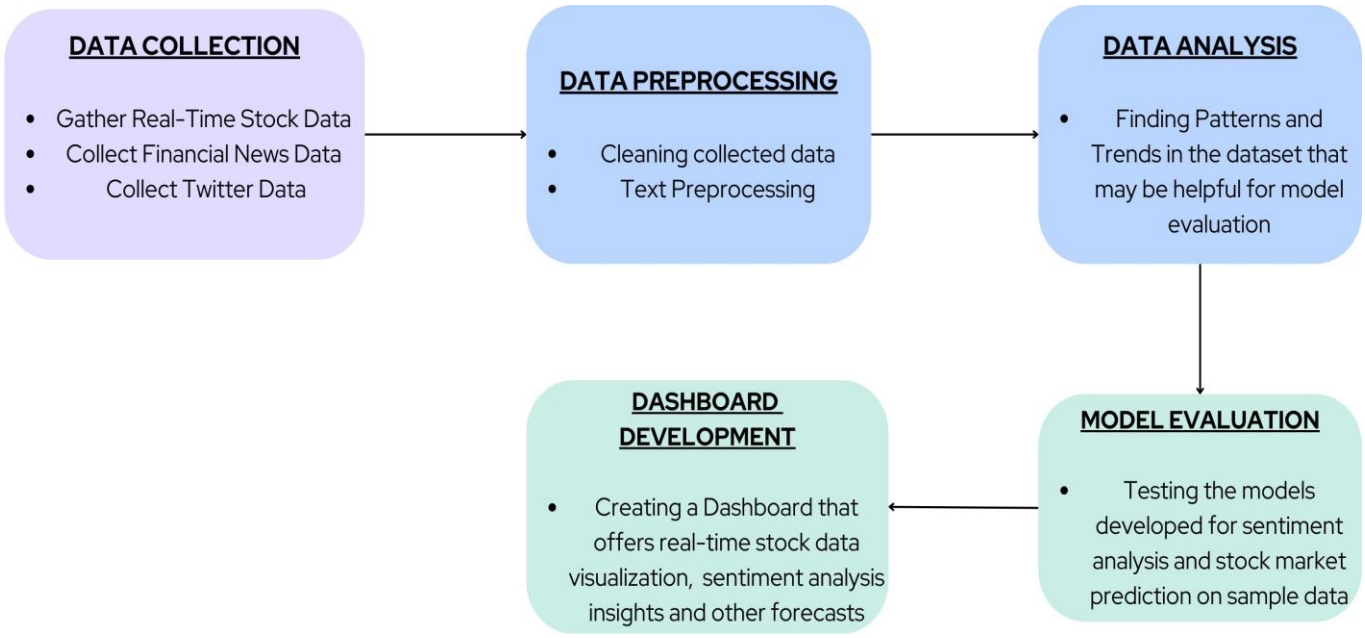
This project aims to build on the already existing techniques and incorporate sentiment analysis on the same to provide better insights of the market data through better visualizations that can provide a more comprehensive understanding of the market dynamics.

Proposed Project:

The proposed project aims to predict stock price movements based on real-time data and sentiment analysis. Specifically, this will focus on developing a predictive model that uses real-time stock data, financial news sentiment analysis, and historical price data to forecast stock price changes over various time intervals. This will explore whether sentiment analysis can enhance the accuracy of these predictions.

In broad terms, the stages of the project can be described as:

1. Collect real-time stock data from reliable sources.
2. Collect News data from reliable Financial News Websites.
3. Perform sentiment analysis on financial news data and Twitter feeds to gauge market sentiment.
4. Develop predictive models that consider a range of factors, including market data, sentiment scores, and historical trends.
5. Create an interactive dashboard to present real-time stock data and prediction results to users.



Some of the notable challenges that would be encountered during the course of the project are:

1. Issues with the data Quality:
    - Ensuring the accuracy and timeliness of real-time data can be challenging due to data source variations and latency issues.
2. Sentiment Analysis Accuracy:
    - Achieving accurate sentiment analysis results, especially in a real-time context.
3. Model Complexity:
    - Developing predictive models that effectively capture the variations and fluctuations of stock markets can be complex, and subsequently the developed model performance may vary.
4. User Interface Design:
    - Designing an intuitive and informative dashboard that meets user needs can be a design and usability challenge.

The dataset that would be initially used to train the developed model is given below:

Huge Stock Market Dataset (kaggle.com)