Gerald Fattah

Alvin Isaac

Max Savasta

Big Data Project Abstract

DataFinance Insights

In an era where financial markets are influenced by an ever-expanding volume of data, this project aims to harness the power of big data analytics to gain deeper insights into stock valuation and market sentiment. The primary idea revolves around the collection and storage of both quantitative and qualitative financial data. Quantitative data, readily available through financial institution APIs, forms the foundation of this project. Additionally, qualitative data, including news articles and sentiment analysis from social media platforms like Twitter, will be incorporated as an added complexity and storage challenge.

The project's data sources encompass a range of financial information. Quantitative data is obtained from established financial institution APIs, offering comprehensive insights into historical stock prices, trading volumes, and financial metrics. Qualitative data involves web scraping of news articles and social media content, followed by natural language processing (NLP) techniques to gauge market sentiment. These diverse data sources provide a holistic view of the financial landscape. The group plans to use primarily AWS tools to build this pipeline.

The challenges in this project encompass efficiently handling and storing large volumes of financial data, ensuring data quality and consistency across diverse sources, developing accurate natural language processing (NLP) models for sentiment analysis from news articles and social media, selecting appropriate machine learning algorithms for stock valuation, and ensuring model interpretability for transparent and trustworthy indicators in the dynamic financial market landscape.