Shuju (Jeffery) Sun
DSCI 510

Final Project Proposal: Stock Sentiment Extraction from News Headlines

Short Description

The project is focused on the rapid dissemination of financial news in the modern, internet-driven age, where information spreads within seconds. This immediacy contrasts sharply with the past, where it took days for such news to reach people through radio, newspapers, and word of mouth. The project notes that financial news articles are now often generated automatically from data like figures and earnings call streams.

This project utilizes data science techniques such as Natural Language Process, Data Extraction and Parsing and Data Visualization, to process this abundant information, aiming to generate investing insights. It's particularly relevant for hedge funds and independent traders who leverage such methods for profit. The project involves analyzing news headlines to assess market sentiment and make investment decisions with a comprehensive visualization dashboard.

Purpose (Question to Solve)

Due to the nature of the Project, the project generally conducts a sentiment analysis on the financial news related to five top stocks in the US and provides a detailed report on the analysis toward potential audiences like hedge funds and independent traders.

Data Source and Collection Method

The project intends to collect data through obtaining financial news, specifying Yahoo Finance and FinViz as the primary sources. By utilizing their APIs, the project aims to systematically collect financial news data, which is crucial for the analysis. APIs facilitate automated and efficient data retrieval, allowing the project to access the latest financial news and market insights. This method is not only time-efficient but also ensures that the data collection is consistent and can be updated or repeated as needed for ongoing analysis.

Yahoo Finance API: https://rapidapi.com/sparior/api/yahoo-finance15/
FinViz: https://github.com/lit26/finvizfinance/tree/master

Ideas about Analysis and Visualization

For the project, key data science techniques are employed for insightful investment analysis. Natural Language Processing (NLP) is used to analyze news headlines, determining their sentiment as positive, negative, or neutral, crucial for understanding market trends. Data extraction and parsing are essential for gathering relevant information from financial news sources, employing methods like web scraping or APIs. Lastly, the project utilizes data visualization and analysis to identify trends and patterns in news data, helping in making informed investment decisions. Specifically, intended visualization methods are involved with statistical graphs like Bar Chart and Stack Bar Charts. These techniques collectively enable the extraction of actionable insights from financial news, influencing investment strategies.

Shuju (Jeffery) Sun
DSCI 510

Data Collection Methods:
https://medium.datadriveninvestor.com/scraping-live-stock-fundamental-ratios-news-and-more-with-python-a716329e0493
https://www.octoparse.com/blog/how-to-scrape-yahoo-finance