CS 410: Project Proposal

Coordinator: Zijie Ku (zijieku2@illinois.edu)
Yanbin Zhang (zhang50@illinois.edu)

Abstract:

Many different sentiment analyses have been conducted to determine and discover the underlying relationship between news and stock price movements since the creation of social media platforms such as Facebook and Twitter. Many research papers have been devoted to addressing the advantage of using Twitter feeds to classify or predict the market movement because of its 5 V's: Volume, Velocity, Value, Variability, Veracity, and Variety. However, in recent years, the abuse of the characteristics of big data, such as the spread of fake news and unrestricted comment, has created a huge hinder to the sentiment analysis or classification by using social media data feeds. Our project is aimed at creating a sentiment analysis tool on stock prices movements from filtered news for short-term trading.

Description:

There are many proprietary AI-powered sentiment analysis tools, such as Bloomberg, Alpaca AI, and Wolfram Finance, available on the market. However, none of them are open-sourced or using public data sources. Our goal is to use open-source API and filtered news sources that are mentioned in the Sources, to create a sentiment analysis of stock movement classification or prediction. Unlike any past or existing project in CS410, our intention was not limited to scrapping unbiased* and professional comments from many financial analysts, but also extract meanings from their professional comments and prediction. These characteristics make this project a novel tool for both the market and the open source community regarding the functionalities it is expected to provide.

Functionality:

Our team would like to perform sentiment analysis on stock price based on historical and real-time filtered news feeds from the various reliable news sources. Our product will be made available and open source to the financial experts and individual investors in their algorithmic trading strategies.

Target Estimation:

We would like to achieve at least 68% accuracy for intra-day trading for the recommended stocks. Instead of crawling news from different search engines for stocks, we will be using filtered news which known to be related to the stocks and hence reduce the number of unrelated news in our analysis. This should significantly increase the accuracy of our analysis.

Minimum Goal:

Create a multi-level classification, (+1, 0, -1), to classify the movement of the stocks for short-term trading.

Preliminary Rough Timeline:

There are two months allocated for this project. A preliminary time contribution is distributed as following:

- At least 70-80% of the time will be devoted to the software implementation. (6-7 weeks)
- 10% of the time will be used to document the implementation and usage of the software and its APIs. (1 week)
- 10% of the time will be used to create a tutorial presentation on instructions and use case of the software. (1 week)

Sources:

Google Finance, Yahoo Finance, IEX, Alpaca Market, Seeking Alpha, Investing.

* Many existing news feeds about financial markets are proprietary and paid (potentially biased).