Sentiment Analysis on Stock Market using Twitter Data

Project proposal

Class: CSCE 5290 Section 002

Group #5

Team Members

Sweety Makwana

Gopi Krishna Sure

Gayatri Katukojwala

Nikhil Rangineni

GitHub Repository:

Motivation

Twitter is a very well-known place where traders tweet about stocks and financial instruments they care about. Their tweets reveal their sentiment about the stocks they are tweeting. Which can help decide overall trend for the particular stock if analyzed correctly. The stock market which is seen relatively high volatile in recent years could provide chances for making money or losing money. There are various factors that affects the direction of the stock market, or a single stock as follows:

- fundamental of the publicly traded company
- technical indicators
- news and media release for the company
- public sentiment - social media posts such as Twitter, Facebook, etc.

While the field of AI has grew tremendously in recent times, an AI engineer could take advantage of available AI technologies to help predict the trend of stock market by analyzing sentiment of traders from twitter.
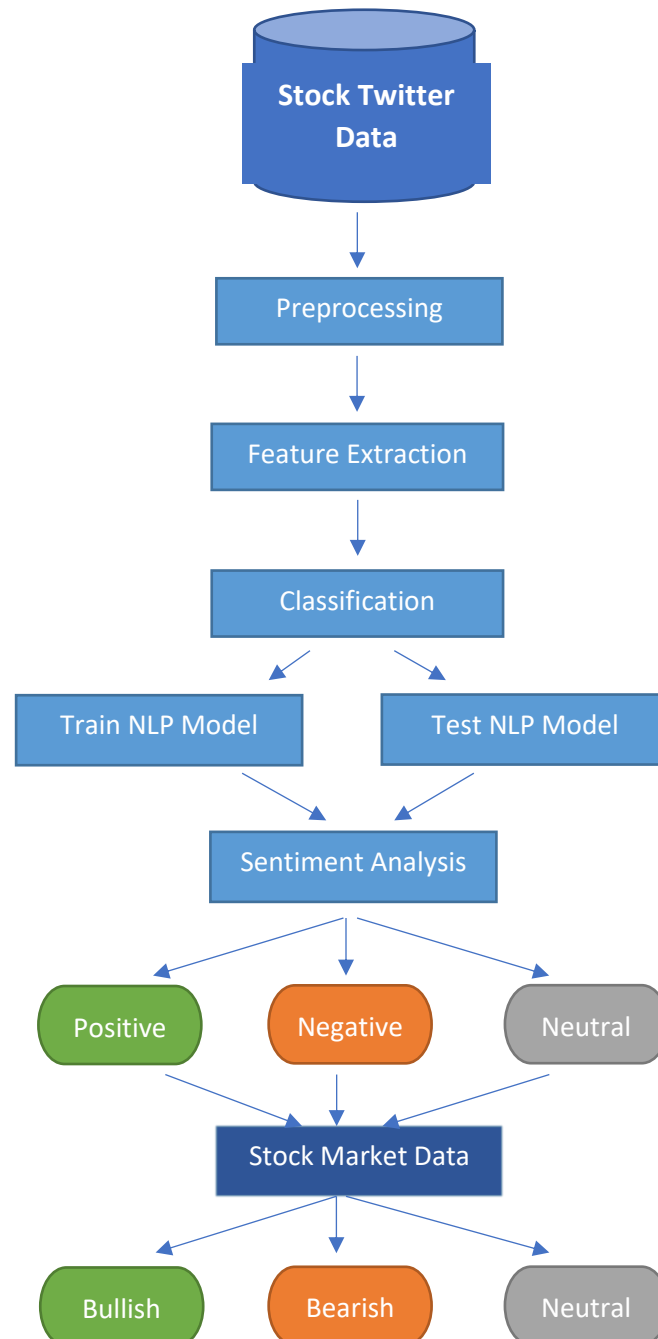
Significance

As Abraham Lincoln said, "With public sentiment nothing can fail. Without it, nothing can succeed." This also applies to the stock market. If public sentiment is positive about the stock market it means the stock market is expected to trend in a bullish manner, else if the public sentiment is negative, once could expect the stock market to trend in a bearish manner. If the sentiment is neutral, then stock market may not see much volatility and may trend neutral. Twitter data is linguistic and unstructured data which can be translated into structured and meaningful data using NLP models. Public sentiment could be classified into three categories as: positive, negative and neutral via text classification and text summarization tools available under the umbrella of Machine Learning field. Without the help of NLP model, the trader needs to read each and every tweet about the particular stock, analyze them manually and then decide the trend of the stock based on the sentiment score he/she comes up with. Not to forget here, the stock market data is constantly changing and so public sentiment about it. It is a very tedious process for anyone to manually keep track of all the tweets and analyze them constantly to help predict the stock prices.

Objectives

To overcome the problems faced by traders and help them decide the trend of a particular stock price using twitter data, our team would like to propose a project to build an NLP based model to derive public sentiment from tweets. Which would help traders to decide the trend of the stock and prevent financial losses. We would acquire pre-labeled Twitter data for a particular financial instrument such as, Apple (AAPL), Tesla (TSLA), Amazon (AMZN), or S&P500 index from Kaggle to train and validate the NLP model. We would obtain historical data for an

interested stock from Yahoo Finance to observe possible co-relation between the tweet sentiment analysis and stock prices. The historical data for stock would include daily price for Open, High, Low, Close, Volume and the date. We would yet to incorporate a type of the classifier to classify twitter data. We may integrate appropriate neural network such as long Short-term Memory (LSTM) or Self Organizing Fuzzy Neural Networks (SOFNN).

Workflow

Group #5

First, we would collect our datasets and visualize it. By using Google Colab platform, we would start implementing data preprocessing and building NLP model to derive sentiments via stock twitter data. Following train, test, and validation the NLP model should be able to detect sentiment of a particular stock or indexes and associated price prediction as such if the sentiment is positive, the stock is poised to be bullish; sentiment is negative, the stock is poised to be bearish; and if the sentiment is neutral, the stock would experience zero or low volatility in short to medium term.

Features:

- Visualize Twitter data
- Preprocess and classify Tweets using NLP model: Positive, Negative, Neutral
- Predict price for a particular stock using public (Tweet) sentiments
- Provide speed and scalability for financial decision to make profit and minimize losses
- Accuracy
- Visualize stock price chart along with Twitter Sentiment Polarity to predict the trend of the stock

Resources:

https://www.kaggle.com/

https://finance.yahoo.com/

https://data.world/crowdflower/apple-twitter-sentiment

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7959635/