Ilkogretim Online - Elementary Education Online, 2020; Vol 19 (Issue 3): pp. 4329-4334
http://ilkogretim-online.org
doi: 10.17051/ilkonline.2020.03.735584

Real-Time Stock Market Prediction Based On Social Sentiment Analysis Using Machine Learning

D. SATHISH KUMAR Assistant Professor Department of CSE Sri Sairam Engineering College

P. KARTHIK RAGAVUG Scholar Department of CSE Sri Sairam Engineering College

A.Arun Software DeveloperProdapt, Chennai

ABSTRACT

Machine learning area unit being employed in conjunction with data processing to solve real world problems. These techniques have evidenced to be extremely effective, yielding most accuracy with nominal quantity of investment and conjointly saving huge amount of time. To increase the annual income, People started looking towards stock investments as a More feasible possibility. With expert guidance and intelligent designing, we are able to nearly double our annual revenue through stock returns stock investment still remains a risky proposition for several people. With the help of Sentiment analysis on the tweets collected from Twitter API and with the closing values of stocks, We create a system which can forecast the stock price movement and we can build a web interface to access the Predicted Information. We seek to develop the system to use real-time data and use it to train the model and update the model using stock data apisor webs craping stock data and twitter apis.

KEYWORDS: Machine Learning, Stocks, Sentiment analysis, Web Scraping, client Interface

1. Introduction

Nowadays, social media has end up a reflect that displays people's thoughts and critiques to any unique event or information. Any tremendous or terrible sentiment of public associated with a particular corporation may have a spin off effect on the price of its shares. We are trying to find to predict the stock market prices of various companies by using sentiment analysis of the social media records which include tweets associated with business and web scrapping the news related to the business. This paper explains how tweets and financial news will assist you to end up a higher investor by making appropriate funding choices with the know-how of the market sentiment. First we will accumulate the tweets, web scrapped stock data and perform sentiment analysis corresponding to that time period we shall examine the stock market values from the past records and use a appropriate system for studying the set of rules to justify a legitimate correlation between social media sentiment and the stockvalues. Eventually with this data we are able to train our version and increase capability to make stock market predictions for future using these tweets and web scrapped

data. Considering the general public reactions to any primary event are to be almost right now on any social media, their temper can be captured fast and estimate the volatility in stock prices can be decided according to it. Since public responses to significant incidents are almost instantaneously visible on social media, their mood can be easily captured and an estimation of stock market volatility calculated.

2.RELATED WORK

[1] This paper used to show the effective use of sentiment analysis in the prediction of Stock Market Prices. Since the stock market is unpredicted no linear machine learning algorithm can predict the stock prices. We use sentiment analysis to predict the stock prices.

[2] The multilayer perceptr on and long short-term memory network are more appropriate for stock price prediction in this paper. This is the first attempt to predict the direction of stock prices using ETE, which can be easily applied in the real world. Using 55 stocks from 11 sectors and six cases of financial crises in the US financial market, we first find that the time- varying ETE based on the 3M and 6M moving windows has market explanatory power. Second, we find changes in the relative size of information flow in the market and the impact of sectors related to the financial crisis. Finally, using ETE network indicators as new features increases stock price direction prediction in all cases of the LR, MLP, RF, XGB, andLSTM models.

[3] This paper presents a novel approach to sentiment analysis for news events. A Word Emotion Association Network (WEAN) is developed based on social media data (i.e., words and emoticons) of a news event to collectively express its semantics and emotions, which lays the foundation for news event sentiment computation. To obtain the initial word emotions, a WEAN-based word emotion computing algorithm is proposed, which is then refined using a standard emotion thesaurus. It can calculate the sentiment of any sentence if it has the word emotions. A Word Emotion Association Network (WEAN) has been developed to collectively express its semantics and feelings, which is the basis for both word and text emotion computations, for word emotion computation via word sentiment association network. A word emotion computing algorithm based on WEAN has been proposed to obtain the initial word emotions through a built iterative process with proven convergence.

[4] Information is becoming the foundation for market decision-making; as a result, market entities expertise is being influenced on this foundation. This paper tries to show that this idea is reflected in social media as well. After conducting a review of the literature, it was determined what type of information makes up the various categories of knowledge and which social media tools are the best sources of such information. The information contained in various types of social media has been analysed and classified as knowledge for customers, knowledge from customers, and knowledge about the customer. The considerations are rounded out by a mention of the importance of customer knowledge in shaping an enterprise's operations strategy and consumer market decision- making.

3.PROPOSED METHODOLOGY

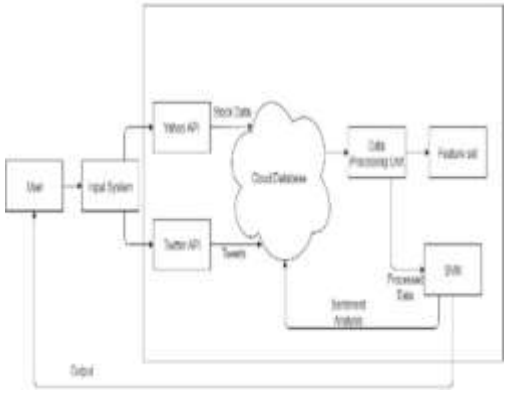

Fig.1. Architecture of Proposed system

3.1. Collection of Data

Twitter has a robust API for collecting tweets. There are two ways to collect tweets: using a streaming API or using a search API. We used search API to get around the limitations of the streaming API. The search API is a REST API that allows users to run queries on recent tweets. The search API allows for finer filtering of queries based on time, region, language, and other factors. The tweet and its metadata are included in the JSON object search. It contains a variety of data, such as username, time, place, and retweets. For the purposes of further research, we have concentrated on time and tweet email. The use of an API necessitates the use of an API key. We were able to login using the Tweedy Python library after authenticating with the key. Each tweet's text includes far too many unnecessary words that have little bearing on the sentiment. URLs, tags to others, and a slew of other marks of no sentiment value are all used in tweets. We must remove noise from the original state of a tweet in order to accurately determine its sentiment. The first step is to divide the text by space, resulting in a list of individual words per text, referred to as a list of words. To train our classifier, we'll use each word in the tweet as a function. After that, we delete stop words from the total word list. Stop words were removed using Python's Natural Language Toolkit (NLT) library. Posts, punctuation, and a few additional terms that have no sentiment value make up Stop words. A stop word dictionary compares each word in the tokenized word list to the dictionary. The word is filtered out if it is a stop word. Collecting the necessary data to Analyze the stock market and social media. So , Web Scraped data from Fin Viz Website for Financial and also collect Api data for Real-time Stock prices from Finn hub website. Web scrap the data and converted the headlines into pandas data frame according to dates and analyzed the Stock Price variation of various Companies Past years Graph.

3.2. Training the model

We developed and trained a classifier for tweet sentiment analysis after gathering a large tweet corpus. We concentrate on two classifiers in particular: Nave Bayes and Help Vector Machine. To identify on each classifier, we extract the same features from the tweets.It classifies the model by scoring each word in the training data and separating the best function. We used Python's Natural Language Toolkit (NLT) for ease of implementation, which helps us to measure with conditional frequency and frequency of each function. We measure the total number of available stock tweets for each company and construct a function matrix that includes positive, negative, neutral, and total tweets for each day. On the other hand, we used historical stock market data for each day to measure market up and down directions, which we used as a dataset label. Having a look at the historical stock market data for each day, measured market up and down path, and used it as a label for the dataset. reprocess the data and Visualize the data format. After Visualizing we Clean the data for model generation. After Cleaning the data, It is split into Training and Testing Data for Model Generation for Stock Data. Also Stock News is Semantically analyzed for using as a feature in model Training. Then the Model is created and Trained for the preprocessed Dataset.

3.3. Predicting the model

After training the model, We examine the relationship between tweet sentiment, web scapping the data from finviz website for financial and stock market prices on a daily basis. To do so, we gathered stock data as well as tweet data for the same timeline we Deploy the Trained model in the server using the Django Frameworks. We also create Business Logic Layer in Django for Server endpoints and Backend Databases using MySQL/Mongo DB.We create REST endpoints in the Business Logic Layer to access the functionalities of the Stock Prediction Model and concentrate on individual company stocks and collect daily data for each one.

3.4. Creating Interface

We create the Front End Interface using HTML, CSS, JS and Bootstrap. After Creating the Frontend Interface, we will connect the front end with business layer for completing the stack. After creating the frontend, we test the application for any errors using Django Testing Tools. After Testing we Deploy the Frontend Interface for final process.

4.RESULT AND DISCUSSION

We observed that using Social Media Sentiment Analysis with the Modern Machine Learning Techniques on Stock Market prediction gives a great Advantage over the Old Machine Learning Techniques without any Social Media Analysis. The Real-Time Usage of this method for Intraday Analysis brings more to the Common people use than Old Methods.

.

Fig.2 output Screenshot

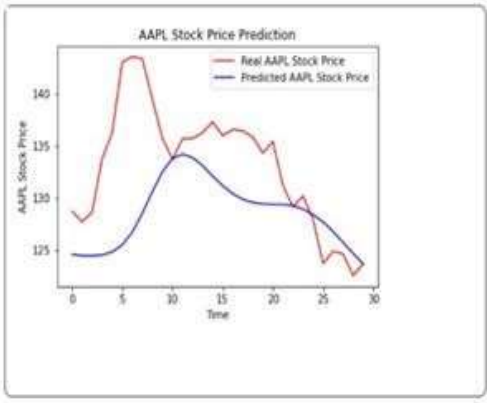

Fig. 3 Stock price Prediction analysis

Fig.4 Amazon Stock price Predictions

5.CONCLUSION

In this paper, Machine learning techniques are being employed with data processing to solve real world problems. These machine learning techniques are being extremely effective, yielding most accuracy with nominal quantity of investment and conjointly saving huge amount of time. With these techniques and expert guidance and intelligent designing, we are able to nearly double our annual revenue through stock returns stock investments till remainsa risky proposition for several people. Using Sentiment analysis on the tweets collected from Twitter API and with the closing values of stocks, We created a system which can forecast the stock price movement, getting the financial news by web srcapping and building a web interface to access the Predicted Information. We developed the system to use real-time data and use it to train the model and updating the model using stock dataapisor web scraping stock data. As opposed to a ground survey that would have been conducted otherwise to gauge public sentiment, using Machine Learning techniques for prediction is less costly.

6.REFERENCES

[1] Salah Bouktif, Ali Fiaz, Mamoun Awad, "Augmented Textual Features - Based Stock Market Prediction",IEEE, vol.08,no.10.1109/ACCESS.2020.2976725,February,2020

[2] Sondo Kim, Seungmo Ku, Woojin Chang, Jae Wook Song, "Predicting the Direction of US Stock Prices Using Effective Transfer Entropy and Machine Learning Techniques", IEEE, vol. 08, no. 10.1109/ACCESS.2020.3002174, June, 2020.

[3] Dandan Jiang, Xiangfeng Luo, Junyu Xuan, Zheng Xu, "Sentiment Computing for the News Event Based on the Social Media Big Data", IEEE, no. 10.1109/ACCESS.2016.2607218, 2016.

[4] Marta Zembik, "Social Media as a source of Knowledge for customers and enterprises", Online Journal of AppliedKnowledge Management, Vol. 2, Issue. 2, 2014.

[5] Srinath Ravi kumar, Prasad Saraf, "Prediction of Stock Prices using Machine Learning (Regression, Classification) Algorithms", International Conference for Emerging Technology(INCET), June, 2020.