Web sentiment analysis: Social media networking

Dr. A Pappu Rajan

Abstract
Social media networking sites are very important in this technological era and the comments, reviews played vital role to make business decision. In general analysis of fundamental and technical analysis can predict a stock to the certain extent but social media platform will reflect further prediction of stocks. In this study three industries such as FMCG industry, Telecom Industry, IT Industry are taken for the analysis. The financial data is extracted from the money control and social media reviews are extracted from Trackur tool. Sentiment analysis is used to extract such opinion and remarks of users by classifying them as positive, negative and neutral sentiment. Then the stock price changes can be compared. Pearson correlation has found and then regression model has built and researcher can find stock price movements and influences to social media in the same.

Keywords: Web sentiment analysis, social media analytics, big data

1. Introduction
Today Technology is introducing several platforms in different business verticals. One of the most important the platform is social networking. Opinion mining is to express people opinions and views about a particular topic such as news, movie, event and remarks related which is new way of understanding the perception of the public in machine learning. Hence sentiment analysis is used to extract such opinion and remarks of users by classifying them as positive, negative and neutral sentiment.

1.1 Social media
Social media is the collective of online communications channels dedicated to community-based input, interaction, content-sharing and collaboration.

1.2 Stock exchange
Most of the trading in the Indian stock market takes place on its two stock exchanges the Bombay Stock Exchange (BSE) and the National Stock Exchange (NSE). The BSE has been in existence since 1875. The NSE, on the other hand, was founded in 1992 and started trading in 1994. However, both exchanges follow the same trading mechanism, trading hours, settlement process, etc.

1.3 Market indices
The two prominent Indian market indexes are Sensex and Nifty. Sensex is the oldest market index for equities it includes shares of 30 firms listed on the BSE, which represent about 45% of the index’s free-float market capitalization. It was created in 1986 and provides time series data from April 1979, onward. Another index is the S&P CNX (CRISIL NSE Index) Nifty; it includes 50 shares listed on the NSE, which represent about 62% of its free-float market capitalization. It was created in 1996 and provides time series data from July 1990, onward.

1.4 Fast moving consumer goods
The fast moving consumer goods (FMCG) segment is the fourth largest sector in the Indian Economy. The market size of FMCG in India is estimated to grow from US$ 30 billion in 2011 to US$ 74 billion in 2018. Food products are the leading segment, accounting for 43 per cent of the overall market.
1.5 Sentiment analysis
It is a type of data mining that measures the inclination of people’s opinion through natural language processing, computational linguistics and text analysis, which are used to extract and analyze subjective information from the Web - mostly social media and similar sources.

1.6 Stock market
Stock market prediction is the act of trying to determine the future value of a company stock or other financial instrument traded on an exchange. The successful prediction of a stock's future price could yield significant profit.

2. Review of literature
Binoy B Nair et al. (2016)\(^1\) this article indicated that the forecasting technique using the sentiments expressed on social media and Artificial Neural Networks can indeed be successfully used for short-term forecasting of stock prices. Rupawari Jadhav and M. S. Wakode. (2017)\(^2\) They list out survey on Sentiment Analysis of Twitter Data for Stock Market Prediction. The Twitter is one of the most useful social media that can help to predict public mood. This paper discussed the different techniques for prediction of the future stock market with help of sentiment score. Ghaith Abdulsattar A et al. (2018)\(^3\) This article discussed about sentiment analysis on predict stock market behaviour based on consumer reactions. They also discussed the filtering, transforming, classifying the polarity of data.

3. Formulation of the research problem
Most stock price predictions have focused on news articles or blogs relating to companies rather than social media and also the prediction is targeted at the average market index not individual stocks. The prediction models are mostly done using Intra-Day minute stock quotes rather than end of day stocks. Hence there is a need to focus on individual stocks prediction using social networking sites or microblogging services. This research tweets are collected using tracker tool and the company data collected from money control. Applying the preprocessing techniques to remove spam and junk tweets. Using correction to find the relation among data. In this research is not including all MNCs data for time concern so convenience sample methods employed for this research to select company for data collection. Therefore researcher has concentrated on FMCG, Telecom and IT industries. This study deals analysing stock price movements of MNC’s using web sentiment analysis on social media.

3.1 Objective of the study
To analyze the stock price movements of MNC’s using sentiment analysis on Social Media. To find the correlation between stock price movements and social media sentiment data. To study about various techniques to predict stock movement.

3.2 Hypothesis of the study
H01: There is no correlation between FMCG sector stock price stock price movements and social media data.
H02: There is no correlation between Telecom sector stock price stock price movements and social media data.
H03: There is no correlation between IT sector stock price stock price movements and social media data.

3.3 Data collection
Money control: The historical stock prices are collected from the money control.com. Data is collected for the period of four months from June to September, 2016. Data consists of daily stock prices with high, low, open and close values are collected for all the companies.

Trackur is the social media analytics tool that offers powerful measurement metrics which help companies to develop better insights over the data. This tool will take conversation about the particular company that is conversed in social media and automatically scored positive, negative, or neutral results. This will also give the influence score for each and every person that might be used for analysis. Hence in this research Airtel, Infosys and HUL data have extracted from the trucker tool for the period of four months from June to September, 2016. Hence here researcher has not evaluated the sentiment score separately.

4. Data analysis
The Calculation is done using the formula: Stock Price Changes = -Open + Close
The total sentiment is calculated by the formula: Total Sentiment = Positive Sentiment + Negative Sentiment

<table>
<thead>
<tr>
<th>Model</th>
<th>R</th>
<th>R square</th>
<th>Adjusted R square</th>
<th>Std. error of the estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.507*</td>
<td>.257</td>
<td>.246</td>
<td>5.21615</td>
</tr>
</tbody>
</table>

The 4.1 gives the R square value for the Airtel sentiment. The value of R square is 0.257. It shows that there is a correlation accuracy of 25.7 %. The correlation accuracy is not effective for prediction but there is a considerable effect. Hence it is possible to predict some extent of stock price using the sentiment value. Hence the model built can predict stock price to some extent.

In the table 4.2, the unstandardized coefficients B has to be considered because it will be using to frame the regression equation. Hence to find the dependent variable, the values of B is used as follows Airtel stock = { -1.459 + [1.080 *Airtel Sentiment]} . So in this study it is found that the social media comments and reviews for Airtel will have the reflection on stock price movements. There is a positive correlation and hence in general it is said that the Telecom

---

\(^1\) Binoy B Nair et al. (2016)
\(^2\) Rupawari Jadhav and M. S. Wakode (2017)
\(^3\) Ghaith Abdulsattar A et al. (2018)
industries stock price movements are mostly affected by Social media sentiments. For this research it is found that the dependent variable Airtel stock has 25 % accuracy with the independent variable Airtel sentiment. The model is good enough to predict the future values having the 25% accuracy. It is not reasonably good for prediction but there is minimum level of accuracy as it can predict to some extent. This research the predictive model is built for the dependent variable Airtel stock and the independent variable Airtel sentiment. It gives the equation as Airtel stock = -1.459 + 1.080 (Airtel Sentiment). Based on this equation it is possible to predict future stock price movements with the accuracy of 25 %. Since the data for sentiment value calculation is only taken for 4 months, the accuracy and correlation value is not that much attractive. Here it is taken because of the availability of data. This can be avoided by taking one-year data which is said to be long term analysis and hence the accuracy can be improved. After eliminating the neutral sentiments in sentiment table, it is necessary to have a look at the comments that is mentioned as positives and negatives. Sometimes the positive comment is wrongly treated as a negative comment.

5. Conclusion
Investment decisions play a vital role in this changing economic condition. The previous analysis of fundamental and technical analysis can predict a stock to the certain extent but social media platform will reflect further prediction of stocks. It is not necessary to consider this tool and there are many ways to extract data from the social media’s that will give better result. For future research can be done by many methodologies like Support Vector Machine, Neural Networks and Decision trees for the purpose of prediction and these type of advanced classifiers will give more accuracy prediction models compared to regression model.

6. References