



International Journal of Engineering, Business and Enterprise Applications (IJEBA)

www.iasir.net

A Hybrid Model for Stock Market Trend Analysis

Ch. Vanipriya¹, Thammi Reddy.K²

¹Department of Information Science and Engineering,
Sir M Visvesvaraya Institute of Technology, Yelahanka, Bangalore, Karnataka, INDIA

²Department of Computer Science & Engineering,
GIT, GITAM University, Vishakhapatnam, Andhra Pradesh, INDIA

Abstract: Stock Market is often viewed as a way to grow the wealth for individuals and companies. Which is the company to invest in, by analyzing the company performance needs a greater insight. Many researchers have attempted to analyze the stock market and to predict the stock price by using some statistical and non-statistical methods such as analyzing news articles using sentiment analysis. In this paper we propose another novel method which combines these two methods and analyze the performance of some of the selected companies listed in BSE. For Technical analysis the moving average indicator, and for sentiment analysis NBC classifiers have been chosen for the proposed work.

Keywords: Sentiment analysis, Technical analysis, Moving Average, NBC.

I. INTRODUCTION

The stock market gives an impression of a legalized gambling for the first time investor. In India, Bombay Stock Exchange (BSE) is the stock exchange where most of the trading takes place. The BSE has been around since 1875. The exchange is home to about 5,300 listed companies as of October 2013. Internet was introduced to India in the early 90's and the use started to increase exponentially. It has the world's third-largest Internet user-base with over 137 million as of June 2012. The stock markets introduced Internet trading (online-trading) in February 2002.

People want to invest their money in stock market and expect high returns in short period of time. They used to analyze the technical predictors to evaluate the performance of the company. Now the online media is playing an important role in the stock market. There are evidences [1] that clearly show that the news articles published through the online media, regarding companies have greater impact on the stock price. A small issue may alarm a great impact on the people's money. By tracking the news about a particular financial firm on various online media, the stock price of that may be evaluated.

Some of the new investors don't have any idea of how the future of the stock market will be like. It's like a show of hands; invest in, any one of the randomly chosen stock, if its price goes up, you are lucky enough or else, better luck next time! But the truth is, it doesn't work so! The astrology of predicting the future stock prices is one of the most trending topics in many fields including trading, finance, statistics and computer science. The motivation for which is naturally to predict the direction of future prices such that stocks can be bought and sold at profitable positions. Stock price and Share price changes depend on many factors like market forces, supply and demand etc. If the demand for a particular stock is more than what is supplied, then the price of the stock goes up, or if it is otherwise, that is, supply is more than demand, the price of the stock goes down!

Most of the investors are great believers! Some believe that it's impossible to predict the stock price while some others say it is completely predictable with a little bit of graphical analysis and a few calculations based on the past. While the latter kind of investors is right in their beliefs, but practically, it's quite more than that. The sentiments, attitude and expectations of the investors also play a major role in predicting the stock price.

The field of Computer Science has the most effective solution to this. The stock prediction can be most efficiently undertaken by combining tools like sentiment analysis and other technical analysis! With its unbelievable popularity, the stock market trending will then become clear and understandable as it brings decent yields with little or no efforts. The stock market is all about dynamics and that is why it is important to accurately forecast further movements of stock bids.

In this paper as mentioned earlier the techniques of Sentiment analysis and Technical analysis were combined to analyze the performance of the chosen company, so that they can take a decision to invest in that company or not

II. RELATED WORK

A pretty good number of authors have already attempted to apply the sentiment analysis technique to predict stock prices, while a few others have used technical analysis for the same. Hence, the combination of the both types of analysis is more likely to give an accurate result, on which the real time investors can rely on!

Masoud Makrehchi *et al.* [2] proposed a novel approach to label the text messages in social media using the major stock market events. They aggregated the net sentiment of a day and proved its predictive power of stock market movement.

Johan Bollen *et al.* [3] published a paper “Twitter Mood Predicts the Stock Market” in the year 2010. They analyzed tweets in the social media website called twitter.

Enric Junque de Fortuny *et al.* [4] discussed about appropriate metrics to evaluate models and gave an insight into evaluating, validating and refining the models for predicting the stock market.

In the work done by [5], they examined the stock message boards for cross-sectional and time-series determinants of message-postings. He showed that variation in message-posting volume is related to underlying firm characteristics and stock market activity. He used a sample of over 3,000 stocks listed on Yahoo! message boards. [6] Developed a system called E-Analyst which collects two types of data, the financial time series and time stamp news stories. It generates trend from time series and align them with relevant news stories and build language models for trend type. In their work they treated the news articles as bag of words.

III. TECHNICAL ANALYSIS

The History of the stock prices that can be used to predict the future forms the basis of technical analysis. Technical Analysis is predicting the future financial price movements by taking into account, its previous movements. With analogy to the weather forecasting, the technical analysis too does not give the absolute predictions about the future. Instead, it can help investors anticipate what is the probability of a certain thing to happen to the prices over time.

Basically, the technical analysis helps the investors to decide on the best time to invest and the best time to sell in order to get the maximum profit. It completely relies on the study of historical price fluctuations. People who do technical analysis before investing in stock market usually study the price charts for price patterns and use price data in different calculations to forecast future price movements.

The technical analysis makes use of the correlation between the price and the company and thus gives the investor a fair chance to decide on the best time to enter/exit the market. The company's fundamentals are not at all a concern for the technical analysts. Technical analysis is variedly applicable to stocks, indices, commodities, futures or any tradable instrument.

A major drawback of technical analysis is that it just considers price forecasts, ignoring the fundamental factors of the company. However, technical analysis assumes that, at any given time, a stock's price reflects everything about that company which necessarily includes fundamental factors.

A. BASIS FOR TECHNICAL ANALYSIS

The Dow Theory laid the foundation for the modern technical analysis. There are three main theorems of the Dow Theory.

i) Price Discounts everything. The technical analysts believe that all the information can be obtained from the current price and it forms the basis for analysis. After all the market price reflects the sum knowledge of all the participants which includes traders, investors, portfolio managers, analysts and the others. Technical analysis utilizes the information captured from the price to analyze what the market is saying with a purpose of forming a view on the future

ii) Price movements are not totally random. Most of the Technical analysts agree that prices trend. Some Technical analysts also acknowledge that there are times when the prices don't trend, in which case it would be difficult to gain profit using Technical Analysis.

iii) “What” is more important “Why”. There are two things that Technical analysts are concerned of.

a) What is the current price?

b) What is the history of the price movement?

The price is the end result of the battle between the forces of supply and demand for the company's stock. The technical analysts believe it is best to concentrate on “what” and never mind “why” since the price goes up when there are more buyers than sellers.

B. TECHNICAL ANALYSIS INDICATORS

There are many Technical indicators, some of them are Trending Indicators, Momentum Indicators, Volatility Indicators, Strength and Sentiment Indicators, Stock Market Indicators, Moving average indicators. Among these “Moving average indicators” was chosen for the proposed work.

The moving average indicator gives the average value of a Stock's price over a period of time. The moving average indicator can be calculated by considering mathematical analysis of a stock's average value over a predetermined time period. It is a trend following or lagging indicator because it is based on the past prices.

The 200- day moving average is a popular technical indicator which the investors use to analyze the price trends. It is nothing but the stock's average closing price over the last 200 days (i.e. (stock price of day 1 + stock price of day 2 up to stock price of day 200)/200).

$$MA = \sum_{i=1}^{200} SP_i / 200$$

Where MA=Moving Average and SP=Stock Price

It says about the market trend, by looking at the previous 200 trading days and taking the average of closing prices of 199 previous trading days.

$$YAP = (CP \text{ of yesterday} + CP \text{ of 199 PTD})/200$$

Where YAP=Yesterday's Average Price, CP=Closing Price and PTD=Previous Closing Days

Mark that point on a chart and make another dot by considering the closing price of yesterday. Again calculate today's average price by dropping a day off backend

$$TAP = (CP \text{ of today} + CP \text{ of 199 PTD})/200$$

Where TAP=Today's Average Price

Again mark that point on a chart and make another dot by considering the closing price of today. This can be repeated every day and both the values can be plotted. It shows the direction of the market and measures the trend. Extremely high readings are a warning that the market may soon reverse to the down side whereas very low readings signify the reverse. The shorter the moving average, the sooner there will be a change in the market. It is used to determine whether the stock is technically healthy or not. The percentage of stocks above the 200- day moving average helps to determine the overall health of the market.

IV. SENTIMENT ANALYSIS

Sentiment analysis also referred to as Opinion Mining is the use of NLP tools, text analysis, machine learning and computational linguistics to identify and extract the sentiment in the source materials which can further be used for various kind of analysis. It is a general binary opposition in opinions. It makes up a great tracker for all kinds of data; a few noteworthy are products, brands, people etc. It usually does the tracking process based on a few keywords and their frequencies in a particular dataset.

It's understood from the previous work that the news items of a particular company have significant impact on stock indices and prices of that company. The people go through the news articles regarding the company before selling or buying a stock of that company. A lot of work was done to extract the sentiment from the static text using Text Mining and NLP techniques. We can analyze news items for sentiments using dynamic data sources such as online news stories and streaming data such as blogs.

Manually going through all the news items and analyzing them is a tedious task. Sentiment Analysis can be used to automatically identify the opinion or sentiments about a company in news article. The sentiments can be positive, negative or neutral. The advantage of this automated tool is its ability to evaluate large quantities of news articles or posts without manual intervention. Earlier this tool was used for analyzing the product from various review sites or discussion forums. Most recently, this tool is being used to measure the sentiment of news articles in the stock market domain.

V. PROPOSED SYSTEM ARCHITECTURE

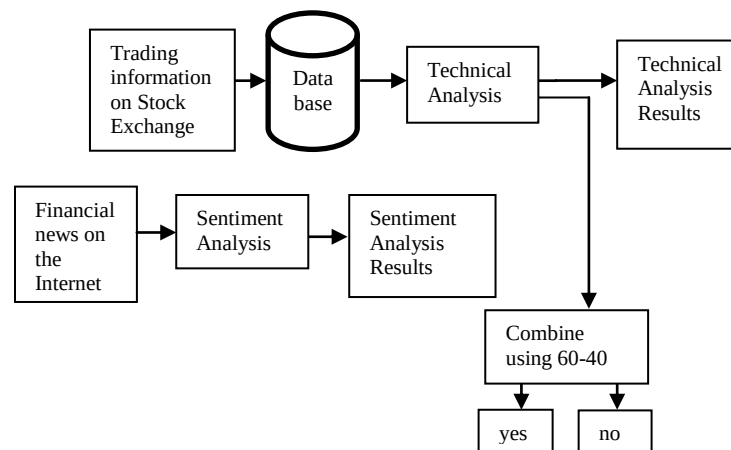


Fig. 1: Proposed System Architecture

The architecture consists of two major modules, one for technical analysis and the other one for sentiment analysis.

Algorithm: Stock market analysis by using technical and sentiment analysis

1. Choose the company.
 2. Extract the stock prices of a period of one year of that company.
 3. Calculate moving average (m) for a day which is
Average value of previous 200 days and compare with
that day's closing price (cp).
if ($m > cp$)
then "it is Uptrend"
else
"it is Downtrend."
 4. Extract the news articles of that company for a period of time.
 5. Manually tag some of these news articles as positive, Negative or neutral.
 6. Train the classifier on these set of data
 7. Predict the sentiment of news articles.
 8. Extract the percentage of good vs bad.
 9. Combine Technical and sentiment analysis to analyze the company trend
-

STEPS INVOLVED IN TECHNICAL ANALYSIS**A) Extract Stock Values**

The historical stock values are extracted for a period of one year from a website in.yahoo.finance.com. The data extracted is in a CSV format. This data is then placed into a database table which can then be queried and processed.

B) Processing the stock data

The 200- day moving average is calculated using a query. The closing price value on the 200th day is then compared with the moving average value.

If it is greater than the latter then it signifies an uptrend. The count of the number of times this occurs is taken. This serves as a parameter to gauge the performance of the selected company. The percentage of time that this occurs is also calculated.

STEPS INVOLVED IN SENTIMENT ANALYSIS**A) Website Reviews**

The initial step in sentiment analysis is the identification of the website containing the news articles of the chosen company. In our work we are using "moneycontrol.com" to extract only the news articles that will be analyzed. Once the required site is chosen, we then begin the process of extraction which includes the extraction of only the news articles over a period of 6 months while ignoring the various ads and other unnecessary contents from the link. Only the link to the news articles are taken into consideration and stored in a file.

B) Extract Reviews

The essential data, i.e., the links to news articles obtained from the previous step is then subjected to extraction and the content of each of the links in the said file is extracted and stored in a separate file. This file containing the articles will serve as an input to the sentiment analysis code.

C) Analysis of each review

In this step the sentiment of a review is predicted. The Naïve Bayes Classifier is used for this purpose. A training set is created which contains of manually tagged articles in a Comma Separated Value (CSV) format. The classifier is trained with this set. After training, the file with the extracted reviews is passed as an input to the classifier which then predicts the sentiment of each article. The final output is a percentage of positive and negative articles which is represented graphically in pie chart.

D) Calculation and Result

The result of the previous 2 analyses has to be combined. The 2 percentages, i.e., the percentage of time that the closing price is higher than the 200- day moving average and the percentage of the positive news articles are combined in a 60-40 ratio to give the final percentage. The final percentage that is obtained from the previous step is then checked to give the final outlook for the company. Depending on the percentage we can say if the company has been consistently seeing an uptrend and the news articles pertaining to the company are on the positive side. If so, then it is safe to invest in the company as there is a positive outlook for the company. On the other hand, if the percentage signifies that the company has been seeing a downtrend, then it can be concluded

that there is a negative outlook and there is a risk in investing in the company. This helps the user make an informed decision while investing in a particular company.

VI. EXPERIMENTS AND RESULTS

The stock prices for the selected company are extracted from the website in.yahoo.finance.com for a period of 1 year. This data is stored in a database and the 200-day moving average is calculated and compared with the closing price. A count of the number of times this occurs serves as a parameter to determine the performance of the selected company. A graph is also plotted using Google Charts to show the variation of 200-day moving average and closing price over the said period of time.

The graph shows the variation of 200-day moving average and the closing price over time. It shows the pattern of the price movement. If the closing price is greater than the 200-day moving average, it is regarded as an uptrend.



Fig 2: Results of Technical Analysis

This screenshot shows the results of the Technical Analysis for the selected company, which in this case, is HCL Technologies. The graph shows the variation of the 200-day moving average and the closing price over a period of one year.

URLs for the news articles for the selected company were extracted from the website www.moneycontrol.com and stored in a file. This file is then used to open each URL separately and extract the news article which it contains. The URLs that do not contain news articles are filtered and only the relevant URLs are written into the file. The end result is file that contains all the articles specified by the collected URLs and this file serves as an input to the Sentiment Analysis module which does the Sentiment Analysis of the news articles pertaining to a particular company. A training set in the Comma Separated Value (CSV) form is also separately created with manually tagged articles which serve as an input to this module. The Naïve Bayes Classifier (NBC) is first trained with this training set. Then the file with the extracted articles is passed to the NBC for test. The percentage of articles that are positive and those that are negative are calculated and also represented in the form of a pie chart.

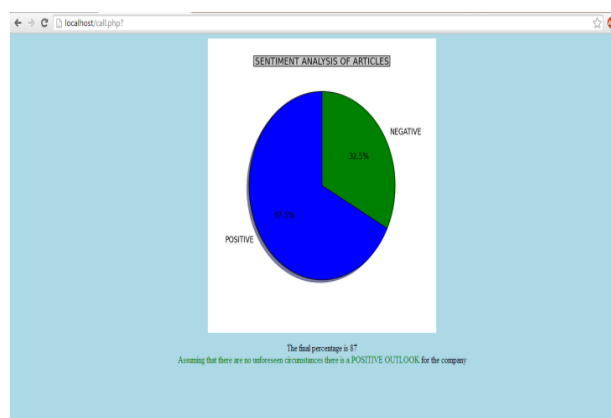


Fig. 3: Result of Sentiment Analysis and Final Results

This screenshot shows the pie chart which depicts the results of the Sentiment Analysis. It shows the percentage of articles that positive and the percentage of articles that are negative. It also shows the final percentage after combination of the Sentiment Analysis results with the results of the Technical Analysis. This page also gives

the final result which specifies whether the outlook for the company is positive or negative. In this case the outlook for the company chosen, i.e., HCL Technologies is positive.

VII. PERFORMANCE ANALYSIS AND EVALUATION.

The accuracy of a measurement system is the degree of closeness of measurements of a quantity to that quantity's actual (true) value. Precision can be defined as the percentage of correct items that are selected. These metrics are calculated for the Sentiment Analysis module. The accuracy calculated by this code is found to be 62%. Precision is found to be 78.5714285714% and recall is found to be 40.7407407407%. These metrics are for the Sentiment Analysis.

VIII. CONCLUSION

The aim of this work is to be able to extract large amounts of data in the field of stock market through the Internet at runtime for the purpose of analysis. The final analysis is the combination of the technical analysis and sentiment analysis and it helps in giving the overall outlook of the company that is chosen so that the user can make an informed decision. In this implementation we have extracted the historical stock prices of a certain company from a list of companies and analyzed those values to find the patterns in the stock values. We have also explored the predictive power of news articles on stock market domain to help the user make an informed decision. The historical prices are analyzed by technical analysis and the news articles are analyzed using sentiment analysis. This work could be extended by adding more companies and the exploration of various other techniques for both sentiment analysis and technical analysis.

REFERENCES

- [1] Indian Stock Market Predictor System, Vanipriya CH, Thammi Reddy K, pp 17-26 Springer International Publishing, 2014, DOI 10.1007/978-3-319-03095-1_3
- [2] Stock Prediction Using Event-based Sentiment Analysis, Masoud Makrehchi, Sameena Shah and Wenhui and Liao, IEEE Computer Society, 2013, pp 337-342
- [3] Twitter mood predicts the stock market, Johan Bollen¹, Huina Mao, Xiao-Jun Zeng., Journal of Computational Science, March 2011, pp. 1-8.
- [4] Evaluating and understanding text based stock price prediction models, Enric Junque de Fortuny, Tom De Smedt, David Martens, Walter Daelemans, Journal of Information Processing and Management, 2014, pp. 426-441.
- [5] Cheap Talk on the Web: The Determinants of Postings on Stock Message Boards, Peter D. Wysocki., Issue 98025 of Working paper (University of Michigan. Business School. Research Support, 1998.
- [6] Mining of Concurrent Text and Time Series, Victor Lavrenko, Matt Schmill, Dawn Lawrie, Paul Ogilvie, David Jensen, and James Allan, Proceedings of International conference on KDD Text mining workshop, 2001.
- [7] Stock Trend Forecasting Method Based on Sentiment Analysis and System Similarity Model", Kaihui Zhang, Lei Li, Peng Li and Wenda, IEEE, 2011.
- [8] Sentiment revealed in social media and its effect on the stock, Hailiang Chen, Prabuddha De, Yu Hu, and Byoung-Hyoun Hwang, IEEE Statistical Signal Processing Workshop, 2011, pp. 25 – 28.
- [9] Financial Text Mining: Supporting Decision Making Using Web 2.0 Content, Hsin-Min Lu and Hsinchun Chen, Tsai-Jyh Chen, Mao-Wei Hung and Shu-Hsing Li, IEEE Intelligent Systems, 2010, pp. 78-82.
- [10]. Sentiment Analysis and Subjectivity", Bing Liu, Handbook of Natural Language Processing, 2010.