RESEARCH ARTICLE

Stock market prediction using Neural Networks and sentiment analysis of News Articles.

Mr. Raj Thakur, Miss. Sanchita Badkas, Miss. Meenakumari Pade and Mrs. Pallavi Khude.
D.Y. Patil College of Engineering, Akurdi, Pune.

…………………………………………………………………………………………………....

Manuscript Info	*Abstract*
……………………..	…………………………………………………………………

Stock market prediction is considered as one of the classic problems of time series prediction due to high volatility of the financial market. Stock prices tend to be dependent on various factors like historical values of the company stock, P/E ratio of the company, news articles related to the company and public opinion which includes faith in the company and rumours related to the activities of the company. The system tries to predict the opening price of stock market for few particular companies based on the company's past performance and information available from various news sources.

…………………………………………………………………………………………………....

Introduction:-

A stock is a type of security that represents a claim on part of the corporation's assets and earning and thus signifies ownership [1] .Companies raise funds by allowing customers to buy the company's shares at an agreed price. Stock Market thus, in turn, affects the economy of a country. Buying and selling (the right) stocks lead to profits. Act of trying to determine the future value of company stock traded on an exchange is Stock Market Prediction. Stock markets have gained speed since the use of internet increased. Trading market analysis aims to predict the market fluctuations and the rise or fall of market in future.

It involves an assumption that the fundamental information that is publicly available in the past is somehow predictively related to the future stock returns. The objective of this study is to extract fundamental information from relevant news sources and use them in analysing or forecasting the stock market from any common investor's viewpoint. Researchers have developed a lot of interest in this area due to its dynamic and unpredictable nature.
Since stock prices are dependent on various parameters there exists no clean equation to determine the movement of the stock price.

Predicting stock market is difficult, since relationship between stock prices and the financial economic variables is not linear. Therefore, nonlinear models could produce more reliable and correct predictions of share prices. Machine learning techniques like Artificial neural networks (ANN) has ability to map nonlinear nature and hence can be used much effectively. This paper investigates systematic usage of neural network to predict stock market prices in future.

Corresponding Author:- Mr. Raj Thakur.
Address:- D.Y. Patil College of Engineering, Akurdi, Pune.

Literature survey:-

A lot of work had been done previously in the field of stock market prediction. Stock market prices are mostly dependent on the principle of demand-and-supply. Demand of a particular company share varies depending on various factors mentioned below:

1. Historical Prices of the company share which includes the opening prices, closing prices, highest prices and lowest prices for a time span varying from a month to an year.[3]
2. Financial status of the company, Dividends and balance sheets, Volatility of the stock[5]
3. Overseas markets that close around the same time when the concerned stock market opens up.[2]
4. Prices of oil, interest rates, inflation, GDP[4]
5. P/E ratio, general index, Volume traded[5]
6. Tweets made by people regarding the company
7. News articles related to the company.[8]

Various methodologies were used to predict the values of the stock. Decision trees were used in-order to determine whether the market will rise or fall. Linear regression approach was used to predict the trading volume of the stocks. Use of Support vector machines has also been done for predicting the rates of NASDAQ market.[2] But, a lot of previous work has proved that use of Artificial neural network is the best choice for predicting the stock markets because:

1. Ability to implicitly detect complex nonlinear relationships between dependent and independent variables.[4][5]
2. ANN is nonlinear model that is easy to use and understand compared to statistical methods.

Background: stock market prediction:-

There are few factors that help to analyse and predict its movement. These forces come under three categories:
1. Fundamental Factors
2. Technical Factors
3. Market Sentiments

Fundamental Factors:-

Stock prices would be determined primarily in an efficient market by these fundamentals, which, at some point, refer to a combination of two things:[5]

1. Earnings base (earnings per share (EPS))
2. Valuation multiple (a P/E ratio).

The level of earning base, the expected growth in the earning base, the discount rate, which is itself a function of inflation and the perceived risk of the stock are the key fundamental factors. EPS means the profit that a particular company has made per share and that too on the last quarter.
An owner of a common stock has a claim on earnings, and earnings per share (EPS) are the owner's return on his or her investment. Whenever a stock is been purchased, a proportional share of an entire future stream of earnings is purchased. That's the reason for the valuation multiple: it is the price one is willing to pay for future earnings stream. Part of these earnings might be distributed as dividends, while the remainder would be retained by the company for purpose of reinvestment.

Technical Factors:-

There are certain technical factors that helps to predict the bumpy ride of stock market, some them are
Expected moving average (EMA), Rate of change indicator, Moving average (MA), Relative strength index (RSI), On Balance Volume etc.

Market Sentiment:-

It refers to the psychology of market participants, individually and collectively. Market sentiment is often subjective, biased and obstinate that helps to make a judgement about a stock's future growth prospects.

The psychology of market participants is hugely dependent on the news and hence, news is another factor that affects the share price. When there is good news about a particular stock or a company which might result in a good future of the company and increase in profits,, people try to invest all their money in that particular company. This leads to increase in the interest of buying the stock whereas there exists a contrary situation too, where news could bring a negative effect where it could ruin the prospect of the particular stock.

Thus, investors depend on various different factors. Short term traders usually prioritize the technical factors of trading and hence incorporate them. While the long-term investors recognize that technical factors play a major role they also prioritize the fundamentals of trading. At the most fundamental level, supply and demand in the market determines stock price. Price times the number of shares outstanding (market capitalization) is the value of a company. It's meaningless to compare just the share values of two companies. Theoretically what affects the investor's valuation of a company is -Earning, but along with this there are certain other indicators also that can be used to predict the stock hence it is investor's sentiments, attitude and expectations play an equally important role in determining the prices of the stock market

Trading analysis can be of following types-
1. Fundamental analysis:-
It mainly focuses on statistical data of the company. The goal of fundamental analysis is to find a stock's intrinsic value. It basically means what a stock is really worth as opposed to the value at which it is being traded in the market. Fundamental analysis may involve examination of financial data, management, business concept and competition. It also includes quarterly balance sheets, daily news and rumours related to the company, return on investments etc. It finds intrinsic value of a stock and if the current value of stock is below intrinsic value, it generates a buy signal and is helpful for long term trading.

2. Technical analysis:-
The Fundamental analysts examine earnings, dividends, assets, quality, ratio, new products, research and the like[11]. Technicians employ many methods, tools and techniques as well, one of which is the use of charts. Using charts, technical analysts seek to identify price patterns and market trends in financial markets and attempt to exploit those patterns.
Technicians using charts search for archetypal price chart patterns, such as the well-known head and shoulders or double top/bottom reversal patterns, study technical indicators, moving averages, and look for forms such as lines of support, resistance, channels, and more obscure formations such as flags, pennants, balance days and cup and handle patterns.[9]
Technical analysis is done considering various market indicators, some of which include price transformations, up and down volume, all time high and low etc. This factor helps to know whether the asset is trending and also its direction of growth in near days. Technical analysis also considers relationships between price/volume indices and market indicators such as moving average, relative strength, and price/volume index.

Goals:-
The system first analyses the past prices of the stock market and news related to the company to predict the opening value of that particular stock. The information from news is scraped from a variety of web sources, the open source selenium web driver is a good choice for accomplishing this task. The system scrapes the financial information like news articles, prices in history, volume of stocks sold etc. The scraped data is then processed to obtain only the data that is useful to us. The system then tries to combine sentiment analysis on this processed data along with fundamental and technical analysis.. This data is then passed through "neural network" in order to forecast the opening value of a particular stock. Output will be the approximate Prediction of highest value or the lowest value for a particular share for the given day.

The system's input consists of the news, public opinion (for rumours and people's faith in the company), volume traded, and public shareholding, stability of the share and various technical parameters of the stock market trading system. Stability of the stock's price of a particular company plays a major role in predicting the future value of a share. It depicts the performance of the share over past time and how the share has recovered from market falls. Thus it provides a measure of how good the stock is likely to be in its present position.

Identified analysis:-

Following are the modules in the system:

1. Scraper:
a. There are three scrapers in the system as follows:
    i. First, the scraper that scrapes the values of the company stock like:
        1. Highest price
        2. Lowest price
        3. Opening price
        4. Closing price.
              The prices scraped are over an interval of one-month
    ii. Second, the scraper that scrapes the news articles related to a particular company.
    iii. Third, the scraper that extracts tweets from the twitter related to a particular company.

2. Sentiment analyser
a. News related to a particular company are passed through the analyser which vectorizes the news articles and also the tweets.
b. The overall sentiment of the news articles and tweets is analysed using bag of words.
c. This is done using 'gensim' library of python which help in the Doc2Vec conversion.

3. Neural Network:
a. The results from the scrapper and the sentiment analyser are passed as input the neural network.
b. Expected results are also fed to the network so as to train it.
c. Error values are back-propagated to update the weights.
d. 'Tensorflow' and 'keras' modules of python are used to implement the neural network

Information gathering:-

Scrapper extracts the news articles regarding the recent company activities and the analyser shows how it will affect its share value. It extracts highest value and the lowest value of a share from the given company so as to get its stability that will help in predicting its future value. The news information will be scrapped from a collection of financial websites. Beautiful Soup[10] is a Python library which is used for scraping the live data and it converts incoming documents to Unicode and outgoing documents to UTF-8. Along with beautiful soup the open source Selenium web driver is used so as to continuously fetch the information as it is visible to a visitor on the webpage. The scrapping would be constrained only to gather information about a particular company or a set of companies.

The process of sentimental analysis comes into picture that will show how the collected news affects the sentiments of people which in turn affects the trading of the stock. We also consider Sentiment analysis on twitter feeds which is performed on the tweets extracted from the Twitter feed. This is done using 'tweepy' library. The system then uses a set of technical analysis methods, variables such as public shareholding which means equity shares of the company held by public. Such Parameters aid in the technical analysis of the trading market so as to minimize the market risk. This parameter along with various other parameters like all-time high, all time low, average returns in the form of numerical value will be input for an artificial neural network.

Experimental setup:-
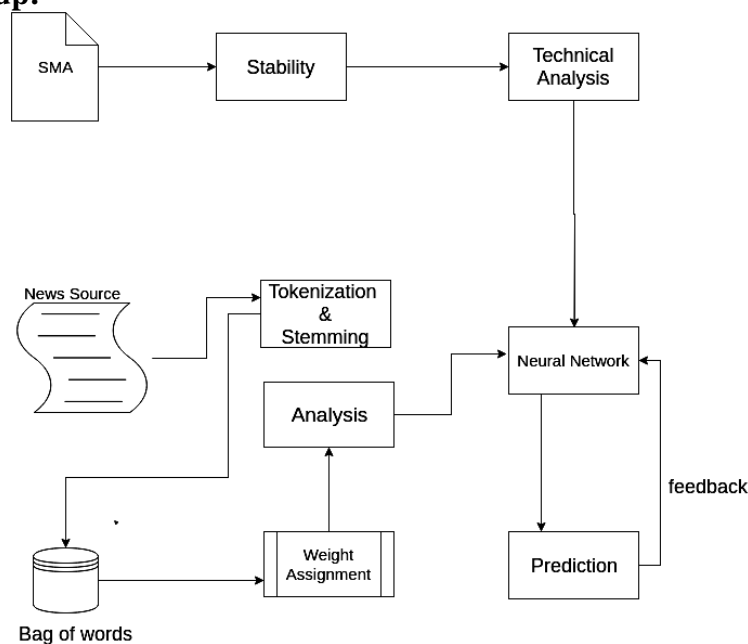
The basic experimental setup of the system is as follows:

1. Scrape the news information, tweets related to the company and other stock related information from the web.
   a. The news articles are scraped from reliable news sources like reuterindia.com
   b. Historical prices of the company shares along with its volatility is scraped from nseindia.com
   c. Tweets related to the company are extracted from twitter using 'tweepy' library

2. The news articles are passed through the sentiment analyser which returns the overall sentiment of the news. The same process is performed on the tweets to quantify the public opinion about the company

3. The values are passed to the neural network where number of nodes, activation function and number of hidden nodes are decided based on the accuracy of the system. After a right amount of training, the neural network is passed test values to check for its accuracy.

Summary:-

The system tries to predict the opening price of the share of a particular company by analysing news related to the company for the next trading day. The system also considers the stability of the stock price of the particular company. Sentimental analysis is used for vectorizing the news, along with technical and fundamental analysis. Artificial neural networks are used as they hold good for pattern recognition, classification and forecasting of the market so as to model nonlinear process without any prior assumptions. The nature of this process helps in obtaining accurate results.

Various stakeholders attempt to have such a system as a strategy that aims towards solving the financial situations as well as improving the efficiency and efficacy of financial risk management. Such proposed model can be a helpful tool for the investors for taking the right decision in regards to their stocks in order to extract any predictive information from that given data. The system will also display a graph of "Stock price" vs. "Date" in which the predicted values and real values in the past are plotted.

Hence it can be concluded that hybridized parameters like combination of technical and fundamental variable gives better prediction accuracy over application of standalone parameters and news being a vital factor that affects investor sentiments can be used to predict the drive of the market.

References:-

1. www.investopedia.com/terms/s/stock.asp
2. Shunrong Shen, Haomiao Jiang, Tongda Zhang, "Stock Market Forecasting Using Machine Learning Algorithms"
3. Farhad Soleimanian Gharehchopogh, Tahmineh Haddadi Bonab, Seyyed Reza Khaze ,"A linear regression approach to prediction of stock market trading volume", International Journal of Managing Value and Supply Chains (IJMVSC),2013
4. Nguyen Lu Dang Khoa, Kazutoshi Sakakibara and Ikuko Nishikawa , "Stock Price Forecasting using Backpropagation Neural Networks with Time and Profit Based Adjusted Weight Factors", SICE-ICASE International Joint Conference 2006
5. Zabir Haider Khan, Tasnim Sharmin Alin, Md. Akter Hussain, "Price Prediction of Share Market using Artificial Neural Network (ANN)", International Journal of Computer Applications, 2011
6. Ruchi Desai, Prof. Snehal Gandhi, "Stock market prediction using data mining", International Journal of Engineering Development and Research, 2014
7. Radu Iacomin, "Stock Market Prediction", International Conference on System Theory, Control and Computing, 2015
8. Minh Dang, Duc Duong, "Improvement Methods for Stock Market Prediction using Financial News Articles", IEEE,2016
9. Murphy, John J., "Technical Analysis of the Financial Markets." New York Institute of Finance, 1999, pp. 1-5, 24-31. ISBN 0-7352-0066-1
10. https://www.crummy.com/software/BeautifulSoup/bs4/doc/.
11. https://en.wikipedia.org/wiki/Technical_analysis.