APPLICATIONS OF DATA MINING TECHNIQUES FOR STOCK MARKET

Prathamesh Patil[1], Rohit Pathak[2], Amit Vaidya[3]
KJ Somaiya Institute Of Engineering & Information Technology, Mumbai, India
[1]prathmeshter@gmail.com,[2]rohatak@gmail.com,[3]amit.vaidya@hotmail.com

***Abstract*— Stock Market trading is a very huge business where lots of money is involved. Along with the money so are the risks and the problems. The basic question an investor faces is whether to invest money or not in a specific stock. Currently majority of the transactions which the investor does is based on human feeling and sentiment and a sudden surge of buying and selling takes place. What is lacking is proper guidance and suggestions based on study of data using technology which can help him in doing so. With the help of Prediction and Data Mining Algorithms the project would be implemented. So the project "APPLICATIONS OF DATA MINING TECHNIQUES FOR STOCK MARKET" hopes to provide a prediction about the prices of the stock as accurately as possible. The investor would get a prediction which is based on the study of some data parameters of stock regarding the expected price so he could decide whether to hold, buy or sell the stock. The knowledge gained through this experiment will be used to make intelligent and informed investment decisions in the future.**

***Index Terms*—Data mining, TPWS, Moving Average, Technical Analysis, Stock Market.**

I. INTRODUCTION

A stock market or equity market is a public entity (a loose network of economic transactions, not a physical facility or discrete entity) for the trading of company stock (shares) and derivatives at an agreed price; these are securities listed on a stock exchange as well as those only traded privately.

The size of the world stock market was estimated at about $36.6 trillion at the beginning of October 2008. The total world derivatives market has been estimated at about $791 trillion face or nominal value 11 times the size of the entire world economy. The value of the derivatives market, because it is stated in terms of notional values, cannot be directly compared to a stock or a fixed income security, which traditionally refers to an actual value. Also more and more companies are being listed on the stock market. So ultimately this is going to grow.

Thus we can see the amount of money involved is huge and so are the risks involved. Thus analysis becomes a major factor in purchasing the stocks or even in selling them. After all the ultimate aim of the investor is to maximize the profits and face minimum risks. Hence Stock Market Analysis is very important for the Investor

II. STOCK MARKET ANALYSIS

Let us now see the basic analysis types which are followed

A. Technical vs. Fundamental Analysis

Technical and Fundamental Analysis are as different as the Firm-Foundation Theory is from the Castles-in-the-Air Theory because while Technical Analysis is used by those who support the latter, Fundamental Analysis is used by those who support the former. Technical Analysis consists of the following of charts to predict where the market will go and, "Most chartists believe that the market is only 10 percent logical and 90 percent psychological"

While the Fundamental analysts don't believe in studying the past or using charts because the, "Fundamental analysts take the opposite tack, believing the market is 90 percent logical and only 10 percent psychological"

B. Technical Analysis

When analyzing stocks using technical analysis there are two principles that the follower must take into the account. "The first principle of technical analysis is that all information about earnings, dividends, and the future performance of a company is automatically reflected in the company's past market prices". "The second principle is that prices tend to move in trends: A stock that is rising tends to keep on rising, whereas a stock at rest tends to remain at rest" . The ability to find and read the charts of a stock and analyze it is paramount to the implementation of the technical analysis of the stock market. Technical analysis concerns the study of charts and followers of this type of stock analysis are known as, chartists, and do not believe in information such as dividends or earnings as the fundamentalist does because they feel that it distracts from the ability to correctly predict what the future of the stock will do. A chart showing the past market prices and the trade volumes of a stock in the past is all the technical analysis would need to analyses where the stock will go in the future. It is important to know that as a stock is rising it is believed that it will keep rising because of the momentum of a stock. Once a stock has an uptrend it will then slightly fall and go up again forming what looks like a shoulder then fall slightly and go up a little higher than the should fall and go up again forming the next shoulder, and then the technical analysis looks for the stock to fall below the line that the valleys have lined up on, called a neckline. If the stock falls below the neckline it will quickly be sold because it will usually result in a downtrend similar to the uptrend occurring before.

C .Data Mining

Data mining is the process of discovering meaningful new correlations, patterns and trends by sifting through large amounts of data stored in repositories, using pattern recognition technologies as well as statistical and mathematical techniques. A process used by companies to turn raw data into useful information. By using software to look for patterns in large batches of data, businesses can learn more about their customers and develop more effective marketing strategies as well as increase sales and decrease costs. Data mining depends on effective data collection and warehousing as well as computer processing. Association Rules would be used in our project. They would help us in making more appropriate selections about the decision to take.

III. EXISTING SYSTEM

The existing system is not that efficient in giving us a good prediction. Although they generated results the results obtained were not accurate. What was lacking is efficiency. Also there were problems with the algorithms used they performed less effectively over very large databases and Failed where Data-set was too large to fit in memory. Drawbacks of traditional clustering algorithms were that results depend on approach used for represent each cluster. Centroid-based approach (using the mean) consider only the cluster centroid.

Other approach is all-points (based on d min) use all the points inside him for cluster representation. This choice is extremely sensitive to outliers and to slight changes in the position of data points. As the first approach can't work well for non-spherical or arbitrary shaped clusters, So there are many problems with the existing system. Thus we can see that problems are being faced by the existing system regarding the prediction process and also the decision giving process. Also efficiency is an issue which needs to be improved. These are some of the problems of the existing system which we would try to improve in our project..

IV. PROPOSED SYSTEM

In the Proposed System we are going to implement a system where the input is raw data in terms of the stock prices and other factors related to the stocks. The prices then are processed by the system to give a prediction about the price and what option to exercise i.e. Buy Sell or Hold. Association mining rules are also applied. Association rule mining, one of the most important and well researched techniques of data mining, was introduced first. It aims to extract interesting correlations, frequent patterns, associations or casual structures among sets of items in the transaction databases or other data repositories. At the beginning of our study we selected nine indicators because of their exactness as well as they provide same type of outputs or their results can be interpreted into common output. Our proposed algorithm also provides similar types of output which is Buy, Sell or Hold where buy means the stock price is going to increase and the investor is suggested to buy some shares, sell means the stock price will decrease soon so the investor is suggested to sell his/her shares if it's in investor's portfolio and hold means investor is suggested to keep their stock unchanged.

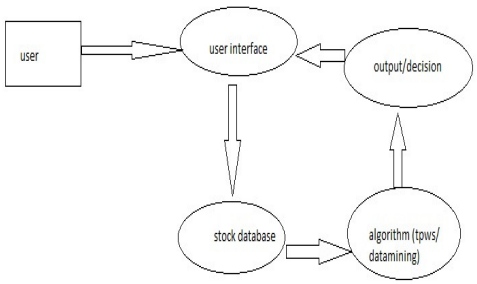

Fig 1. System Design

V. METHODOLOGY

The stochastic oscillator only works on the closing price of a security. It actually draws two lines to predict future movement of the security. But from our analysis we found out closing price in not an appropriate measure of a stock's value because it represents only the last price when the trading ends or stops. Therefore we decided to use typical price, which is the average of high, low and close price. Moreover only price data is not sufficient to predict a stock's pattern because it is important factor for a security to determine how much volume traded at which price. For the above reason we customized the original stochastic algorithm and proposed typical price weighted stochastic (TPWS).

Similar to original stochastic algorithm it works on stage over two periods. It is calculated by the following ways: Typical price= (High + Low + Close)/3 TPW= ((Typical Price-lowest typical price in last d1 period)* volume traded on lowest typical price)/(Highest high in d1 period * volume traded on highest typical price – lowest typical price in last d1 period)* volume traded on lowest typical price) * 100 TPWS= Moving average of TPW on d2 period From the study, we have chosen 14 days for d1 and 5 days for d2.It provides decision based on result (buy, sell or hold) and also a point to represent the significance of the decision. If today's TPWS is greater than Highrange then: Point= (TPWS - Highrange) /Highrange * 100 If today's TPWS is smaller than Lowrange then: Point= (Lowrange-

TPWS) / Lowrange * 100 Else Point=Mod (Midrange-TPWS) / midrange * 100 For this study, TPWS point greater than 10 was considered as 10 and negative values are considered as 0. A details pseudo code is given to understand the technique more clearly.

```
Input: 14 days high,low,close price and volume traded
Output: Decision signal and point

procedure typicalprice(high,low,close)
        return (high+low+close)/3

procedure tpws(high[],low[],close[],volume[])
        for each day, i=1 to d1 do
                tp= typicalprice(high[1],low[1],close[1])
                stoc_d1[i]= (tp * vomume[i] – minimum typical price in d1 period* volume[]
                traded on lowest typical  price)/((maximum high[] in d1 period * volume[]
                traded on maximum high[]) –( minimum low[]in d1 period* volume[] traded on
                minimum low[]) )* 100
        for each day, i=1 to d2 do
                 sum_d1+=stoc_d1[k]
        stoc_d2= sum_d1/d2
        point=point_calculation(stoc_d2)
        if(stoc_d2>highrange) then
                decision= Sell
        else if(stoc_d2<lowrange) then
                decision= Buy
        else
                decision= Hold
return decision and point
```

Fig 2. TPWS Algorithm

VI. IMPLEMENTATION

Our System is currently live through an Internet Website. We are currently giving the basic Buy Sell Hold options for some companies and are in the process of trying to improve and add the system functionalities. We are performing tests and doing system testing and trying to improve the efficiency of the system and would achieve the desired results as we expect.

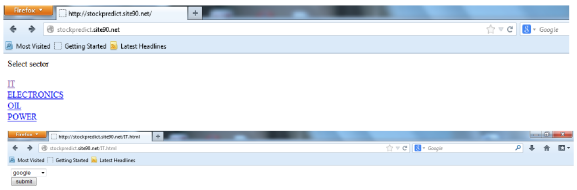

Fig. 3 Select Sector and Select Company

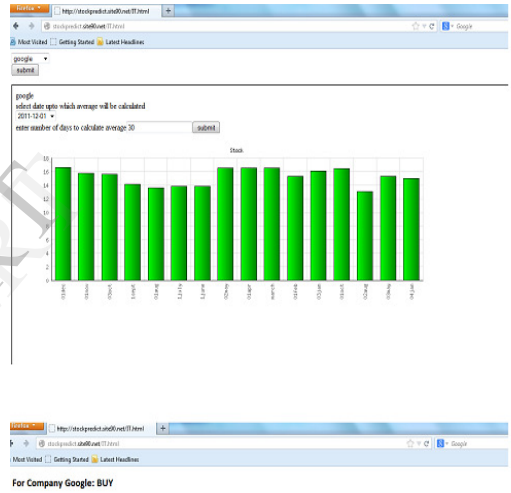

Fig. 4 The Graph and The Result

VII. REFERENCES

[1] Stock Market Prediction model using TPWS and Association Rules Mining
Sheikh Shaugat Abdullah and Mohammad Saiedur Rahaman Department of Computer Science American International University-Bangladesh Dhaka, Bangladesh
shaugat, saied}@aiub.edu
[2] "Extracting Trading Rules from the
Multiple Classifiers and Technical Indicators in Stock Market,"
K. Kim, I. Han and J. S. Chandler
The Korea Society of Management Information Systems, '98 International Conference, pp. 147-163(17), 1998.
[3] "An Optimized Approach to Analyze Stock market using Data Mining Technique,"
D. P. Gandhmal, R. B. Parihar and R. V. Argiddi,
IJCA Proceedings on International Conference on Emerging Technology Trends (ICETT) (1), pp. 38-42, 2011.
[4 "Stock Price Forecasting by Hybrid Machine
Learning Techniques,"
] C.-F. Tsai and S.-P. Wang, International MultiConference of Engineers and Computer Scientists 2009, vol. I IMECS 2009, pp. 755-760, March 2009.
[5] V. H. Shah, "Machine Learning Techniques for Stock Prediction,"
www.vatsals.com retrieved: June 2012.
[6], "A Decision tree-Rough set Hybrid System for Stock Market Trend Prediction,"
B. B. Nair, V. P. Mohandas and N. R. Sakthivel
International Journal of Computer Applications 6(9), pp. 1–6, September 2010.
[7] R. V. Argiddi and S. S. Apte, "Future Trend Prediction of Indian IT
Stock Market using Association Rule Mining of Transaction data,"International Journal of Computer Applications 39(10), pp. 30-34, February 2012.
[8] "Technical Analysis for Stocks," posted: May 2009,
www.tradegreeks.com , retrieved: July 2012.
[9] R. Agrawal, T. Imieliński and N. Swami, "Mining association rules
between sets of items in large databases," ACM SIGMOD International
Conference on Management of Data, vol. 22 pp. 207-216, June 1993.
[10] S. Kotsiantis, D. Kanellopoulos, "Association Rules Mining: A Recent
Overview," GESTS International Transactions on Computer Science and
Engineering, vol. 32 (1), pp. 71-82, 2006.
[11] S. A. S. Olaniyi, Adewole, Kayode S. and R. G. Jimoh, "Stock Trend
Prediction Using Regression Analysis – A Data Mining Approach," vol.
1 no. 4, pp. 154-157, July 2011.
[12] "Technical Analysis: Indicators and Oscillators," www.investopedia.com , retrieved: July 2012.