Prediction Engine for Stock Market

Pavithra S, Pradeep Kumar S, Khishore V.S

II-year M.sc Decision and computing Science, Department of Computing, Coimbatore Institute of Technology, Coimbatore, India.

Abstract:

Stock market is a promising financial investment that can generate great wealth. However, the volatile nature of the stock market makes it a very high-risk investment. Thus, a lot of researchers have contributed their efforts to forecast the stock market pricing and average movement. Researchers have used various methods in computer science and economics in their quest to gain a piece of this volatile information and to make great fortune out of the stock market investment. Our prediction engine (PE) will predict near real values of the stocks based on market behaviour using artificial intelligence and predictive analysis. The sources for our data will be 'money control' which shows the stock values of Indian companies and 'Nation stock exchange' (NSE). Our PE uses certain concepts and algorithms such as correlation, Time series/trend data, analysis of variance, Multi Correlation of data, saturation, and the probable cause analysis to validate and predict the market values. The forecasts are predicted using the R software. This will enable the investors to make decisions accordingly prior to investing in the firm. Thereby the PE will be a great tool for the share buyers as this implies a clear picture of how the firm's stock market will be in the next coming years.

Keywords- Time series of trend data, Multi correlation of data, Headroom Analysis, Market saturation, Probable cause analysis.

I. INTRODUCTION

Generally, a stock market or an equity market or share market is an aggregation of buyers and sellers where a loose network of economic transactions of stocks or shares which represent claims of ownership on businesses Some of the examples of the latter include shares of private companies which are sold to investors through equity crowd funding platforms. Stock exchanges list shares of common equity as well as other security types, e.g. corporate bonds and convertible bonds. We are predicting an near to real time value of stock for various commodities. By using those values one can decide the investment he or she has to make if so in future. So, what is stock market prediction? It 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 for the investors. One of the efficient-market hypotheses suggests that stock prices reflect all currently available information and any price changes that are not based on newly revealed information are inherently unpredictable. Others disagree and while those with this viewpoint possess multitude methods and technologies which intentionally allow them to gain future price information. There can be various methods or tools through which one can predict the future value of any commodity. The values predicted aren't the real values. They are only the approximate prices which could be useful at some point prior any decision making or investments. In view of this, tools such as python, R Cassandra, r and similar software's can be used to predict the future prices of commodities. Here, we have used the r software to predict the value of a stock for a specific year

and interpret it. The r Language provides a lot of statistical tools like linear and nonlinear modelling, classical statistical tests, time series analysis, classification, clustering and graphical techniques which are really expensive. R has an advantage of having a well-designed publication of quality plots which can produced and interpreted; over the other software's. So, getting into the r software, the trend lines give us a visual interpretation of how the progress is up to. The methodologies used are;

Time series of data using trendlines, Correlation of data, probable cause analysis, headroom analysis.

1.1 TRENDLINES

They give us the trends of variations based on the varying time period. The "trendline()" is a function used to plot different models.

For lin2P, the formula goes like this:

y=a*x+b

For lin3P,

y=a*x^2+b*x+c

For log2P,

y=a*ln(x)+b

For exp3P

y=a*exp(b*x)+c

For power3P,

y=a*x^b+c

On specifying the model type, we get the trendline based on it. A simplified form of code is;

```
trend.line(x, y, type = "linear", plot = TRUE)
```

The type could be either linear, quadratic, exponential, and so on. The trendline plotting is used in our predictions of stock prices to track the prices graphically. The 'type' varies accordingly.

1.2 FUNCTIONS

Creating user defined functions is one of the greatest strength any programming language. In r, a user defined function can be created and it can be called by the function name itself and its parameters if they are specified in the declaration.

variable←function_name(parameters)

Here we have used functions wherever required, to calculate the future values of stocks.

"Sum" is the function name.

```
sum<-function(a,o){
   trendline(a,o,model="line2P",plot=TRUE)
}
```

"predi" is the function name.

```
predi<-function(u){
  h<-u$Year
  av<-u$Avg
  relation<-lm(av~h)
  print(relation)
  print(summary(relation))
  newdata<-data.frame(h=2018)
  print(newdata)
  result<-predict(relation,newdata)
  print(result)
}
```

1.3 CORRELATIONS IN FUNCTIONS

The values of stocks predicted, are correlated with one another for the purpose of analysis and auditing. It is necessary to compare the stock values to into a conclusion which is higher and which is relatively lower.

"Cor()" is the function to produce the correlation and "Cov()" is to produce the covariances.

A simplified form is **cor (x, use=, method=)**

The method takes three values as, Pearson, spearman or Kendall.

II. READING THE DATA

The data is read from the excel document or text file. Based on the source of data, it is read in different formats. If the data is in the excel file, then it is read into the r using the read.csv()

method. If the data is in a text file, then it is read using read.delim() method.

out the results for any data we input. For example, if the data is in csv format, the engine will predict the future values according to the csv file we input.

The Engine is designed in such a way that, it will be predicting the stock prices of a commodity at any year we enter and will give

```
x<-read.csv("gold.csv",header=TRUE)

y<-read.csv("crudeoil.csv",header=TRUE)

z<-read.csv("silver.csv",header=TRUE)

m<-read.csv("aluminium.csv",header=TRUE)
```

2.1 RESULTS

Fig1:

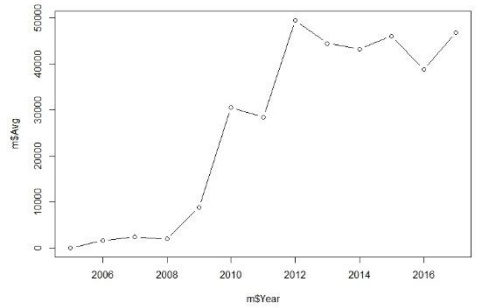

Fig2

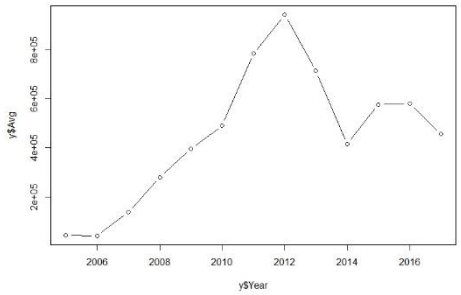

Fig3:

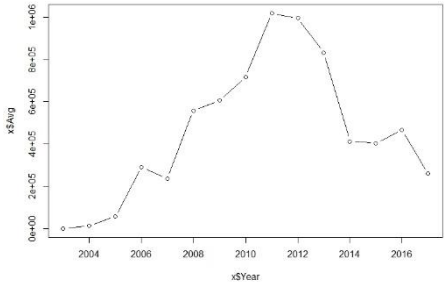

Fig.4:

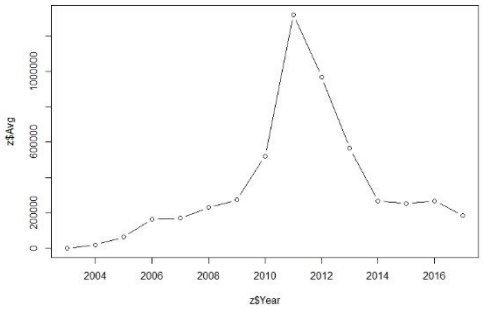

Fig1: Aluminium Graph

Fig.2: Crude Oil

Fig.3: Gold

Fig.4: Silver

The trendlines for the stock commodities corresponding to the years are obtained below;

Fig.1:

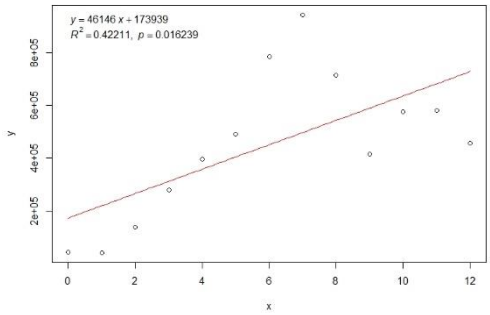

Fig.2:

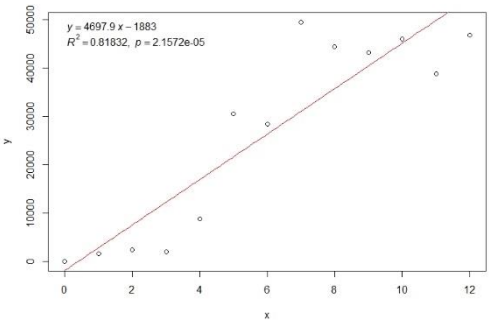

Fig.3:

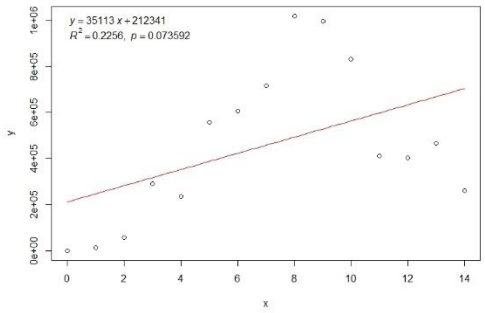

Fig.4:

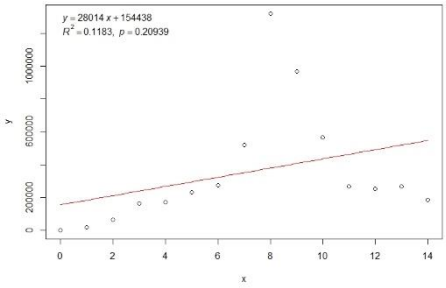

Fig.1: Crude oil

Fig.2: Aluminium

Fig.3: Gold

Fig.4: Silver

2.2 INTERPRETATION

From the above graphs it is evident that the stock prices were all at peak during the year 2012. The trends makes leads us to the decision making process by which we decide to go for it or not to; which is the main idea of headroom analysis.

III. ANALYSIS

The stock prices change from time to time and year to year. We have taken the past year's data of prices for each commodity and predicted that the future prices of the stocks of various categories differ from each other and individually as well.

3.1 PROBABLE CAUSE ANALYSIS

There is always an uncertainty in the prices of the stocks in through out the world. They vary time to time. It becomes difficult for the people who try to invest and spend money on the commodities without being sure how will the

stock change in the future, whether it will decrease or increase. This prediction engine will enable the customers who invest on the stocks to make decisions according to the value interpreted by the engine. The Customer will thereby stop investing or invest at a lower cost if the prices aren't getting better day by day.

3.2 HEADROOM ANALYSIS

In this type of analysis, we determine which competitive markets will enter the most critical strategic decisions facing energy service providers (ESPs). Thus this analysis which determines the difference between the current electric change and the cost of serving customers.

This is generally decision making process which makes us to choose between a "GO" or a "NO GO" decision in the changing market activity.

3.3 CONCLUSION

Using the engine designed using the tools in r, we have predicted the approximate values of stocks. This will thereby yield us a value on which we can rely on to an extent. As the tools such as trendlines of data, multi correlation of data, user defined are used; the results predicted are of course reliable for the customers.

REFERENCES

1. Abhyankar, A., Copeland, L. S., & Wong, W. (1997). Uncovering nonlinear structure in real-time stockmarket indexes: The S&P 500, the DAX, the Nikkei 225, and the FTSE-100. Journal of Business & Economic Statistics, 15, 1–14.
2. Forecasting of Indian stock market using artificial neural networks, by Munna Majumder, Md Anwar Hussain.
3. Desai, V. S., & Bharati, R., (1998). A comparison of linear regression and neural network methods for predicting excess returns on large stocks. Annals of Operations Research, 78, 127–163.
4. Malkiel, Burton G. (2003). "The Efficient Market Hypothesis and Its Critics". CEPS Working paper No 91.
5. Xiaodong Li a, Haoran Xie a,⇑, Li Chen b, Jianping Wanga, Xiaotie Deng "News impact on stock price return via sentiment analysis", ScienceDirect 2014
6. Shou-Hsiung Cheng, ―Forecasting the Change of Intraday Stock Price by Using Text Mining News of Stock‖, IEEE, 2010.
7. Fung G.P.C., Yu, J.X., and Lu, H., 2005. ―The Predicting Power of Textual Information on Financial Markets.‖ IEEE Intelligent Informatics Bulletin, 5(1), pp.1-10.
8. M. Mittermayer, G. Knolmayer, ―NewsCATS: A News Categorization And Trading System‖, IEEE Proceedings of the Sixth International Conference on Data Mining (ICDM'06), 2006.
9. Mitchell, M.L., Mulherin, J.H., 1994. The Impact of Public Information on the Stock Market. Journal of Finance, 49(3), pp.923-950.
10. Mendelsohn Louis B. (2000) Trend Forecasting with Technical Analysis: Unleashing the Hidden Power of Intermarket Analysis to Beat the Market, Marketplace Books.