



INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X

Impact factor: 4.295

(Volume 5, Issue 2)

Available online at: www.ijariit.com

Share market analysis for share selection using data mining technique

Munaf Patel

pmunaf123@gmail.com

Thakur College of Engineering and Technology,
Mumbai, Maharashtra

Zahir Aalam

zahir.aalam@thakureducation.org

Thakur College of Engineering and Technology,
Mumbai, Maharashtra

ABSTRACT

The share market has been a field of vast interest both for those who wish to make money by trading shares in the share market. Generally, there is an opinion about share markets like high risk and high returns. Even though we have a huge number of potential investors, only very few of them are invested in the share market. The main purpose is they are not able to take risk of taking the skill of investors. Though get low returns they want to save their money. One important reason for this problem is that they don't have proper guidance for making their portfolio. In this paper we focus the real-world problem; we had selected three indices such as SENSEX, NIFTY. The analysis is purely based on the data collected from the past three years. The Data mining technique, Time series interpretation is applied for the Data analysis to show the ups and downs of a particular index. The correlation and Beta are the tools which give the suggestion about the share and its risk. The correlation tool is used to identify the relationship between the index and the company individually. This Beta is used to identify the risk associated with the share.

Keywords— Linear regression, Data analysis, Correlation, Beta, Time series

1. INTRODUCTION

The principal segments of the monetary market are cash and capital markets. The Security advertise is isolated into two classes, essential market and auxiliary market. The essential market is that piece of the capital markets that arrangements with the issuance of new securities. The way toward pitching new issues to financial specialists is called endorsing. On account of another share issue, this deal is called as the first sale of share (IPO). The auxiliary market is an on-going business sector, which is furnished and sorted out with a place, offices and different assets required for exchanging securities after their underlying advertising. It alludes to a particular place where securities exchange among numerous and unspecified people are conveyed out through intermediation of the securities firms, i.e., an authorized intermediary, and the trades, a specific exchanging association, as per the tenets and controls are built up by the trades. Securities exchange determining incorporates revealing business sector patterns, arranging venture procedures,

distinguishing the best time to buy the shares and what shares to buy. There has been a basic requirement for robotized ways to deal with powerful and productive usage of enormous measure of money related information to help organizations and people in key arranging and speculation basic leadership. It is fundamental and needs to discover need of great importance in the field of Market conduct is to uncover the market patterns, venture choices, speculation techniques, distinguishing the best time to buy shares and what shares to buy.

The share trading system is a complex, non-stationary, clamorous [1][13] and non-direct powerful framework. Gauging securities exchange, cash swapping scale, insolvencies, understanding and overseeing budgetary hazard, exchanging prospects, FICO score, advanced administration, bank client profiling, and tax evasion investigations are the centre testing errands to be considered. The clever thought is to give an understanding of the innocent financial specialists. The outcomes investigation and clarification will give important astute to every single new financial specialist. This paper presents a calculation, which incorporates the devices Correlation and Beta, to translate the relationship and the hazard related to the share.

2. BACKGROUND STUDY

The Analysis was made, in light of the previous authentic information of BSE and NSE. The current research, in a relationship with information digging strategies for Time Series utilizing the calculations like "ARMA" and AR, is more valuable for the above said gigantic information examination of a securities exchange. Yet, what choices arranged outcome it will give in credulous speculators' viewpoint. The principle expectation is to make attentiveness for the new hopeful for securities exchange and furthermore to evacuate the dread about share related issues. There are numerous information mining procedures, with its own calculations will bolster for monstrous information investigation. [2] Data mining is a strategy of finding valuable examples in information that are covered up and obscure in an ordinary position. [3] It originates from a few fields like insights, [4] database machine learning. It is more important to, comprehend the conduct of the share trading system. This can be an awesome test for all securities exchange financial

specialists. In the context of credulous financial specialists, there is no standard framework or rules for them to get it. It is an approach where it demonstrates money markets stream [5] [6] by perusing the information from late years. The data mining strategy predicts [3] the inevitable or sudden falls met by the share trading system. The customary calculation like ARMA is running with its own particular disservices. To recognize designs in time arrangement information, The Correlation and Beta estimations are the two new changed.

Table 1: NSE Indices

CNX MIDCAP
NIFTYMIDCAP 50
CNX INFRA
CNX REALTY
S&P CNX DEFTY

Table 2: Banks under Nifty

Unione Bank of India	HDFC Bank
Axis Bank	ICICI Bank
Bank of Baorda	Kotak Mahindra Bank
Bank of India	Oriental Bank
State Bank of India	Punjab National Bank

3. PROBLEM DEFINITION

In this paper the objective is to meet out the general challenge, i.e., the main goal is to improvise the decision-making power and improvement about the investment in the share market from the new user’s perspective. The new investors are having the problem of choosing a valuable share. The reason for this concern is the lack of knowledge about the market and lack of knowledge. It is very essentials to identify them.

4. THE RESEARCH PROPOSAL

4.1. Time Series Analysis

The definition Time arrangement expresses that it is a requested grouping of estimations of a variable at similarly divided time interims [7]. It can be acquired from any framework at the decided time interval [10]. The day by day value change of a market, Process and quality control, Economic Forecasting, Census Analysis, Share Market Analysis might be considered as a Time arrangement

For Example, consider the equation (a)

$$X = \{x_i \quad i = 1 \dots N\} \tag{a}$$

In this equation, *i* is the time index and *N* is the total number of observations. The critical occasions are created over some stretch of time. Consider the quick changes, for example, fall and raise met by the offer market. It researches and anticipates the procedures. The Box Jenkins or Autoregressive Integrated moving normal (ARIMA) is a traditional time series used to model such time series. Nevertheless, the ARIMA method is limited by the requirement of stationary of time series and control of residuals. The event depiction function changes according to the forecast plan. For Example, *x_i* represents today’s closing price of the share and it is necessary to predict the percentage changes of tomorrow’s price, the event depiction function can be defined in the following equation (b)

$$g(t) = \frac{x_i + 1 - x_i}{x_i} \tag{b}$$

The main drawback is that the time series should be converted into stationary [9] and periodic series to analyze it. Data mining is a tool for identifying hidden data from a pool of data. Please find below Nitfy index of top 5 banks

4.2 Proposed algorithm

Declaration
 X = list of banks in bank nifty
 Mc = Market cap
 Initialize
 A[i] = BankNiftyclosingprice
 B[i] = Individual Bankclosingprice

Loop (1...n months)
 Calculate
 Avg = ΣA[i] / no. of working days;
 End;
 Find the correlation {A [i] B [i]}
 End;

Beta calculation

Declaration
 Closing price (bank) = *r_p*
 Current closing price = *r_a*

$$\beta = \frac{Covar(r_a, r_p)}{Var(r_p)}$$

If $\beta < 1$ then the share has a low risk
 Else
 The share has a high risk
 End if;
 End

5. DATA ANALYSIS AND INTERPRETATION

As mentioned early, there are three indices chosen for the study. From these each index we had chosen five shares.

Table 3: Company Name vs. M Cap

Company Name	MCap..Rs (In crores)
HDFC Bank	509,516.17
ICICI Bank	217,245.59
Kotak Mahindra	208,099.36
Axis Bank	148,471.13
IndusInd Bank	105,290.45
Yes Bank	80,006.97
RBL Bank	20,520.06
Federal Bank	18,951.89
IDFC Bank	18,409.31
City Union Bank	10,708.47
Karur Vysya	8,120.19

1. SBI 2. HDFC Bank 3. ICICI Bank 4. Axis Bank 5. Kotak Bank

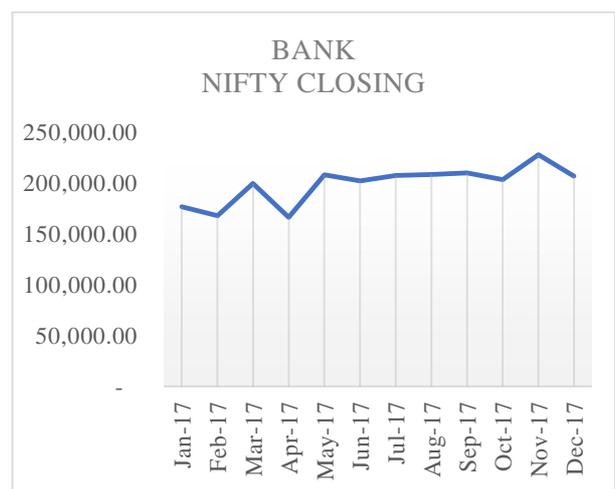


Fig. 1: Nifty Month vs. closing point

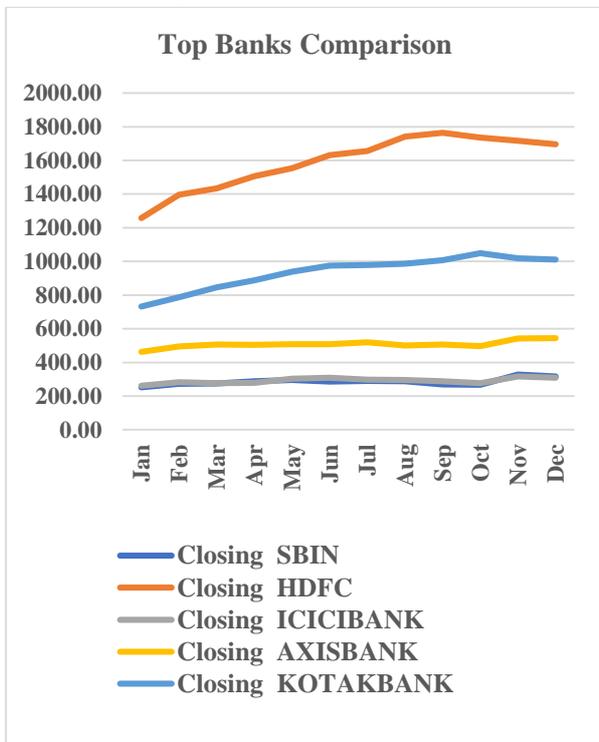


Fig. 2: Comparison of Nifty Closing Top Banks

Table 4: Nifty month vs. closing index figure 2. Nifty Month vs. closing point

Date	Bank Nifty/CNX Closing Point
Jan -2017	176110.3
Feb -2017	167453.45
Mar -2017	199035.25
Apr -2017	165862.3
May -2017	207613.7
Jun -2017	201746.05
Jul -2017	206852.5
Aug -2017	207924.85
Sep -2017	209536.25
Oct -2017	202773.55
Nov -2017	227144.6

Table 5: Correlation between Bank Nifty/CNX with top 5 Banks

Date	NIFTY/CNX	SBIN Closing	HDFC Closing	ICICI BANK Closing	AXIS BANK Closing	KOTAK BANK Closing
Jan -2017	176,110.30	253.13	1,257.84	262.50	462.23	732.64
Feb -2017	167,453.45	272.69	1,395.22	282.81	494.33	786.39
Mar -2017	199,035.25	275.23	1,433.75	277.03	505.64	845.21
Apr -2017	165,862.30	289.16	1,506.65	278.71	505.01	886.90
May -2017	207,613.70	295.48	1,553.32	304.31	507.48	939.50
Jun -2017	201,746.05	285.65	1,631.61	308.33	509.17	974.53
Jul -2017	206,852.50	289.35	1,655.47	298.13	519.21	978.71
Aug -2017	207,924.85	288.88	1,742.14	296.06	501.60	986.24
Sep -2017	209,536.25	268.55	1,764.43	288.83	506.55	1,006.98
Oct -2017	202,773.55	267.59	1,735.43	277.44	496.27	1,048.41
Nov -2017	227,144.60	327.73	1,716.06	316.15	543.12	1,017.87
Dec -2017	206,445.25	315.81	1,696.04	309.37	544.33	1,010.49
Correlation		0.984	0.759	0.993	0.988	0.191

6. CONCLUSION

Numerous Clear examinations were made in view of the recorded information of the NSE securities exchange to recognize the stream and changes which occurs in securities exchange for each list. From the three years information, which had started from January 2017 and up to the primary seven day

stretch of December 2017. The diagram which gives a thought about the different changes occurred in the share exchange is appeared. The above chart demonstrates a sudden fall in the share trading system amid the subsidence time frame. The stream demonstrates a consistent development after 2017 which gives the positive sign for general financial specialists and furthermore for the innocent speculator. We wish to close from the paper that putting resources into the managing an account record in securities exchange will dependably give gainful answers for the news financial specialists. It is a creative start and proposed framework to present the basic leadership control for the credulous financial specialists through the following exploration work

7. REFERENCES

- [1] Basaltoa N, Bellottib R, De Carlob F, Facchib P, Pascazio S (2005). Clustering share market companies via chaotic map synchronization, Physica A.
- [2] Chi-Lin L, Ta-Cheng C (2009). A study of applying data mining approach to the information disclosure for Taiwan’s share market investors, Expert Systems with Applications.
- [3] David E, Suraphan T (2005). The use of data mining and neural networks for forecasting share market returns, Expert Systems with Applications.
- [4] Hsiao-Fan W, Ching-Yi K (2004). Factor Analysis in Data Mining, Computers and Mathematics with Applications.
- [5] Jar-Long W, Shu-Hui C (2006). Share market trading rule discovery using a two-layer bias decision tree, Expert Systems with Applications.
- [6] Jie S, Hui L (2008). Data mining method for listed companies’ financial distress prediction, Knowledge-Based Systems.
- [7] R. J. Pavinelli, “Time Series Data Mining: Identifying Temporal Patterns for Characterization and Prediction of Time Series Events”, PhD Dissertation, Marquette University, 180 p., 1999.
- [8] X. Feng and H. Huang, “A Fuzzy-Set-Based
- [9] Reconstructed Phase Space Method for Identification of Temporal Patterns in Complex Time Series”, IEEE Trans. on Knowledge and Data Engineering, Vol. 17, No. 5, pp. 601-613, 2005.
- [10] R. J. Pavinelli and X. Feng, “Data Mining of Multiple Non-stationary Time Series, in Proceedings of Artificial Neural Networks in Engineering”, St. Louis, Missouri, pp. 511-516, 2005.
- [11] H. Kantz and T. Schreiber, Nonlinear Time Series Analysis, Cambridge: Cambridge University Press, 388 p., 1997.
- [12] J. Han and M. Kamber, Data Mining: Concepts and Techniques, San Francisco: Academic Press, 800 p.,2007
- [13] S.M. Weiss and N. Indurkha, Predictive Data Mining: A Practical Guide. San Fransisco: Morgan Kaufmann., 228 p., 1998. Clustering share market companies via chaotic map synchronization –Elsevier physical A 345–(2005)196-206.
- [14] A heuristic forecasting model for shared decision making – Math ware and soft computing – 2005 – 3339.
- [15] A review on Data mining Applications to the performance of Share marketing – IJCA-(0975-8887) Volume 1 – No.3
- [16] World Federation of Exchanges “2007 “Annual report and market statistics – Retrieved on 2008-09-30
- [17] Fama E.F., (1970) Efficient capital markets: A review of theory and empirical work “ a journal of finance 25.pp-384-417
- [18] Han: Jiawei and Kamber Micheline 2006: Mining concepts and techniques –second edition
- [19] Eng.W.F.1988.the technical analysis of shares options and futures –advanced trading systems and techniques

- [20] An approach for generating financial forecasting using data mining techniques – ICETT-2011. 60558-131.News Headlines" ([http:// www. cnbc. com/ id/ 27166818](http://www.cnbc.com/id/27166818)). Cnbc.com. October 13, 2008. Retrieved March 5, 2010
- [21] Analysis and clustering of nifty companies of share market using data mining tools –journal of Engineering Research and studies. [23] Learn About Indian Share Market - & How Shares are Traded in India for NRIs, OCIs and PIOs?? By NriInvestIndia.com Dated: May 24, 2000.
- [22] Technical market indicators optimization using evolutionary Algorithms – GECCO 2008 - ACM - 978-1-