Analyzing Different Stock Options and Support in Decision Making for First Time Investors

[1] Ch. Padmini, [2]A. Tejaswi, [3]Ch. Nitheesha, [4]Bh. Venkata Sai Tejaswi, [5]M. Sravan

[1]Assistant Professor,
[1,2,3,4,5]Department of CSE, Dhanekula Institute of Engineering & Technology,
Ganguru, Vijayawada, Andra Pradesh, India

ABSTRACT

Now-a-days many youngsters and retired professionals want to invest their savings in stock market. But youngsters don't have sufficient knowledge about stock market and its fluctuations. They often confused where to invest their savings and how much to invest in which form they should buy their shares and what is the best time to buy and sell the shares. They need a proper guidance from an unbiased and reliable resource. This project provides an analysis report on the share prices of six different banks from the previous six months which can help in deciding where to invest and how much to invest to get profitable returns. This analysis is useful for both shareholders speculation decision making and also bankers risk congnizance. The methodologies that we are used are support vector machine and heteroskedasticity. A chasewillden stock market is a taken as an experimental platform where the share values of various banks are collected. And the result shows the share value volatility and also the close price. The estimate result shows the better bank for shareholders to speculate.

Keywords: *Support vector machine, heteroskedasticity, volatility, share values*

I.   INTRODUCTION

Shareholders often found that it is a time taking process to gain useful information to help them in speculation decision making. Efficient speculation decision making today is based on a variety of information sources including historical financial data series. There have been a number of studies showing that the sentiment contained in the stock market based on share values.

Our work develops a novel decision-support system using sentiment analysis, support vector machine (SVM) and generalized autoregressive conditional heteroskedasticity (GARCH) modeling. Chasewillden stock market, a widely used financial website has been selected as an experimental platform where corpuses of financial data of share prices were collected. For the analysis of share values we collect the data from various banks such as ICICI, SBI, PNB, Axis, Canara, BOB.

Important questions include: Is there an obvious correlation between shareholder sentiment and stock market performance? How can shareholder sentiment estimate future stock price? Solution for these questions is a sentiment index would be applicable. Research on sentiment index selection has recognized both direct sentiment and indirect sentiment indices. However, a direct sentiment index based on examine and an indirect sentiment index based on related stock market data can be imprecise. Share value information can reflect investor sentiment and can play an important role in investor decision making. These opinions could alter the way in which shareholders expend, dealing, gain, and share information. Express shareholder sentiment and producing a sentiment index from share values would be valuable. Overall, our study shows that share value sentiments do contain valuable information for speculating decision making and supports the shareholder sentiment hypothesis that irrational shareholders do impact stock markets.

The residue of the paper is structured as follows. We highlight previous research related to stock markets in the related work section and introduce our method in the architectural design of GARCH-SVM based on sentiment index section. We collect data from the Chasewillden stock market and conduct experiments to quantitatively assess our method in the data description and experiments and evaluations sections. Finally, we conclude the paper with possible further research.

II.   RELATED WORK

Stock price trend prediction is an interesting research area because more accurate predictions are directly related to more returns in stocks. That's why, in recent years, significant efforts have been put into action to develop new models that can predict for future trend of a specific stock or overall market. Most of the existing techniques use technical indicators. Some of the researchers portrayed that there is a strong relationship between news article about a particular company and its stock prices fluctuations.

Tina Ding, Vanessa Fang, Daniel Zuo, "Stock Market Prediction based on Time Series Data and Market Sentiment", 2012., the authors developed a system which predicts stock market movements on a given day, based on time series data and market sentiment analysis. They collect share prices from the internet into Excel spreadsheet. For sentiment analysis, the authors obtained Twitter Census stock Tweets data-set from Info-chimps, a privately held company which offers a "data marketplace". Naive Bayes Classifiers is used to analyze sentiment in the tweet data set. The SVM, Logistic and Neural network techniques would be used for predicting market movement.

Phillip Tichaona Sumbureru, "Analysis of Tweets for Prediction of Indian Stock Markets", International Journal of Science and Research (IJSR), 2015 mainly focuses on the prediction of daily stock movements of three Indian companies listed on National Stock Exchange (NSE). The Support Vector Machine (SVM) was used for the prediction of the stock market. The tweets were collected directly from twitter using Twitter API which were futher filtered using keywords. The relevant stocks were downloaded directly from internet.

III.   SYSTEM ARCHITECTURE

Figure.1 shows the conceptual flow chart for the methodology we use to conduct context-sensitive sentiment analysis of stock market based on share values of various banks. We first, collect data from various banks. . Then, using sentiment analysis, we identify features from share values to automatically predict sentiment polarity. We then aggregate postings for each share on a daily basis. Thereafter, we use a coarse sentiment index and SVM classifiers to build GARCH-SVM models to estimate future share value volatility.

First, we collect the previous share values of the various banks which were posted by shareholders and it is used to evaluate the time series of the share values and also the sentiment analysis

Data PreProcessing
Data preprocessing is an important step in the data mining process. It is a method used to perform normalization and also remove the unwanted data, duplicates etc.

Sentiment Analysis
Sentiment analysis is used to express shareholder sentiment because manual collection would be impractical. We use sentiment analysis technology to spontaneously classify unstructured reviews as positive or negative, and then identify shareholder sentiment as either bullish or bearish. We built a sentiment index to incorporate the sentiment value for each day.

GARCH-SVM Model
Volatility modeling and forecasting have been important in finance research over the past decades. There is a large number of well-established financial models for measuring and forecasting volatility. GARCH models were developed as the extensions of autoRegressive conditional heteroskedasticity (ARCH) models. GARCH models have been widely used to forecast time series containing features of autocorrelation and heteroskedasticity. Such applications include electoral behavior, finance, and electricity price
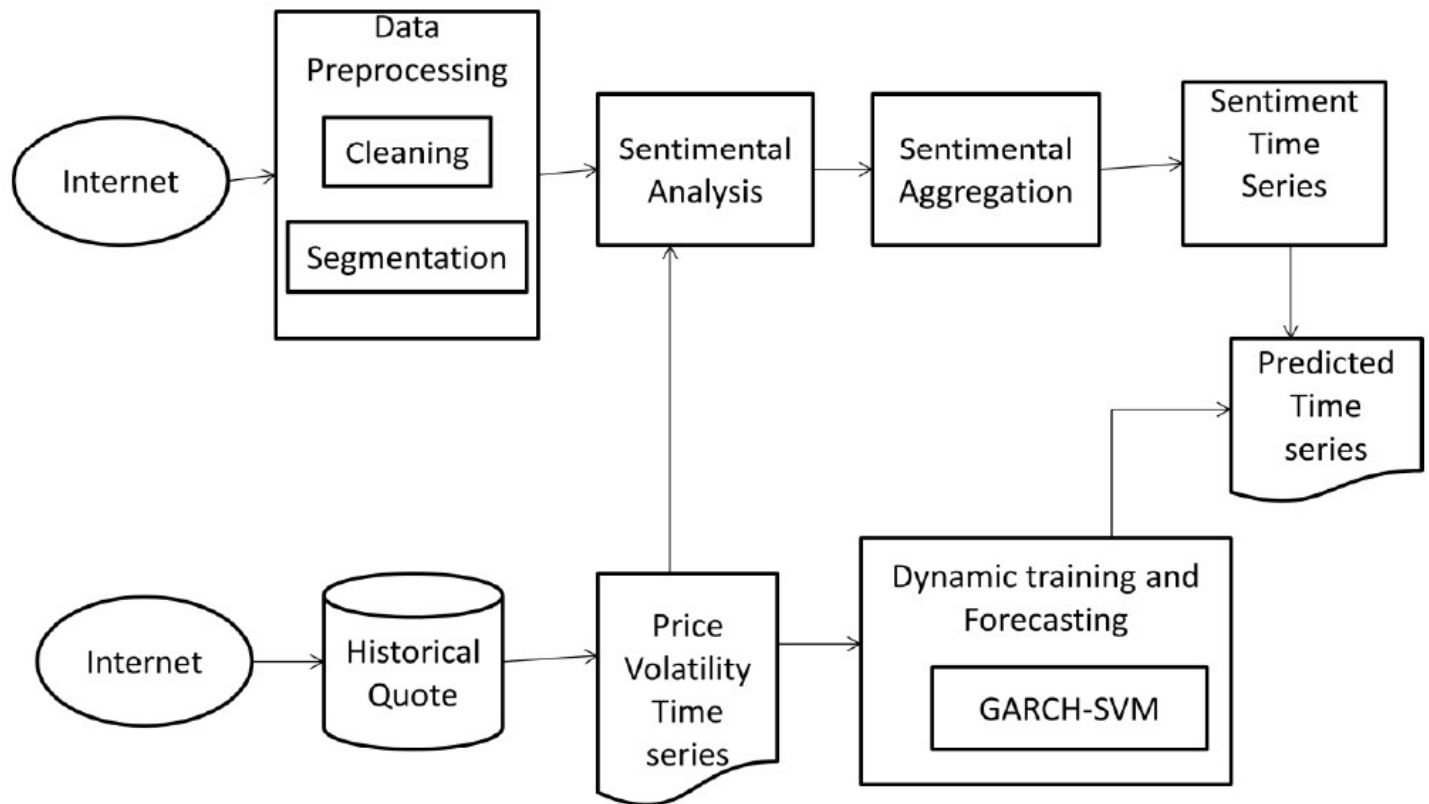
Fig.1: Flowchart to associate sentiment and volatility

SVM models are regressions seeking to maximize separation among sets. SVM models have been combined with GARCH models in many applications, to include forecasting electricity price, and many financial studies.SVM are supervised learning methods capable of generating input-output mapping functions from training data. SVM functions can be used for classification (categorization of data) or regression (estimation of dependent variable value). SVMs belong to a family of generalized linear models that classifies based on the value of a linear function.

We employ GARCH modeling and integrate results into an SVM model to accommodate complicated nonlinear and asymmetric relations embedded in volatility forecasting. This is because the information contained in a volatile series' history may involve many events.

Support vector machine(SVM) is a method for the classification of both linear and non linear data. If the data is linearly separable, the SVM searches for the linear optimal separating hyperplane (the linear kernel), which is a decision boundary that separates data of one class from another.

IV.   DATA DESCRIPTION

One of the key assumptions in finance theory (specifically behavioral finance) is that not all shareholders are fully rational and their demand for risky assets is influenced by their beliefs or sentiments that are not absolutely justified by fundamental news. Basically, the investors are subject to sentiment. We use both share values and financial time series data.

Stock Forum Data

We created a corpus of share values, which are taken from a well known finance website Chasewillden Stock Market.

Bank Stock Monthly Prices

A large number of share values are updated in this stock market every day, but some of them had no value for our research. Fig. 2 shows the monthly share values of the stock market of various banks with respect to close price.
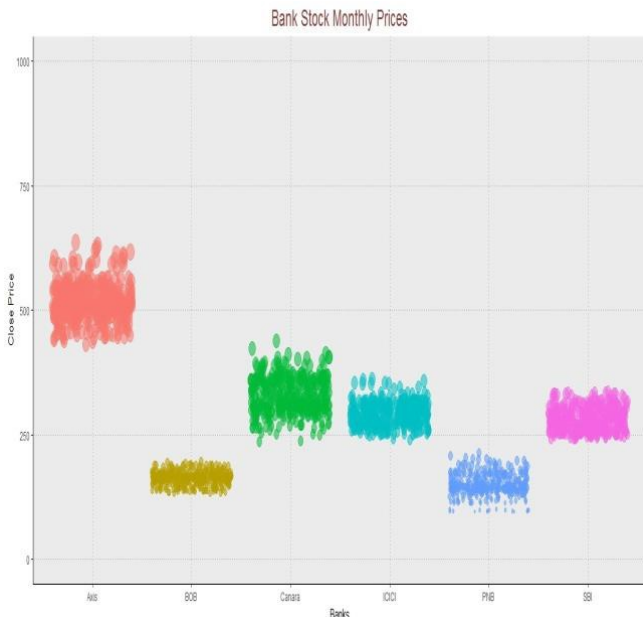
Fig.2: Bank Stock Monthly Prices

Daily close price

Fig.3 shows the daily share values of a month with respect to the close price. Each dot represents a day share value.
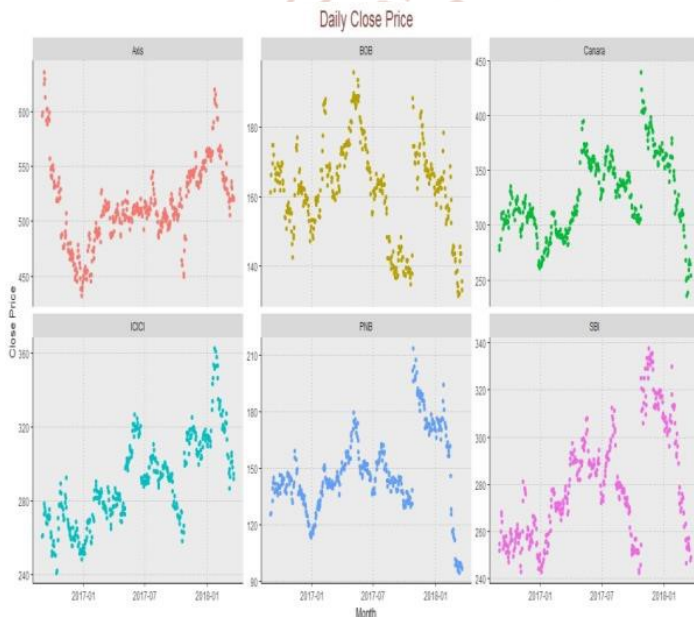

Fig.3: Daily Close Price

Monthly Traded Quantity with Price

Fig.4 shows the traded quantity i.e., the total number of shares that exchanged hands in a given trading day with respect to the price.
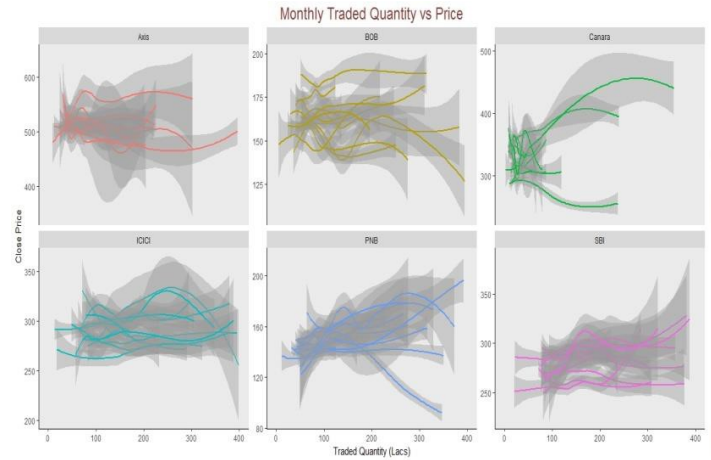

Fig.4: Monthly Traded Quantity with price

V. CONCLUSION

Chasewillden stock market is taken as a financial website, was selected as an experimental platform to obtain a corpus of share values of a stock market of various banks. By statistically analyzing these data the shareholder can decide which bank is better to those and invest. This project is really helpful for first investors to get a honest review about the company. The data collection doesn't take much effort as the required data is available on the internet. Even though the first investor doesn't have much knowledge about the market or company. It is easy to make an assumption about the company which can be proved more accurate. Thus this platform will be more useful for those investors who want to invest in small amount to secure their future.

REFERENCES

1. IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEMS, VOL. 44, NO. 8, AUGUST 2014.
2. Vivek Rajput et al, International Journal of Computer Science and Mobile Computing, Vol.5 Issue.6, June- 2016.
3. D. D. Wu et al., Enterprise Risk Management in Finance, Desheng Dash Wu and David L. Olson 2015.
4. Emma Haddi et al. / Procedia Computer Science 17 (2013) 26 – 32
5. Emma Haddi et al. / Procedia Computer Science 17 (2013) 26 – 32
6. B. Watkins,"Riding the wave of sentiment: An analysis of return consistency as a predictor of future returns," J. Behav. Fin., vol. 4, no. 4, pp. 191–200, 2003.

7. T. Zhang, J. Li, and P. Malone, "Closed-end fund discounts in Chinese stock markets, Chin. Econ., vol. 37, no. 3, pp. 17–38, 2004.

8. H.-J. Sheu, Y.-C. Lu, and Y.-C. Wei, "Causalities between sentiment indicators and stock market returns under different market scenarios," Int. J. Bus. Fin. Res., vol. 4, no. 1, pp. 159–171, 2010.

9. S.-H. Poon and C. W. J. Granger, "Forecasting volatility in financial markets: A review," J. Econ. Lit., vol. 41, no. 2, pp. 478–539, 2003.