Chapter 7

Stock Market Prediction Based on Machine Learning Approaches

V. Lalithendra Nadh and G. Syam Prasad

Abstract Forecasting stock market based on the information available with high precision is not so consistent because of its unsteady nature. There are numerous approaches in the anticipation of stock markets. Machine learning systems are a standout among other methodologies in expectation. Numerous researchers have done wide research over the years using different machine learning algorithms. In this paper, the written work examines on different computational tools such as genetic algorithms (GAs), support vector machine (SVM), artificial neural networks (ANNs) are used for stock market forecasting.

Keywords Genetic algorithms (GAs) Support vector machine (SVM) Artificial neural networks (ANNs)

7.1 Introduction

Forecasting of a stock market is an uncertain undertaking in the business environ ment. The fundamental standard of a stock market is price rising of a share based on the increase in investment in business and the other way around. Forecasting the stock price may prompt benefit or misfortune that the trader needs to tolerate. Any anticipation will help in decision making either to buy or sell that particular share. Based on several reliable anticipation strategies and news items, numerous traders are trading by compelling some calculated risks. The informative news from various media channels has a high impact on share value movement. Fundamental analysis is enough for investors. Technical analysis and shortterm Predictions are required for trading. Many inputs are to be considered to anticipate the future value in the stock market. But it is absolutely confusing the traders while taking decisions. The successful formula for the anticipation of a stock market is getting the finest outcomes with less number of data sources. Forecasting is very difficult due to its volatile nature. The movement average of various stocks was incorporated in the stock market index. Market movement is reflected by the index rather than an individual stock movement. To such an extent, numerous researchers focused on forecasting

Satyanarayana et al., Computational Intelligence and Big Data Analytics, SpringerBriefs in Forensic and Medical Bioinformatics

76 7 Stock Market Prediction Based on Machine Learning Approaches

individual stock prices. There are a few methodologies that are utilizing machine learning algorithms. The most wellknown approaches are genetic algorithms (GAs), support vector machine (SVM), artificial neural networks (ANNs).

7.2 Literature Review

Kyoungjae Kim [1] proposed a genetic algorithm (GA) for the expectation of securi ties exchange in an unexpected way. The main advantage of GA is to propel the learn ing calculation and manysided quality diminishment in feature space. Kyoungjae Kim [1]. Authors proposed GA in a different way to calculate the affiliation weights for feature discretization (FD) and for ANN during the forecast of stock value. In gen eral, for getting the affiliation weights in ANN, gradient descent algorithm is utilized, and backpropagation algorithm is utilized as a local search algorithm. According to [1], author point of view, gradient descent algorithm performs very poorly when contrasted with GA. The learning procedure will be simplified when the information is appropriately discretized, and learned consequences of enhanced generalizability will diminish the redundant and noisy information adequately. Prior to the search process, the affiliation weights and inception for feature discretization are initialized to arbitrary values. The genetic algorithm feature discretization (GAFD) in Fig. 7.1 has three stages mainly. Three sets of parameters have been taken in the primary stage. To start with a set of affiliation weights in the middle of the information layer and hidden layer of the system, the second set affiliation weights in the middle of a concealed layer and yield layer are edges for FD in the third set. The search ing parameters are encoded on chromosomes to influence the best utilization of the fitness function. The determined affiliation weights with feedforward computation from the primary stage and the linear function are consolidated in the second stage. The thresholds for FD and derived affiliation weights are valuable to the proposed information in the final stage.

Kim [2] proposed expectation of consistent value change in stock record utilizing support vector machine (SVM). SVM is a sort of learning algorithm in Fig. 7.2 that locates a unique kind of highest margin hyperplane and linear model. Support vectors are only the preparation illustrations that are near to highest margin hyperplane. The binary decision classes are isolated by a hyperplane for the linearly divisible element. Deciding the parameters and finding the support vectors are nothing but unraveling a quadratic program which is linearly compelled. SVM alters the contributions of highdimensional component space to normalize nonlinear class by building a direct model. Among different parts, the best model must be picked which limits the estimate. Polynomial kernel and Gaussian outspread basis function are the normal cases of a kernel function. If the two parameters are chosen inadequately, it will cause underfitting or overfitting issues. The information factors that are utilized as a part of this approach are Korea Composite Stock Price Index (KOSPI) on a regular schedule and technical indicators. The traditional change in stock price index is delegated as 1 or 0. 0 denotes that todays index is prominent than following

7.2 Literature Review 77

Fig. 7.1 Framework of genetic algorithm feature discretization

Fig. 7.2 Support vector machine

day, whereas 1 denotes that todays index is lesser than the following day. Every feature component is standardized independently to the specific range through linear scaling which ensures that greater value inputs do not amaze inputs of smaller value, to reduce errors during forecasting. To recover similar cases, the closest neighbor strategy is utilized for casebased reasoning (CBR). The closest neighbor strategy is normal recovery technique which can be simply applied to financial information. Euclidean distance is the assessment function of closest neighbor strategy. At last, the outcomes were satisfied when contrasted with backpropagation network.

de Oliveira and Nobre [3] proposed a technique for stock market forecasting using artificial network approach. This approach comprises three stages in Fig. 7.3 to antic ipate the tendency and behavior of a stock. The three stages are (1) getting samples,

78 7 Stock Market Prediction Based on Machine Learning Approaches

Fig. 7.3 Artificial neural network structure

(2) input preprocessing, and (3) forecast. The foremost objective is to anticipate the closing price of a particular stock. Samples were taken from a stock named PETR4 which trades on BM&FBOVESPA. The preprocessing was done on day byday quotes of that specific stock together with some technical indicators. The technical indicators are nothing but mathematical computations which will be con nected to a series of information. For each date, if there is any misplaced information, that specific date ought to be erased from the sample set. The linear intersection strat egy was picked during the process of normalization for a series of sample sets having different esteem scales. One of the machine learning tasks called supervised learn ing was chosen for investigating the training data. The backpropagation algorithm was chosen for training the network to calculate the error contribution. Windowing process is utilized to pass the sample sets in addition with forecast horizon and size of the window picked. At last, the training process was finished and the performance was measured utilizing validation and test.

7.3 Conclusion

For any trader or investor, the stock market forecast will help in making appropriate decisions and furthermore gives some clarity during buying and selling the stocks. This paper is reviewed on a stock market forecast by utilizing the genetic algorithms (GAs), support vector machine (SVM), artificial neural networks (ANNs). The behavior and trends of the stock are explained, the closure of the next day is the ref erence function, the fiveday annotations Successfully obtaining appropriate results. ANN mainly focused on improving the learning algorithm and association weights. Combination of genetic algorithm with ANN has been proposed to overcome the limitation of ANN. GA looks for close optimum thresholds of feature discretization

7.3 Conclusion 79

and close to best possible solutions of association weights in learning algorithm. The study of the genetic algorithm also has a few constraints while processing the elements in hidden layer and optimization of a learning process. To overcome the ordinary procedures, SVM anticipated structural risk minimization standard for the enhanced model. SVM got the better accurate outcomes contrasted with the back propagation in neural networks. Each algorithm has some advantages as well as few limitations. Hence, the combination of some technical indicators and the machine learning algorithms will give best results.