International Journal of Research and Applications

ISSN (online): 2349-0020 ISSN (print): 2394-4544 http://www.ijraonline.com/

Research Article

Genetic Algorithm for prediction of financial performance

Sravanthi Kurapati

Corresponding Author:
sravanthi.k1982@gmail.com

DOI:
http://dx.doi.org/
10.17812/IJRA.6.21(3)2019

Manuscript:
Received: 17th Jan, 2019
Accepted: 28th Feb, 2019
Published: 20th Mar, 2019

Publisher:
Global Science Publishing Group, USA
http://www.globalsciencepg.org/

ABSTRACT

This paper is to produce a GA-based methodology for prediction of stock market performance along with an associative classifier from numerical data. This work restricts the numerical data to stock trading data. Stock trading data contains the quotes of stock market. From this information, many technical indicators can be extracted, and by investigating the relations between these indicators trading signals can discovered. Genetic algorithm is being used to generate all the optimized relations among the technical indicator and its value. Along with genetic algorithm association rule mining algorithm is used for generation of association rules among the various Technical Indicators. Associative rules are generated whose left side contains a set of trading signals, expressed by relations among the technical indicators, and whose right side indicates whether there is a positive ,negative or no change. The rules are being further given to the classification process which will be able to classify the new data making use of the previously generated rules. The proposed system introduces a new genetic algorithm for prediction of financial performance with input data sets from a financial domain.

Keywords: Genetic Algorithm, Associative rule mining, Technical Indicators.

University College of Engineering- Kothagudem,
Kakatiya University, Telangana State, India -507115.

IJRA - Year of 2019 Transactions:
Month: January - March
Volume – 6, Issue – 21, Page No's: 1110-1115
Subject Stream: Computers

Paper Communication: Author Direct
Paper Reference Id: IJRA-2019: 6(21)1110-1115

DOI: 10.17812/IJRA.6.21(3)2019

International Journal of Research and Applications
Jan - Mar © 2019 Transactions 6(21): 1110-1115

eISSN : 2349 – 0020
pISSN : 2394 – 4544
www.ijraonline.com

COMPUTERS RESEARCH ARTICLE

Genetic Algorithm for prediction of financial performance

Sravanthi Kurapati

University College of Engineering- Kothagudem,
Kakatiya University, Telangana State, India -507115.

ABSTRACT

This paper is to produce a GA-based methodology for prediction of stock market performance along with an associative classifier from numerical data. This work restricts the numerical data to stock trading data. Stock trading data contains the quotes of stock market. From this information, many technical indicators can be extracted, and by investigating the relations between these indicators trading signals can discovered. Genetic algorithm is being used to generate all the optimized relations among the technical indicator and its value. Along with genetic algorithm association rule mining algorithm is used for generation of association rules among the various Technical Indicators. Associative rules are generated whose left side contains a set of trading signals, expressed by relations among the technical indicators, and whose right side indicates whether there is a positive ,negative or no change. The rules are being further given to the classification process which will be able to classify the new data making use of the previously generated rules. The proposed system introduces a new genetic algorithm for prediction of financial performance with input data sets from a financial domain.

Keywords: Genetic Algorithm, Association rule mining, Technical Indicators.

1. INTRODUCTION

Applications of data mining techniques for stock investment include clustering, decision tree etc. Moreover, researches on stock market discover trading signals and timings from financial data. Because of the numerical attributes used, data mining techniques, such as decision tree, have weaker capabilities to handle this kind of numerical data and there are infinitely many possible ways to enumerate relations among data. Stock prices depend on various factors, the important ones being the market sentiment, performance of the industry, earning results and projected earnings, takeover or merger, introduction of a new product or introduction of an existing product into new markets, share buy-back, announcements of dividends/bonuses, addition or removal from the index and such other factors leading to a positive or negative impact on the share price and the associated volumes. Apart from the basic technical and fundamental analysis techniques used in stock market analysis and prediction, soft computing methods based on Association Rule Mining, fuzzy logic, neural networks, genetic algorithms etc. are increasingly finding their place in understanding and predicting the financial markets. Genetic algorithm has a great capability to discover good solutions rapidly for difficult high dimensional problems. The genetic algorithm has good capability to deal with numerical data and relations between numerical data. Genetic algorithms have emerged as a powerful general purpose search and optimization technique and have found applications in widespread areas.

Classification is a well-known task in data mining that aims to predict the class of an unseen instance as accurately as possible. While single label classification, which assigns each rule in the classifier the most obvious label, has been widely studied, little work has been done on multi-label classification. Most of the work to date on multi-label classification is related to text categorization. In existing associative classification techniques, only one class label is associated with each rule derived, and thus rules are not suitable for the prediction of multiple labels. However, multi-label classification may often be useful in practice.

Although associative classification has better prediction accuracy than traditional classification approaches, it has a weak capability of handling numerical data and its relations. To improve the capability of handling numerical data in associative classification, there are two issues that must be addressed, including constructing a suitable relation representation method of numerical data and building associative classifiers from numerical data with suitable relation representations. The major contributions of this study are to propose a simple yet powerful structure for relation representation of numerical data in associative classification problem and to improve the capability of handling numerical data in associative classification.

2. LITERATURE REVIEW

To conduct our research, we first need to specify an appropriate universe of trading rules from which the current GA may have been applied to. In stock market, when the brokers or dealers want to buy or sell a share, some of them will depend on a technical trading rule. Robert Edwards and John Magee[5] defined Technical trading rules as "the science of recording the actual history of trading (price changes, volume of transactions, etc.) in a certain stock or in "the Averages" and then deducing from that pictured history the probable future trend."

A Technical Indicator [1] is a series of data points that are derived by applying a formula to the price data of a security. Price data includes any combination of the open, high, low or close over a period of time. Some indicators may use only the closing prices, while others incorporate volume and open interest into their formulas. The price data is entered into the formula and a data point is produced. A technical indicator offers a different perspective from which to analyze the price action. Indicators serve three broad functions: to alert, to confirm and to predict.

The Technical Indicators used in this papers are:

SMA--Simple Moving Average: A simple, or arithmetic, moving average that is calculated by adding the closing price of the security for a number of time periods and then dividing this total by the number of time periods. A simple moving average is formed by computing the average price of a security over a specific number of periods

EMA--Exponential Moving Average Calculation: EMA reduce the lag by applying more weight to recent prices. The weighting applied to the most recent price depends on the number of periods in the moving average. There are three steps to calculating an exponential moving average. First, calculate the simple moving average.

MACD--Moving Average Convergence-Divergence: MACD indicator is one of the simplest and most effective momentum indicators available. The MACD turns two trend-following indicators, moving averages, into a momentum oscillator by subtracting the longer moving average from the shorter moving average. As a result, the MACD offers the best of both worlds: trend following and momentum.

Genetic Algorithm

The Genetic Algorithm was proposed in 1975 and its framework is based on a direct analogy to Darwinian natural selection and mutations in biological reproduction [2]. It belongs to a category of heuristics known as the stochastic method, which employs randomized choice operators in the search strategy [3]. The appeal of GAs comes from their simplicity and elegance as strong search algorithms, as well as their ability to discover good solutions rapidly for difficult high-dimensional problems. The genetic algorithm is a popular method which has been applied in different data mining tasks, such as clustering [4]. Selection, crossover, and mutation are the three major GA operations.

The genetic algorithm can be summarized as:

Randomly generate Initial population;
Evaluate fitness of each chromosome in the population;

While (result doesn't achieve the goal)
{
Perform selection operation;
Perform crossover operation;
Perform mutation operation;
Evaluate fitness of each chromosome in the population;
}

3. PROPOSED SYSTEM

The diagram below (Fig. 2) Gives the overview of the project. The numerical Stock Market data: DJIA is being used. Stock trading data contains the quotes of stock market. From this information, many technical indicators can be extracted, and by investigating the relations between these indicators trading signals can discovered. Genetic algorithm is being used to generate all the relations among the technical indicator and its value. Along with genetic algorithm association rule mining algorithm is used for generation of association rules among the various Technical Indicators. Associative rules are generated whose left side contains a set of trading signals, expressed by relations among the technical indicators, and whose right side indicates whether there is a positive negative or no change. The rules are being further given to the classification process which will be able to classify the new data making use of the previously generated rules.

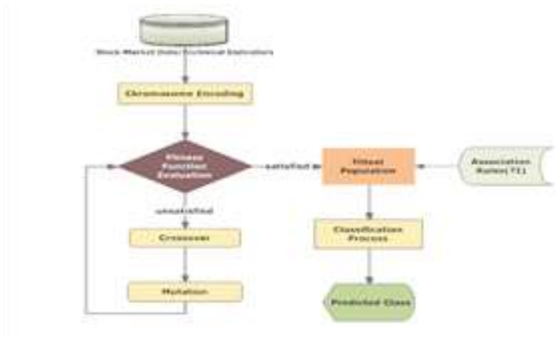


Figure 1: Genetic Algorithm

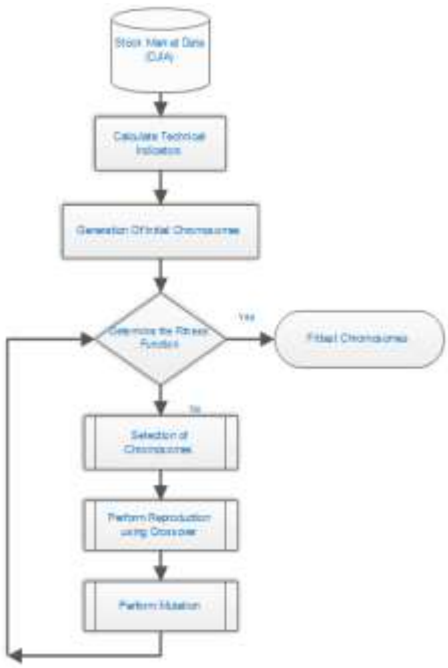


Figure 2: Flow of the proposed system

Genetic Algorithm

The genetic algorithm starts with the generation of initial chromosomes. Here, the chromosomes are represented as 64bits. The chromosome structure is as mentioned below:

RSI	EMA	MAC	K	ROC	CCI	Willi	am%	LIBO

Where each of the technical indicators mentioned is assigned 8bits each. Relations generated from the technical indicators are as shown in Table 1.

The last step is the mutation function wherein specific amount of chromosomes which are being determined by mutation rate are undergone mutation where one random bit of the chromosome is being flipped if this results in a new chromosome then it's being added to the final chromosome list. Now the final new chromosomes list in turn gives us the rules or relations of various technical indicators. This is being given as an input to the next module.

Table 1: Technical Indicators relations

Technical Indicators	Rules
RSI	RSI>70 , RSI <30
EMA	EMA12>adj close, EMA26>adj close, EMA12>EMA26, EMA26>EMA12, EMA12(t)>EMA26(t-1), EMA26(t)>EMA12(t-1), EMA12 (t) > adj close (t-1), EMA26 (t) > adj close (t-1)
MACD	MACD>0, MACD>100, MACD>K, MACD>ROC, MACD(t) > K(t-1), MACD(t)>ROC(t-1), MACD>adj close, MACD(t)>adj close(t-1)
	K>70, K(t-1)>70, K<30, K(t-1)<30, K(t-1)>K(t), K(t)>D(t-1), K(t-1)>D(t), K(t-1)>D(t-1)
ROC	ROC>3, ROC<-1, ROC>K, ROC(t)>k(t-1)
CCI	CCI>100, CCI>high, CCI>Low, CCI>adj close, CCI(t)>Adj close(t-1), CCI<-100
Williams %R	%R >ROC, %R>ROC(t-1), %R>-30, %R<-70, %R>adj close, %R>adj close(t-1)
LIBOR	LIBOR>0.7

Association Rule Mining Algorithm

Figure 3: Finding final set of frequent rules

The Apriori Algorithm takes in the rules and generates all possible combination of frequent rules sets. During each iteration rules are being assigned to each of the class label and then the support count of the rule along with the class is being calculated. All the rules class label set is

being pruned which are below the Support and confidence threshold.

The support and confidence is being calculated as below:

Support (rule⊚class)=count(rule⊚class/ |total number of chromosome| *100

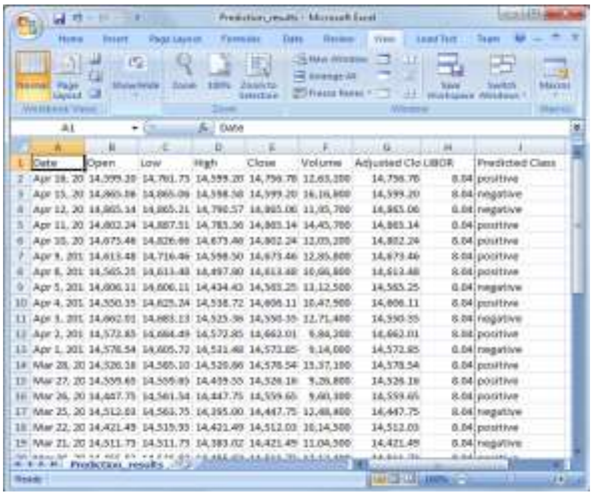
Confidence = support(rule@label)/ support (rule)

At the end of the module we are left with the all possible frequent combination of the rule set along with the class label. These sets of frequent rule will help us in classification of the new test data.

Prediction

Figure 4 Flowchart for Prediction Process

In the Classification step, the new test data is being collected and is being given to the genetic algorithm to generate the initial set of chromosome. The chromosomes obtained will help us in obtaining the rules which can be used to predict the class label. The rules obtained from the new set of chromosomes are mapped against the rules obtained from the previous step i.e. association rule mining algorithm.



The Prediction performance is being calculated by comparing the actual results to the predicted ones.

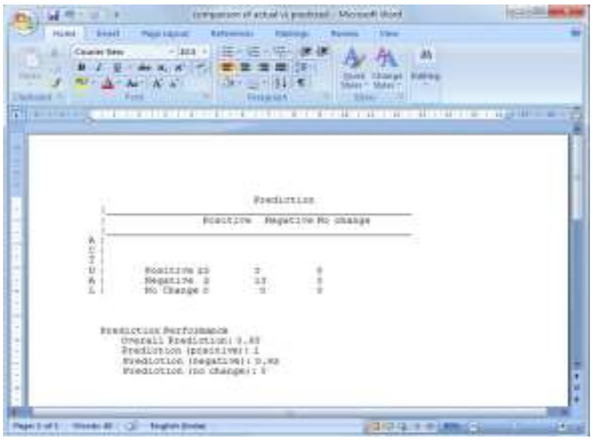


Figure 5: Prediction Performance Results

The above figure gives the prediction performance of the application as compared to the actual results. From this it can be concluded that the overall prediction performance was 95%. Individually, The Positive Class Label showed 100% accuracy while Negative Class Label had 98% accuracy.

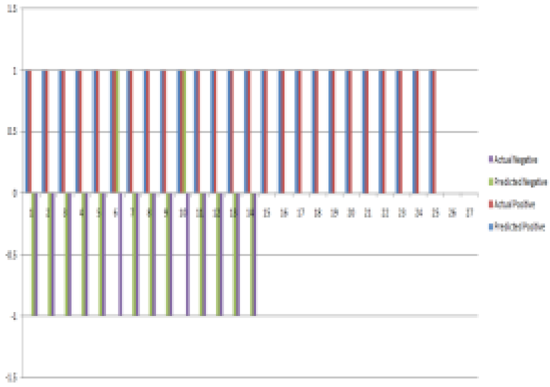


Figure 6: Prediction Performance Graph

The above Figure shows the prediction performance graph, which shows the predicted positive to the actual positive results and the predicted negative to the actual negative.

4. CONCLUSION

A detailed Study was conducted on the DJIA stock market and the various technical indicators being used for Stock Market data which help in analysing the Stock Market. The Literature survey was conducted on Genetic Algorithm and Association rule Mining Algorithm and a combine approach was being implemented. As Association Rule Mining Algorithms cannot handle numerical data efficiently a modified Genetic Algorithm was being used for the representation of the stock market data and to generate rules among the various technical indicators. Association rule mining Algorithm help in generation of the frequent rule set along with the class label and then predict the class label for the new test data i.e. if it is positive, negative or no change.

Overall a new method was proposed for Stock Market Prediction using a combination of Genetic and Association rule Mining Algorithm which can handle numerical data.

The Prediction Performance was also being calculated and a comparison was carried out with the actual performance. The overall prediction process had 95% of accuracy.

REFERENCES

1) Chen, S. H. Genetic Algorithms and Genetic Programming in Computational Finance. Boston, MA: Kluwer.2002
2) Neely, C., Weller, P. & Dittmar, R. Is Technical Analysis in the Foreign Exchange Market Profitable? A Genetic Programming Approach. Journal of Financial and Quantitative Analysis. 32:405-26.1997.
3) Thomas, J. & Sycara, K. The Importance of Simplicity and Validation in Genetic Programming for Data Mining in Financial Data. Proceedings of the Joint AAAI-1999 and GECCO-1999 Workshop on Data Mining with Evolutionary Algorithms. 1999.
4) Allen, F. & Karjalainen, R. Using Genetic Algorithms to Find Technical Trading Rules. Journal of Financial Economics. 51:245-271. Studies 27:221-334. 1999.
5) Robert Edwards, John Magee. Technical analysis of stock trends. Seventh edition, page 4.1997.
6) Ryan Sullivan, Allan Timmermann, Halbert White. Data-snooping, Technical Trading Rule Performance, and the Bootstrap. The Journal of Financial, 54, (5):1647-1692.1999.
7) M. Mitchell. An Introduction to Genetic Algorithms. MIT press, Cambridge, USA. 1996.
8) Packard, N., "A genetic learning algorithm for the analysis of complex data," Complex Systems, 4, 1990, pp.543-572.
9) Rutan, E., "Experiments with optimal stock screens," Proceedings of the 3rd International Conference on Artificial Intelligence Applications on Wall Street, 1993, pp.269-273.
10) Syswerda, G., "Uniform crossover in genetic algorithms," In Schaffer, J.D. (Eds.), Proc. 3rdInt.Conf. Genetic Algorithms. San Maeto, CA: Morgan Kaufmann, 1989.
11) Trippi, R. R. and DeSieno D., "Trading equity index futures with a neural network," The Journal of Portfolio Management, 1992.