



Stock Market Trend Prediction Model Using Data Mining Techniques

Oyelade Iyinoluwa* and Adewale Olumide Sunday

Department of Computer Science, Federal University of Technology, Nigeria

*Corresponding author: Oyelade Iyinoluwa, Department of Computer Science, Federal University of Technology, Akure, Nigeria

Received:  November 26, 2019

Published:  December 12, 2019

Abstract

Stock market prediction is essential and of great interest because successful prediction of stock prices may promise smart benefits. These tasks are highly complicated and very difficult. Many researchers have made valiant attempts in data mining to devise an efficient system for stock market movement analysis. This research has developed an efficient approach to stock market trend prediction by employing Frequent Pattern growth and Fuzzy C-means clustering algorithms. This research has been encouraged by the need of predicting the stock market to facilitate investors about when to buy, sell or hold a stock in order to make profit. Firstly, the original stock market data were converted into interpreted historical (financial) data via technical indicators. Based on these technical indicators, datasets that are required for analysis was created. Subsequently, Frequent Pattern Growth algorithm was used to generate frequent patterns. Based on these frequent patterns, Fuzzy C-means clustering technique was used to formulate the prediction model. Finally, a classification technique, K-Nearest Neighbor classifier was employed to predict the stock market trends. The results from the stock market trend prediction were validated through Hit ratio evaluation metric to estimate the prediction accuracy. Comparative analysis was carried out for the proposed model and a neural network model was used to benchmark the proposed model. The obtained results showed that proposed model produced better results than the neural network model in terms of accuracy. This paper has provided a novel approach which combines FP-Growth, Fuzzy C-means and K-Nearest Neighbor algorithms for stock market trend prediction.

Keywords: Fuzzy C-means Algorithm; FP-Growth Algorithm; K-Nearest Neighbor Algorithm; Stock Market Prediction

Introduction

Stock market prediction is one of the most trending topics in finance and business. However, the unpredictable nature of the stock market creates a problem for investors to do profitable investments. Several research efforts have been carried out to predict the market in order to make profit using different techniques ranging from statistical analysis, technical analysis and fundamental analysis among others, with different results. These techniques however cannot provide deeper analysis that is required and therefore not effective in predicting stock market trends. However, finding patterns in stock market can provide insight into market behavior, buying or selling habits and co-movement of stock shares. This research aims at creating an approach to discover inference knowledge from the relationships among stock index indicators to provide useful information about market trends for investment decisions Ehsan et al. [1] defined stock market as a private or public market for the trading of company stock and derivatives

of company stock at an agreed price; these are securities listed on a stock exchange as well as those only traded privately. Stock market is very volatile in nature and prices of stocks change almost instantly. It strongly depends on demand and supply. The prices will be high when the demand is high, and the prices will be low when the demand is low [2]. Uncertainty is the main characteristic of all stock markets, which is related to their future state. This feature is undesirable and unavoidable for the investor whenever stock market is selected as the investment tool. Predicting the stock market is the best option to reduce uncertainty. Stock market prediction includes uncovering market trends, planning investment, investment strategies, determining the perfect time to purchase the stocks and what stocks to purchase [1]. Many factors may influence the stock market including political events, general economic conditions and traders' expectations [3]. Therefore, predicting price movement on stock data seems to be quite difficult.

Due to technological development in information technology and software abilities, gathering sufficient data on stock traded by the hour and even by the minute have become much easier. This is the reason why data mining techniques have attracted the attention of investors interested in predicting the trends of the stock market. According to Ravindranath [4], Data Mining is the process of selecting, exploring and modeling large quantities of data to unravel previously unknown patterns for the purpose of business and commercial advantage. It allows users to analyze data, categorize it, and summarize the relationships among data Ehsan et al. [1] defined data mining as the science and technology of exploring data in order to discover previously unknown patterns. Finding frequent patterns plays an essential role in mining associations, correlations, and many other interesting relationships among data. Moreover, it helps in data indexing, classification, clustering, and other data mining tasks as well. Frequent pattern mining is an important data mining task and a focused theme in data mining research. It was first introduced by Agrawal et al. [5] to find frequently purchased items by customers in a "market basket analysis". It became more popular due to its high applicability in various data analysis fields such as DNA pattern recognition, web data mining, clinical data mining, software bug analysis and stock market analysis. Apriori and FP-Growth algorithms are two most frequent pattern mining algorithms in literature [6]. Apriori algorithm which was introduced by Agarwal and Srikant [7] adopts the apriori property and candidate generation process to generate association rules. FP-Growth is a tree-based approach introduced by Han et al. [8] which follows a two-step procedure; it scans the database once and generates the Frequent Pattern tree in the first step and then it discovers frequent patterns from the tree. Apriori and FP-growth algorithms generate huge number of frequent patterns which are not able to produce direct knowledge or inference. Fuzzy Logic can play an essential role in discovering inference knowledge from huge number of frequent patterns by using Fuzzy Inference. Fuzzy inference is the process of formulating the mapping from a given input to an output using fuzzy logic. This paper adopts the fuzzy c-means clustering approach for matching facts with frequent patterns which performs excellently in either exact or partial matching to generate inference knowledge. The inference knowledge generated helps investors to know when to buy, sell or hold a stock. Therefore, the main purpose of implementing the fuzzy inference approach for association rule mining on stock market data is to discover inference knowledge by analyzing the random data and to use the knowledge as a reference during decision making.

Related Works

Stock market prediction is considered to be a challenging task for both investors and researchers, due to its profitability and intricate complexity. The predictability of the market is an issue that has been much discussed by researchers and academics. Several researches have been done for predicting the price direction of a stock using

different data mining techniques and have emerged with various results Sachin et al. [2] worked on an association rule mining model for finding the interesting patterns in stock market dataset. The objective was to predict if the stock prices of a specific company will be up or down by the next day. The work used apriori algorithm for generating frequent patterns from the stock data and rules were generated from the frequent patterns. However, the prediction was not accurate because the rules generated were too ambiguous. Krittithee and Pakorn [9] worked on association rule mining on stock index indicators. In their work, a performance study was done using association rule mining to model index indicators and trading volume and their effect on price change. The work only described the relationship between stock indexes and how they affect price change but did not go further to give conclusion on the state of the particular stock. Shubhangi and Nandgaonkar [10] proposed Using Association Rule Mining: Stock Market Event Prediction from Financial News. Their work considered only the closing price of stocks and it was used to calculate the stock index indicators and for defining the rule. Naive Bayes algorithm was used for training the technical index indicators and association rule mining was used for generating buy, sell or hold signals. However, the accuracy of the model was low because they only considered one field of the stock data among five which is closing price. However, relations among stock market data are not sufficient for predicting the trends in the market. This paper improves on the reviewed related works by modeling a fuzzy clustering mechanism for predicting the future price direction of a stock.

Methodology of the Proposed Model

The proposed model adopted frequent pattern growth algorithm to find frequent patterns among stock index indicators and also adopted Fuzzy C-means clustering approach to develop a prediction model using the frequent patterns dataset. Historical data of a Nigerian bank for three years were obtained online from a stock broking firm with the Nigerian Stock Exchange. The dataset represented in Table 1 contains fields such as such as open price, close price, high price, low price and volume for each trading day. Five technical indicators such as Relative Strength Index (RSI), Rate of Change (ROC), Moving Average Convergence Divergence (MACD), Stochastic Oscillator (SO) and On Balance Volume (OBV) were calculated from the historical dataset in. The parameters for the technical indicators are specified according to the standard default guidelines. The data points were converted into symbols, that is, numeric-to-symbolic conversion to reduce the large number of distinct values. These indicators are described below.

Moving average convergence divergence (MACD)

The MACD is the difference between these two moving averages plotted against a centerline. In this work, the most common moving average values are used in the calculation. These are the 26-day and 12-day exponential moving averages (EMA). The signal line used is zero.

Table 1: Sample of raw dataset.

Date	Close	Open	High	Low	Volume
3/2/2014	28.4	27.45	28.49	27.45	17,491,810
4/2/2014	27.52	28.94	28.02	27.52	15,829,780
5/2/2014	27.5	27.45	27.71	27.45	23,873,228
6/2/2014	27.32	27.9	27.5	27.32	6,564,140
7/2/2014	27.48	27.5	27.5	27.4	15,182,202
-	-	-	-	-	-
25/01/2017	23.89	23.94	23.94	23.8	5,328,624
26/01/2017	24.5	23.89	24.5	23.88	3,495,576
27/01/2017	24.35	25.6	24.5	24.32	2,469,043
30/01/2017	23.97	23.85	23.97	23.7	2,934,064
31/01/2017	23.7	24.35	23.9	23.57	5,636,309

Formula:

$$MACD = EMA_{n_1} + EMA_{n_2}$$

$$EMA_{n_1} = \frac{2}{n_1 + 1} (C_{n_1} - EMA_{n_1-1}) + EMA_{n_1-1}$$

$$EMA_{n_2} = \frac{2}{n_2 + 1} (C_{n_2} - EMA_{n_2-1}) + EMA_{n_2-1}$$

Where n_1 is 12 days, n_2 is 26 days, C_{n_1} is the closing price at day 12 and C_{n_2} is the closing price at day 26. The trading rules of the MACD to be used are:

- I. IF MACD is above the signal line then buy.
- II. IF MACD is below the signal then sell.

Relative strength index (RSI)

The Relative Strength Index (RSI) considers whether an asset is over bought or oversold. RSI is described in equation 3.4.

Formula:

$$RSI = 100 - \frac{100}{1 + RS}$$

$$RS = \frac{\text{Average gain of } n\text{-day up periods}}{\text{Average loss of } n\text{-day down periods}}$$

Where n is the number of days. The default number of days used by analyst is 14 days which will be adopted in this work. The trading rules to be used are:

- i. IF RSI increases to above 70 (implies overbought) then sell.
- ii. IF RSI is between 30 and 70 (implies normal) then hold.
- iii. IF RSI decreases to below 30 (implies oversold) then buy.

Rate of change (ROC)

ROC is a momentum oscillator that measures the percentage of change in price. It calculates the percentage difference between the closing price of the current day and the closing price of days.

Formula:

$$ROC = \left[\frac{C_t - C_n}{C_n} \right] \times 100$$

where n is the number of trading days which is 12 days, C_t is closing price of current day, C_n closing price of 12 days ago. The trading rules to be used are:

- i. IF ROC is negative then sell.
- ii. IF ROC is positive then buy.
- iii. IF ROC is zero then hold.

Stochastic oscillator (STO)

The Stochastic Oscillator gives an indication of the stock's last closing price relative to the stock's recent trading range [11]. The STO is plotted within a range of 0 to 100 and signals over-bought conditions above 80 and oversold conditions below 20 [12]. STO is described in the equation below.

Formula:

$$STO = \left[\frac{C_t - L_n}{H_n - L_n} \right] \times 100$$

where n is the number of trading days which is 14 days, C_t is the closing price of current day, L_n is the lowest price over 14 days and H_n is highest price over 14 days. The trading rules of the stochastic oscillator are summarized as follows.

- i. IF STO increases above 80 (implies overbought) then sell.
- ii. IF STO is between 20 – 80 (implies normal) then hold.
- iii. IF STO is below 20 (implies oversold) then buy.

On-balance volume (OBV)

The On-Balance Volume indicator reflects movement in volume of stocks. The OBV is calculated by taking the total volume for the trading period and assigning it a positive or negative value depending on whether the price is up or down during the trading period. OBV is described in the equation below.

Formula:

If Today's close is greater than Yesterday's close, then:

$$OBV = Yesterday's\ OBV + Today's\ Volume$$

Else, if today's close is less than yesterday's close then:

$$OBV = Yesterday's\ OBV - Today's\ Volume$$

Else, if today's close is equal to yesterday's close then:

$$OBV = Yesterday's\ OBV$$

the On-Balance Volume are summarized as follows:

- i. IF OBV is upward then buy
- ii. IF OBV is downward then sell
- iii. IF OBV is the same then hold

The indicators for each trading day are calculated and a sample is presented in Table 3 together with the trading volume for each day. They are referred to as features and serves as the input dataset. The indicators for each trading day are calculated and a sample is presented in Table 2 together with the trading volume for each day. They are referred to as features and serves as the input dataset.

Conditions applied

Yesterday's OBV is yesterday's volume. The trading rules of

Table 2: Sample of technical Indicator Dataset.

Date	MACD	RSI	ROC	SO	OBV
3/2/2014	0.13	51.59	1.18	67.63	17,491,810
4/2/2014	0.07	42.98	-3.44	25.12	15,829,780
5/2/2014	0.02	42.8	-2.62	24.15	23,873,228
6/2/2014	-0.03	41.18	-5.14	15.46	6,564,140
7/2/2014	-0.06	43.24	-2.73	23.19	15,182,202
-	-	-	-	-	-
25/01/2017	0.07	50.82	-2.89	20.71	4,379,942
26/01/2017	0.11	58.2	3.99	81.03	43,246,896
27/01/2017	0.13	55.98	1.46	69.3	34,907,437
30/01/2017	0.12	50.69	-0.04	28.72	13,928,281
31/01/2017	0.08	47.28	-0.84	13.98	28,607,226

Table 3: The range of each Indicator.

MACD		RSI		ROC		STO		OBV	
aMACD	< 0	aRSI	>= 70.00	aROC	< 0	aSTO	>= 80	aOBV	downward
bMACD	> 0	bRSI	<= 30.00	bROC	> 0	bSTO	<= 20	bOBV	upward
cMACD	0	cRSI	30.01- 69.99	cROC	0	cSTO	20.01 - 79.99	cOBV	statics

Model Architecture

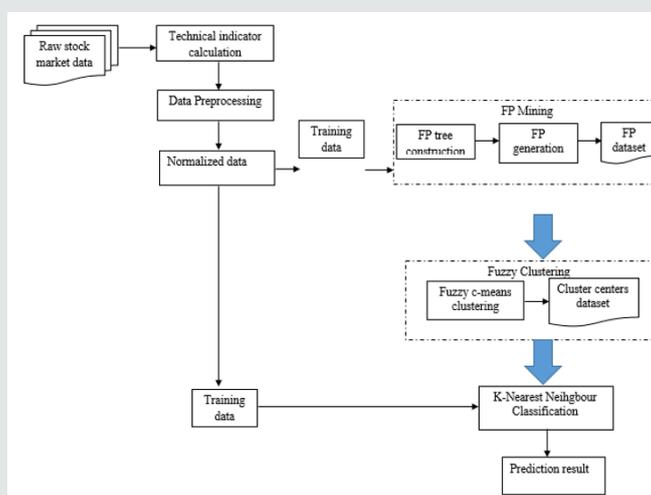


Figure 1: Stock Prediction Model Architecture using Data Mining Technique.

Figure 1 shows the architecture of the Stock Market Trend Prediction Model. The different components that make up the model work together to achieve the model objective. Data preprocessing is in two phases. First, the technical indicators are calculated from raw data retrieved from the internet. The technical indicator dataset formed the dataset required for mining. Secondly, data discretization reduces the large distinct values by dividing the data record into three ranges. Frequent pattern mining is performed on the preprocessed data to retrieve frequent patterns. The frequent pattern dataset is clustered using Fuzzy C-Means clustering algorithm to extract three cluster centers and K-Nearest Neighbor algorithm is used as the classification technique to classify new data points to the cluster centers in order to predict the trends of the new data points.

Data Preprocessing

Time series data is difficult to manipulate, but when they are treated as symbols (item units) instead of data points, interesting patterns can be discovered, and it becomes an easier task to mine them. Thus, it is suggested to convert the basic unit into symbols, i.e., numeric-to-symbolic conversion to reduce the large number of distinct values. The numeric-to-symbolic conversion transforms the available features (e.g. MACD, RSI, ROC, STO and OBV) of a financial instrument into a string of symbols. Each indicator is divided into three ranges (a, b, c) as shown in Table 3.

Frequent Pattern Mining

In order to mine the mapped dataset to discover frequently occurring features, frequent pattern mining has to be performed on preprocessed data. Under this process, frequent pattern growth algorithm reads the mapped dataset and generates frequent patterns as per predefined minimum threshold value and store in the frequent pattern base.

The step by step procedure for constructing a FP-tree is stated below.

Step 1: The mapped dataset is scanned once to determine the support count of each feature. Infrequent features are discarded, while the frequent features are sorted in decreasing support counts.

Step 2: Create the root of an FP-tree T and label it as null. For each trading day in the database, select the frequent items and sort them according to the order of the frequent item list.

Step 3: Let the sorted frequent item list be $[p|P]$ where p is the first element and P is the remaining list. Call $insert_tree([p|P], T)$. The function is performed as follows. If T has a child N such that $N.item-name = p.item-name$, then increment N 's count by 1, else create a new node N , with its count initialized to 1, its parent link linked to T and its node-link linked to the nodes with the same item-name via the node-link structure.

Step 4: If P is nonempty, call $insert_tree(P, N)$ recursively

After the FP-tree is constructed, frequent itemset are mined from it. Minimum support was experimented from 0.1 to 1. 0.1-0.5 had a lot of uninteresting patterns. 0.6 to 1 had concise and interesting patterns but the numbers of patterns generated for each 0.6 to 1 minimum support were almost equal and did not much effect on the result. Therefore, 0.6 was selected as the minimum support for this paper.

Input: The transaction database represented by the constructed FP-tree and a minimum support of 0.6.

Output: The complete set of frequent patterns.

Method: call FP-growth (FP-tree, null)

Procedure FP-growth (Tree, α)

- a) If tree contains a single prefix path, then
- b) Let P be the single prefix-path of tree
- c) Let Q be the multipath part with the top branching node replaced by a null root
- d) For each combination (denoted as β) of the nodes in the path P do
- e) Generate pattern $\beta \cup \alpha$ with support = minimum support of nodes in β
- f) Let freq-pattern-set(P) be the set of patterns generated
- g) Else let Q be Tree
- h) For each item a_i in Q do
- i) Generate pattern $\beta = a_i \cup \alpha$ with support = a_i support
- j) Construct β 's conditional pattern-base and then β 's conditional FP-tree $tree_\beta$
- k) If $tree_\beta \neq \emptyset$
- l) Then call FP-growth (Tree, β)
- m) Let freq-pattern-set (Q) be the set of patterns generated
- n) Return (freq-pattern-set (P) \cup freq-pattern-set(Q) \cup freq-pattern-set(P) * freq-pattern-set(Q))

Fuzzy C-Means Clustering Algorithm

Fuzzy c-means (FCM) algorithm is applied on the frequent pattern dataset. Each data point is classified into the cluster with the highest membership value. FCM is based on the minimization of the following objective function. The objective function is calculated in each iteration.

$$O_m = \sum_{i=1}^D \sum_{j=1}^N \mu_{ij}^m \|x_i - c_j\|^2$$

Where D is the number of data points, N is the number of clusters, x_i is the i^{th} data point, C_j is the center of the j^{th} cluster, μ_{ij} is the degree of membership of x_i in the j^{th} cluster. For a given data point, x_i , the sum of the membership values for all clusters is one.

FCM performs the following steps during clustering.

Step 1: The cluster membership value u_{ij} is initialized to 0

Step 2: The predefined number of clusters N is assigned. The number of clusters is 3 and each cluster represents each trend which are uptrend, downtrend and static trend.

Step 3: The cluster centers of each cluster is calculated as follow.

$$C_j = \frac{\sum_{i=1}^D \mu_{ij}^m x_i}{\sum_{i=1}^D \mu_{ij}^m}$$

Step 4: The degree of membership of i^{th} data point in each cluster is calculated and updated as follows.

$$\mu_{ij} = \left[\sum_{k=1}^N \left(\frac{\|x_i - C_j\|}{\|x_i - C_k\|} \right)^{\frac{2}{m-1}} \right]^{-1}$$

Step 5: The objective function O_m is calculated.

Step 6: Step 3 to 5 is repeated until O_m improves less than the specified minimum threshold.

Model Implementation

This section describes the analysis and implementation results of the Fuzzy clustering-based model for stock market prediction. The implementation of this model follows the following steps.

1. Frequent patterns are generated from the pre-processed

Table 4: Cluster Centers and their corresponding trend.

Technical Indicators						
Cluster Centre	MACD	RSI	ROC	STO	OBV	Corresponding Trend
Cluster Centre for Cluster 1	0.17	33.8	1.2	27	-2548244.28	Static Trend
Cluster Centre for Cluster 2	0.16	29.5	3.11	55.5	147866291.1	Uptrend
Cluster Centre for Cluster 3	-0.04	21.2	-8.48	9.96	-183110835.1	Downtrend

Result and Discussion

The performance of the model was studied using financial data of three banks from the period of September 2017 to January 2018. The dataset comprises of opening, high, low and closing prices as well as the traded volume of one Nigerian Bank stock. Thus, a total of 300 data points, that is 100 data points for each bank were available for testing. The five chosen technical indicators, MACD, RSI, ROC, STO and OBV were computed from the sampled data. A K Nearest Neighbor (K-NN) classifier using $k=1$ is adopted as the

dataset.

2. The frequent patterns are then clustered into three clusters.
3. K-Nearest Neighbour Classifier assigns new data points to the cluster centers.

The model parameters comprise of both the frequent pattern mining parameters and the fuzzy clustering parameters which are outlined below.

- I. Minimum support = 0.6
- II. Minimum support count = 60
- III. Number of clusters = 3

The total number of frequent patterns generated for minimum support of 0.6 is 94 frequent patterns.

The model has five inputs (MACD, RSI, ROC, STO and OBV) and one output (TREND). The data is plotted in 2D and clusters are computed using fuzzy c-means clustering function (fcm). Fuzzy C-Means clustering is an iterative process. First, the initial fuzzy partition matrix is generated, and the initial fuzzy cluster centers are calculated. In each step of the iteration, the cluster centers and the membership grade point are updated, and the objective function is minimized to find the best location for the clusters. The process stops when the maximum number of iterations is reached, or when the objective function improvement between two consecutive iterations is less than the minimum amount of improvement specified. The three cluster centres achieved are elaborated in Table 4. Each cluster centre is a data point and it corresponds to a trend. Data points that fall into cluster 1, 2 and 3 have static trend, uptrend and downtrend respectively.

classification technique. The cluster centres are the classes that new data points will be assigned to. Each new data point is classified by a majority vote of its neighbor with the data point assigned to the class of one nearest neighbor. Static trend is assigned 1, uptrend is assigned 2 and downtrend is assigned 3. Uptrend means a price fall which is an indication to buy a stock. Downtrend means price is high and rising which is a good time for a trader to sell stock. However, when the trend is static, that is price did not go up neither did it fall, the best decision to take at that point is to hold stock.

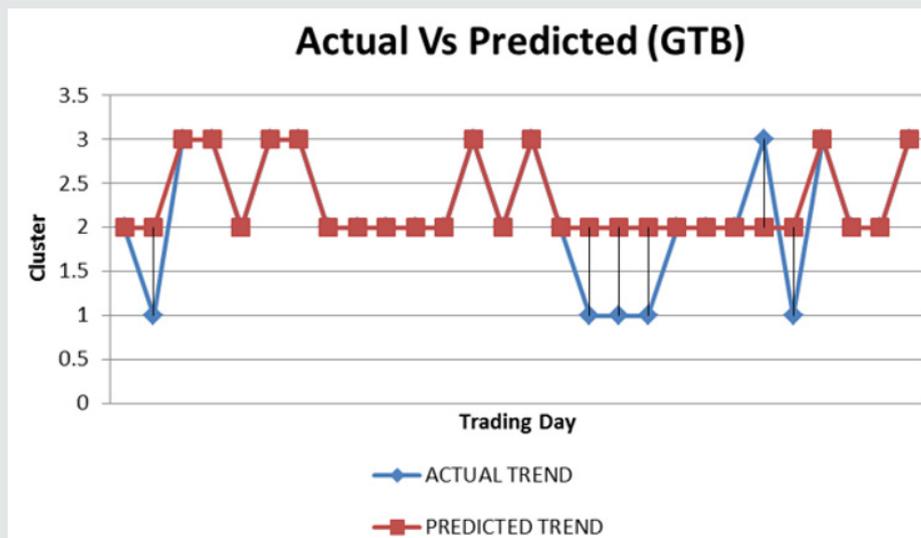


Figure 2: Actual Vs Predicted curve for GTB.

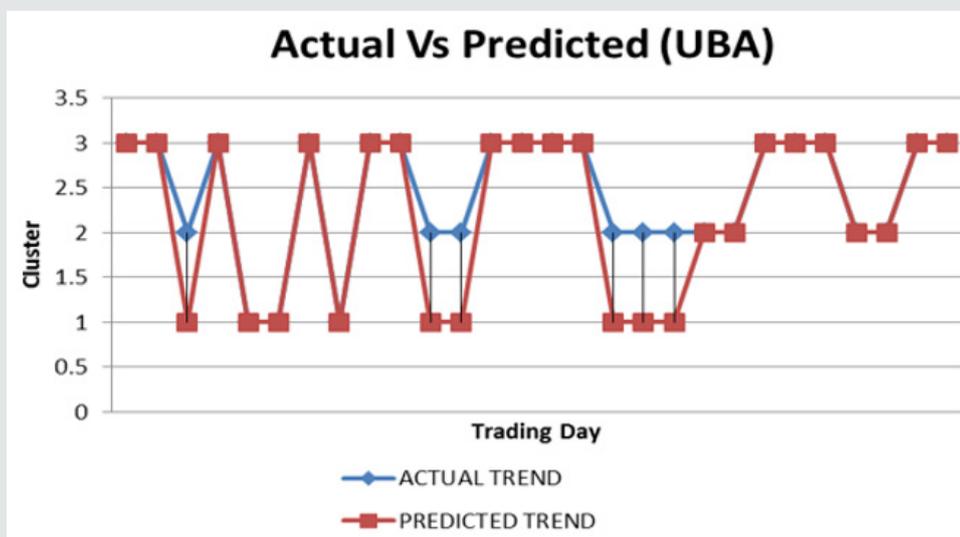


Figure 3: Actual Versus Predicted curve for UBA dataset.

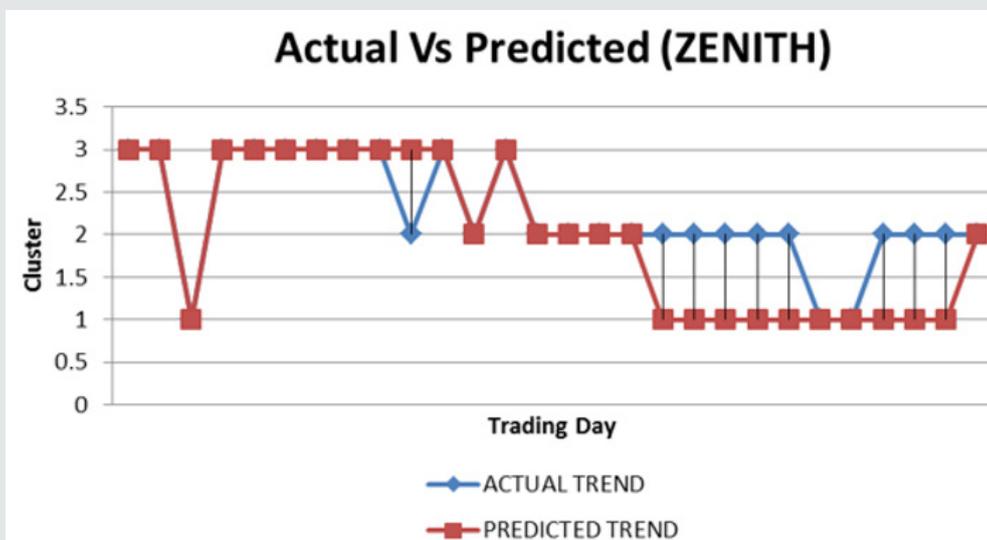


Figure 4: Actual Versus Predicted curve for ZENITH dataset.

- For GTB dataset, 17 data points misclassified out of 100 data points. Therefore, the accuracy of prediction is 83%.
- For UBA dataset, 89 data points classified correctly while 11 are misclassified. Therefore, the accuracy is 89%.
- For ZENITH dataset, 82 data points classified correctly while 18 are misclassified. Therefore, the accuracy is 82%.

Table 5 summarizes the number of data points that were classified under each cluster for each bank. For GTB and UBA,

Table 5: Classification of data points per cluster.

	GTB	UBA	ZENITH
Cluster 1	-	13	13
Cluster 2	64	46	41
Cluster 3	36	41	46

Performance Evaluation

The evaluation criteria used in this paper is Hit Ratio and its expression is shown in equation 13.

$$Hit\ ratio = \frac{1}{N} \sum_{i=1}^N H_i$$

Where

$$H_i = 1 \text{ if } pred_i = Act_i \text{ otherwise } 0.$$

$pred_i$ is the predicted output value,

cluster 2 has the highest number of data points classified to it which corresponds to uptrend. This means that there were generally more uptrends during the testing period, meaning that prices fell mostly within this period, hence buying GTB and UBA shares during this period would have been more profitable than selling. However, cluster 3 has the highest classification for ZENITH which corresponds to downtrend. This means that prices rose more during this period, hence selling ZENITH shares would be more profitable than buying.

Act_i is the actual output value

N is the number of test data points. N is 100 for each bank.

Table 6 and Figure 5 shows the hit ratio of each bank and also the number of 1's and 0's for each bank.

Table 6: Hit ratio for each bank.

Bank	1's	0's	Hit ratio
GTB	83	17	83 out of 100
UBA	89	11	89 out of 100
ZENITH	82	18	82 out of 100

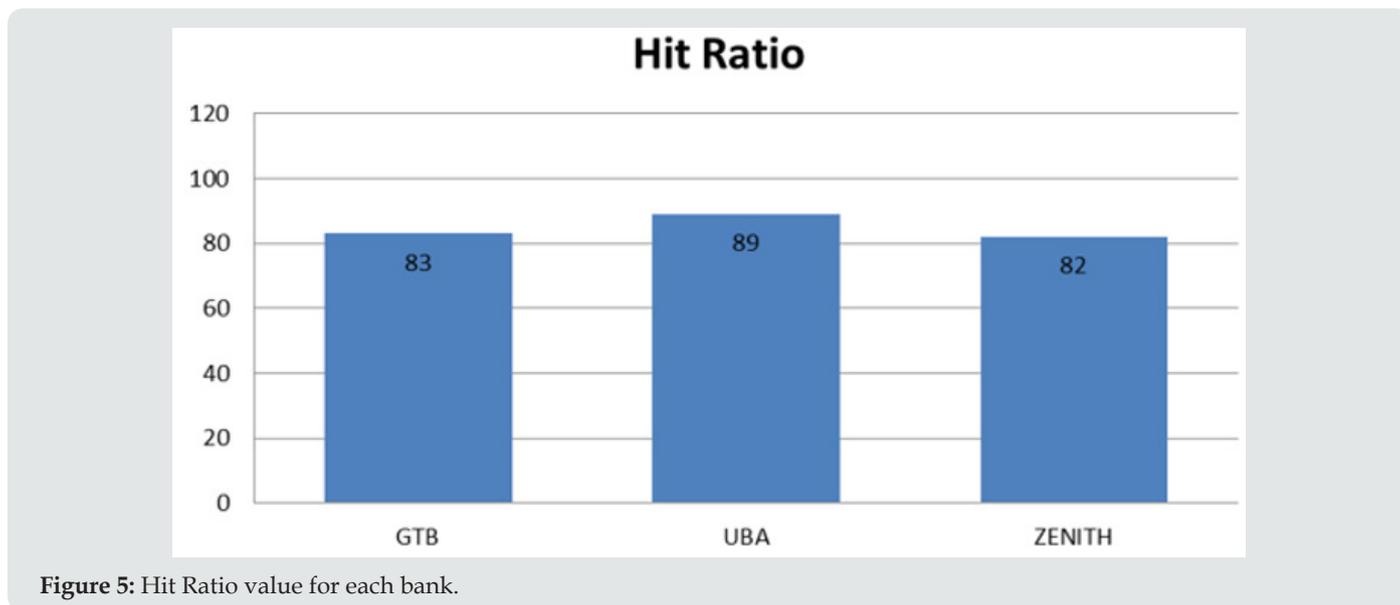


Figure 5: Hit Ratio value for each bank.

Therefore, the model has an average accuracy of 84.67%.

Conclusion

This paper has proposed a novel approach for predicting stock market trends using frequent pattern mining, fuzzy clustering and k-nearest neighbor algorithms. A total of 742 data points were obtained online. Five technical indicators were chosen for this

paper. Based on these indicators, dataset required for analysis was created. The dataset was preprocessed and then mined for frequent patterns. By choosing a minimum support 0.6, 94 frequent patterns were obtained. Fuzzy C-means clustering algorithm was applied to the frequent pattern dataset and three cluster centres were obtained and interpreted to represent either uptrend, downtrend or static trend. The performance of the model was studied using

financial data of three banks. A total 100 data points for each bank were available for testing. The five chosen technical indicators, MACD, RSI, ROC, STO and OBV were computed from the sampled data. K-Nearest Neighbour model using $k = 1$ was adopted to test the model. The cluster centres are the classes that new data points will be assigned to and each new data point was classified by a majority vote of its neighbour with the data point assigned to the class of one nearest neighbour. The predicted trends were compared with the actual trends to those that classified correctly and otherwise. The work was benchmarked with a neural network model that used the following parameters; 668 samples of training data, 100 samples of testing data, 5 inputs, 10 hidden neurons, 3 output layers and 40 iterations. The comparison shows that this research work performed better than the neural network model.

At the end of this research;

a) A stock market trend prediction model that can predict satisfactorily the stock market trend using a combination of five technical analysis indicators was developed.

b) The prediction model showed an average prediction accuracy of 84.67%.

Results from the prediction can help in the decision making of individual investors of when to buy, sell or hold a stock.

Future Work

Based on the result obtained from this paper, the model has shown an efficient way of predicting real time stock trends and assisting decision making of investors. In future research, more technical indicators can be included in the model. The experiment reported in this paper was based on modeling; future work can concentrate on developing a real time prediction system which will be easy to use by individual investors.

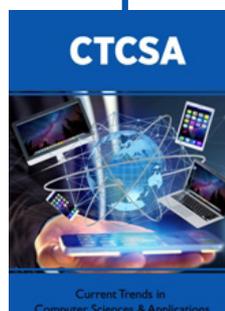
References

1. Ehsan H, Hamed D, Jamal S (2010) Application of Data Mining Techniques in Stock Markets: A Survey. *Journal of Economics and International Finance* 2(7): 109-118.
2. Sachin k, Shailesh J, Thakur RS (2014) An Association Rule Mining Model for Finding the Interesting Patterns in Stock Market Dataset. *International Journal of Computer Applications* 93(9): 10-20.
3. Karpio K, Lukasiewicz P, Orłowski A, Zabkowski T (2013) Mining Associations on the Warsaw Stock Exchange. *Proceeding of the 6th Polish Symposium of Physics in Economy and Social Science* 123: 553-599.
4. Ravindranath B (2003) *Decision Support Systems and Data Warehouses*. New Age International Limited.
5. Agrawal R, Imielinski T, Swami A (1993) Mining Association Rules between Sets of Items in Large Databases. *Proceedings of the ACM SIGMOD Conference on Management of Data* 22(2): 207-216.
6. Chaturvedi K, Patel R, Swami DK (2014) A Fuzzy Inference Approach for Association Rule Mining. *IOSR Journal of Computer Engineering* 16(6): 59-66.
7. Agrawal R, Srikant R (1994) Fast Algorithms for Mining Association Rules in Large Databases, *VLDB Conference*, Los Angeles, California, USA, pp. 487-499.
8. Han J, Kamber M, Tung AKH (2001) Spatial clustering methods in data mining: A survey.
9. Krittithae U, Pakorn L (2015) Association Rule Mining on Stock Index Indicators. *International Journal of Computer and Communication Engineering* 4(1): 46-49.
10. Shubhangi SU, Nandgaonkar SS (2013) Using Association Rule Mining: Stock Market Event Prediction from Financial News. *International Journal of Science Research (IJSR)* 4(6): 319-7064.
11. Hussein D, Pepe S (2002) Investment using Technical Analysis and Fuzzy Logic. *Fuzzy Sets and System* 127(2): 221-240.
12. Achelis SB (2000) *Technical Analysis from A to Z* (2nd ed.). NY: McGraw-Hill, New York, USA.

 This work is licensed under Creative Commons Attribution 4.0 License

To Submit Your Article Click Here: [Submit Article](#)

DOI: [10.32474/CTCSA.2019.01.000122](https://doi.org/10.32474/CTCSA.2019.01.000122)



Current Trends in Computer Sciences & Applications

Assets of Publishing with us

- Global archiving of articles
- Immediate, unrestricted online access
- Rigorous Peer Review Process
- Authors Retain Copyrights
- Unique DOI for all articles