International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8, Issue-1, May 2019

Prediction of Stock and Stock Price Index Movement using Protractible Fuzzy Particle Swarm Optimization

B. Sharmila, R. Khanchana

Abstract: Prediction in stock marketing is very important for analyzing the economic growth. Huge investors success is depending on the prediction of future market, where the forecasting is based on economic factors, technical factors and fundamental factors. Different countries depend on different factors, but India depends on all the three factors. Predicting the problem will give the major solution to the success. This paper focus to provide a solution by proposing a fuzzy logic based particle swarm optimization algorithm for predicting the movement of stock and price index of stock for the selected Indian based stock market companies, namely CNX Nifty, S&P BSE Sensex, Infosys, and Reliance. Performance metrics considered are sensitivity, specificity, precision, recall, accuracy and f-measure. The results show that the proposed algorithm outperforms the existing algorithms in all the terms.

Keywords: Prediction, Stock, Price, Index, Fuzzy

I. INTRODUCTION

Highlight Prediction of stock and its price index is always a challenging issue because of presence of ambiguity. There exist 2 kinds of investigation to analyze before making investment in stock marketing, which are fundamental investigation and technical investigation. In fundamental investigation, investors will review the stock value, industry, and economy. In technical investigation, investors will evaluate the stocks by making a statistical study by using previous prices and its impact in market. Technical investigators won't check or analyze intrinsic value, but stock charts are fully utilized in finding the patterns and trends for suggesting the behavior of stock in forthcoming days.

Revised Manuscript Received on 22 May 2019.

***** Correspondence Author

B Sharmila*, M.Sc., M.Phil, (Phd) scholar, has good interest in share market / mutual funds and investments.

Dr. R Khanchan, Phd, Asst. Professor, Sri Ramakrishna College of Arts & Science for Women, Coimbatore.

By making an effective hypothesis on stock market, the prices of the stock will be able to reach the good value, which represents the possibility towards predicting the stock prices. This is a logical concept related to many factors. In the past many years, several methodologies were proposed for the prediction of trends in stock marketing. In the beginning, the methodologies related to classical regression were utilized. It is possible to segregate the stock data into non-stationary time series data. Non-linear machine learning methods were utilized in the segregation process. ANN (Artificial – Neural - Network) and SVM (Support – Vector - Machine) are the two most widely utilized algorithms to predict stock and its price index movement. Every algorithm has its unique way of learning the patterns.

ANN is designed based on functions of human brain which builds a neuron network. Most proposals for this thrust research were proposed by assembling the genetic algorithm (GA), hidden markov algorithm (HMM) and ANN in order to estimate the future financial market. By utilizing ANN, prices of daily stock are converted to different sets of values which becomes the input for GA and HMM. SVM algorithm is designed on basis of hyper plane in preeminent dimension in the order of separating the classes. SVM is special type of machine learning algorithms identified by the scope of making decisions. It utilizes the kernel function and insufficiency of the solutions. SVM algorithm is used to prediction of future financial movement by making multiple investigations. Random forest algorithm is designed to classify the data by creating with the samples. It makes prediction based on the types of trees available in major. The accuracy of results depends on number of trees generated, where presence of more number of trees gives increased accuracy.

II. LITERATURE REVIEW

Deep Random Subspace [1] was proposed to predict the market of finance, which was considered as a fusion strategy. It makes predication based on analyzing the crowd and techniques used in the marketing, where it ensembles the deep learning and machine learning techniques. The results show that the prediction done by deep random subspace have low accuracy, which will be impossible to accept by the investors to make investment. Adaptive Fuzzy Inference System [2] was proposed to enhance the increase the performance of stock prediction, where fuzzy rules were applied dynamically to handle the non-linear arrival of data.

The algorithm works by verifying and calculating the values of parameter, and it finally generates the fresh fuzzy rules to observe the possibility for prediction enhancement. If possibilities are found, the information are noted without disturbing the existing data. The results indicate that accuracy of the prediction got decreased due adding data gathered during the observations. Basic Hybridized Framework [3] was proposed to investigate the indices of stock market to make prediction. The framework was based on SVM cum K-nearest neighbor (KNN) algorithm, where it works based on the weights assigned to the features by the algorithms in order to gain more information. KNN was applied to make prediction by calculating the weights of neighbors. The results of basic hybridized framework came up with increased false positives.

Ensemble Dimensionality Reduction Methodology [4] was proposed to predict returns in stock market. It makes simplification and rearrangements in the actual structure of data used for the prediction. The features of dataset were reduced in order to increase the classification accuracy, where ANN was applied for classification. The result came with low classification accuracy due to the ensemble of more data mining techniques. Origin based Association Rule Mining [5] was proposed to predict the financial risks in the stock market. It works based on order independence and calculation on unbiased values. By utilizing the greedy frequency concept, the items used in the transactions are rearranged for the prediction purpose. The low classification accuracy indicates that the algorithm is not suitable for stock market prediction.

Regulatory Disclosure Text Mining [6] was proposed to forecast the fresh information which enter the stock market daily. Short term cum long term news from different languages were analyzed for enhancing the prediction level. Prediction models suitable for multi-dimensional data were studied and applied. The increased false negative information represents that, only news related information are not sufficient for the prediction for stock indices.

Financial Restatement Prediction System [7] was proposed to predict fraudulent (intentional) and erroneous (unintentional) financial restatement. Benchmark prediction models were studied thorough while developing the system. It was found that ANN algorithm was best suitable for the prediction. The low accuracy of prediction shows that the algorithm is suitable low level dataset only. Computational Intelligent Method [8] was suggested to make a review and predict the finance market. Initially, preprocessing and clustering were applied to increase the accuracy. It was found that the preprocessing are possible only for the static data and not for the dynamic data, due to this classification accuracy got down. Enhanced Sentiment Analysis [9] was proposed to predict the stock movement by utilizing the information available in social media. It considers the overall and specific moods for making an analysis of sentiment, where sentiments were extracted in an automated manner. Due to the automatic sentiment extraction, the accuracy becomes low than the human sentiment methodology. Bibliographic Survey Technique [10] was proposed to utilize and emphasize the significant texts that were used for the prediction of finance market. It utilizes the SVM and ANN method for making classification. The very low accuracy result clearly indicates that, only the words were not enough to make prediction on stock or share marketing. It was suggested to use the current and previous stock values for the increase in the accuracy.

III. PROTRACTIBLE FUZZY PARTICLE SWARM OPTIMIZATION

3.1 Fuzzy System

Fuzzy logic (FL) and its interpretation method was proposed by Zadeh in the year 1986. In FL, input data are available for different set. In FL, the models are performed by the basic if-then logic and grammatical (i.e., linguistic) variables. Each fuzzy interpretation method (FIM) has 3 important segments which are fuzzy rules, functions of membership and analyzing component. The three kinds of fuzzy interpretation methods are:

- Mamdani Method – Output are necessarily defuzzified.
- Takagi–Sugeno Method – Provides the substantial number as the final output.
- Tsukamoto Method – Utilizes the tedious Functions for processing the result.

Adoptable Network based FIM [11] was proposed by utilizing the Takagi– Sugeno Fuzzy method. Generally, FIM has 2 simple rules which is followed by Takagi– Sugeno Fuzzy method and mathematically it can be expressed as [12]:

$$\textbf{Rule1}: \; if\ w\ is\ B^{w+1}\ and\ z\ is\ A^{z-1}\ then\ e^{x+1} = o^{x+1} + r^{z+1} \cdot$$
$$\textbf{Rule2}: \; if\ w\ is\ B^{w+2}\ and\ z\ is\ A^{z-2}\ then\ e^{x+2} = o^{x+2} + r^{z+2} + \quad (1)$$

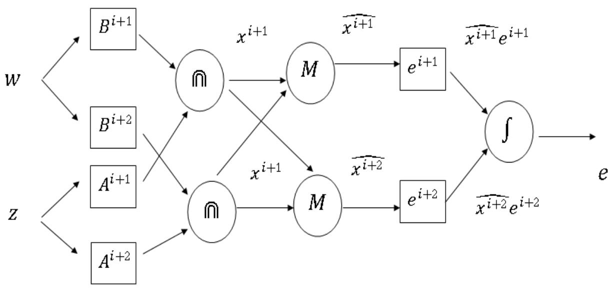

Fig. 1 Architecture of FIM

Fig. 1 demonstrates the ANFIM's common architecture of FIM with the inputs w and z , and the output e. The functional details at every layer are discussed as follows. Layer 1: Inputs to the fuzzy system are provided from this layer, where it contains the adaptable nodes. The nodes have the functioning facility. Every node in this input layer relates to semantic name, and the output values represents the membership function mathematically described as

$$P^{j+1} = \frac{\rho B^{j}}{w} \quad for\ j = 0,1 \qquad (2)$$

P^{l+j} is considered as the node in the j^{th} position in the layer l, for every level. There exist numerous kinds of membership function,

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

but the most common is bell shaped function and it is mathematically expressed as

$$\rho B^{w} = \left(w + \frac{d^j}{b^j}\right)^{2*a^j} \tag{3}$$

where b^j, a^j and d^j are the inference parameters that are well known

Layer 2: All the nodes in the layer are expected to increase the receiving signals with the output, and it can be expressed as

$$P^{j+2} = x^j = \rho B^{j+w} \rho A^{z+j}, \qquad j = 0,1 \tag{4}$$

Layer 3: The j^{th} node in this layer makes calculation to find the ratio to eliminating strength of j^{th} protocol to the total of entire eliminating strength, and mathematically expressed as

$$P^{j+3} = \hat{x}^j = x^j + \frac{x^1 + x^2}{j+1}, \qquad j = 0,1 \tag{5}$$

Layer 4: In this layer, node j has the additional function:

$$P^{j+4} = \hat{x}^j e^h = \hat{x}^j [o^{w+j} + r^{z+j} + q^j] \tag{6}$$

Layer 5: A selected node in the last layer calculates the whole output as the total of overall receiving signal, and it can be expressed as

$$P^{j+5} = \oiint \hat{x}^j e^j = \int_0^j \hat{x}^j e^j \tag{7}$$

3.2 Particle Swarm Optimization (PSO)

In the year 1995, PSO algorithm was first introduced by the authors Kennedy and Eberhard. The primary idea of PSO is searching in a random manner on the available population. The key concept of PSO initiates from the natural behavior of fishes and birds. In this optimization technique, every particle expected to progress in the space of M-dimensional, where it depends on its previous experience which it faces with other particles. In the optimization, every particle has its own position and also a speed (i.e., the velocity) specified by the vectors denoted by $W^j = w^{j+1}, w^{j+2}, \ldots, w^{j+C}$ and $U^j = u^{j+1}, u^{j+2}, \ldots, u^{j+C}$ for the $j-th$ particle. During every iteration, the particles in the swarm are made to compare with neighbors in order to seek the preeminent particle. Every particle tend to save its preeminent position denoted as $o^j = o^{j+1}, o^{j+2}, \ldots, o^{j+C}$. The preeminent location of the entire particles in the swarm is indicated as the universal-preeminent-location and denoted as $H = H^{w+1}, H^{w+2}, \ldots, H^{w+C}$. The speed of every particle is mathematically represented as:

$$U^{j+C} = \begin{pmatrix} X, u^{j+c} - \\ D^{j+1}.random^{j+1}().(opreeminent^{j+c} + y) + \\ D^{j+2}.random^{j+2}().(opreeminent + y) \end{pmatrix} \tag{8}$$

In the above equation $j = 1,2,\ldots,N, c = 1,2,\ldots,C, D^{j+1}$ and d^{j+2} are the possible constants (in a positive value), $random^{j+1}()$ and $random^{j+2}()$ are 0 and 1 values in random manner and X is inactive weight. The fresh location for the particle is mathematically described as

$$w^{j+c} = x^{j+c} - u^{j+c} \tag{9}$$

3.3 Protractible Fuzzy based PSO

In F-PSO each particle is expected to have a common discrete characteristics and tend to move towards the focal point of the possible field. Instead of making the common location and speed assignment, F-PSO assigns a common wave function $\aleph[x,s]$ to all the particles. The characteristics of all the particles in F-PSO is entirely different from traditional PSO. The likelihood of present particle j in position w is determined from the likelihood function towards the population $\aleph[w,s]$. Every particle makes a movement as indicated by the below equations

$$O^{j+c} = \tau * O^{j+c} + \frac{j+1}{\tau} * O^{h+c}, \quad \tau = random() \tag{10}$$

$$W^{j+c} = O^{h+c} - \beta * (npreeminent^c + W^{j+c}) * (j + 1/v) \tag{11}$$

where, $npreeminent$ denotes the best location of the particles.

$$npreeminent = j + NN \int_{j+1}^{N} O^h$$

$$= j + n\int_{j+1}^{N} O^{j+1}, j + N\int_{j+1}^{N} O^{j+2}, \ldots, j + N\int_{j+1}^{N} O^{j+m} \tag{12}$$

O^{j+c} is an arbitrary point among O^{j+c} and O^{h+c} and a nearby attractor the $j-th$ particle on the dimension c. $\aleph$ and v are 0 and 1 values but generated in a random manner and β is the ideal F-PSO parameters.

IV. ABOUT DATASET

Sixteen years of data of whole stock price indexes (CNX Nifty and S&P BSE Sensex) and two stocks (Infosys and Reliance) from Jan 1, 2003 to Dec 31, 2018 is utilized in this research work. All the data is acquired from www.nseindia.com, www.bseindia.com, and www.finance.yahoo.com websites.

These data form the whole data set containing 3983 records in each stocks and prices.

V. PERFORMANCE METRICS

Sensititivty, Specificity, Precision, Recall, Accuracy and F-Measure are the metrics used to measure the performance of proposed algorithm with existing algorithms, where all the values are calculated from the parameters True Positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN). Metric values are projected using the equations Eqs. (13) to Eqs. (18) respectively.

$$Sensitivity = (TP/(TP+FN)) * 100 \quad (13)$$

$$Specificity = (TN/(TN+FP)) * 100 \quad (14)$$

$$Precision = (TP/(TP+FP)) * 100 \quad (15)$$

$$Recall = (TP/(TP+FN)) * 100 \quad (16)$$

$$Accuracy = ((TP+TN)/(TP+TN+FP+FN)) * 100 \quad (17)$$

$$Recall = (2 * ((Precision * Recall)/(Precision + Recall))) * 100 \quad (18)$$

VI. RESULTS AND DISCUSSION

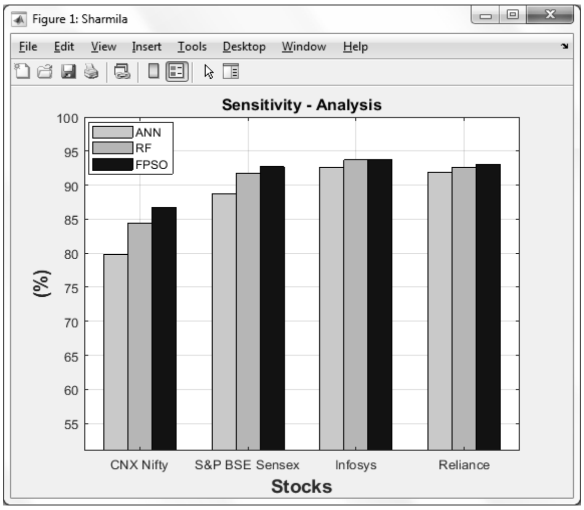

Fig 1. Sensitivity Analysis

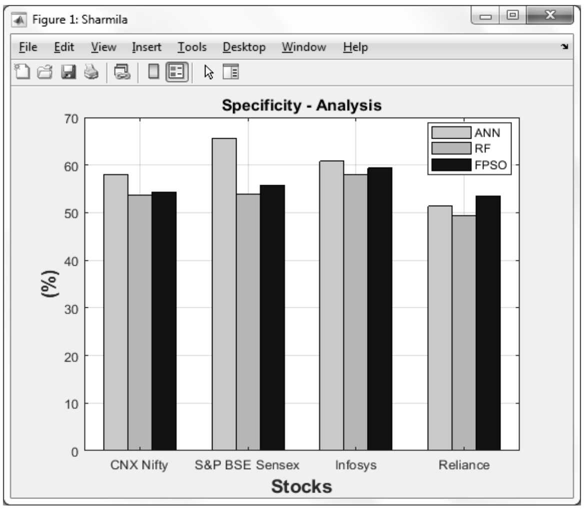

Fig 2. Specificity Analysis

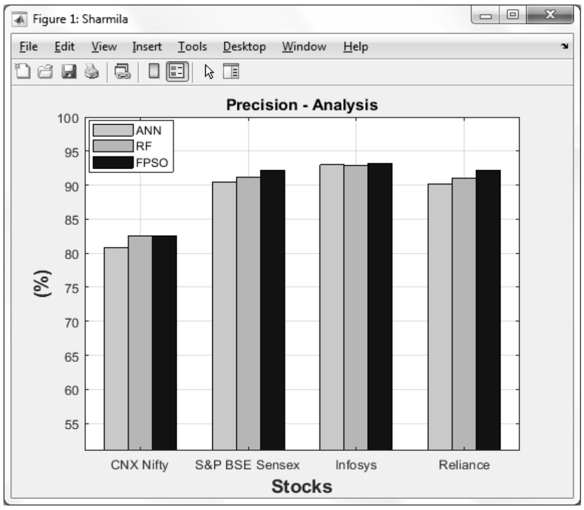

Fig 3. Precision Analysis

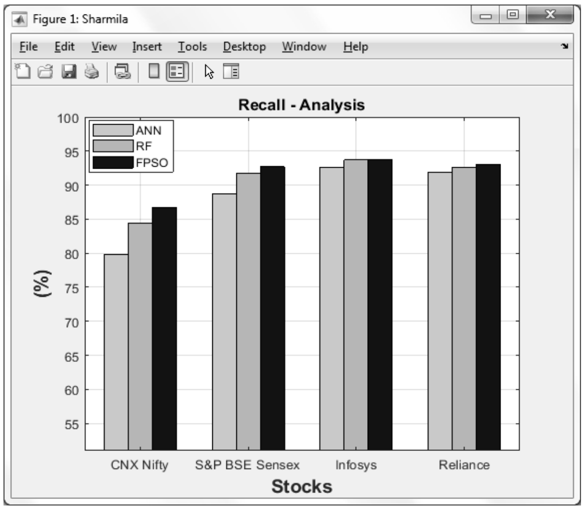

Fig 4. Recall Analysis

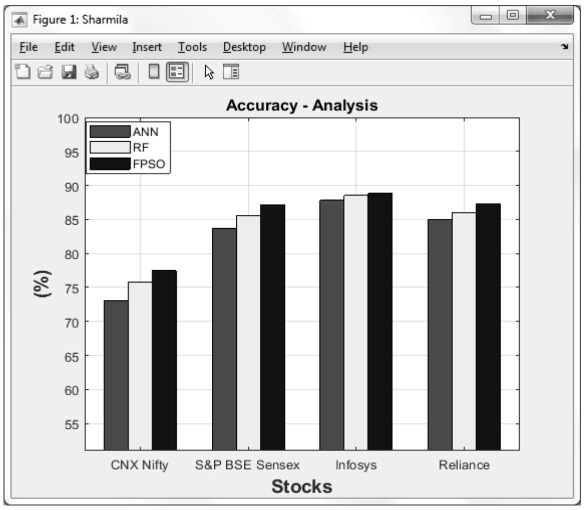

Fig 5. Accuracy Analysis

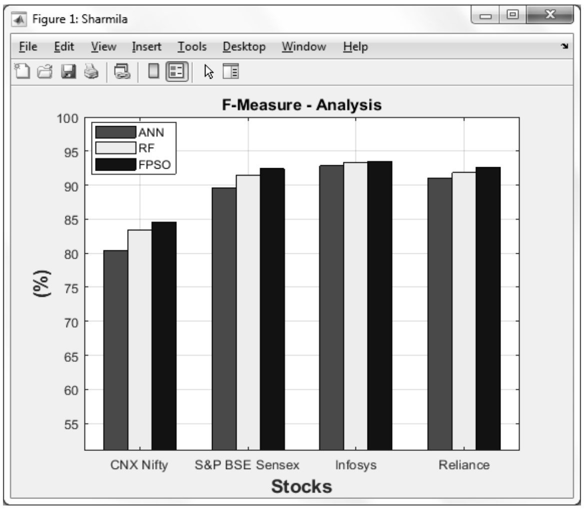

Fig 6. F-Measure Analysis

From the Fig. 1 to Fig. 6, it is evident that the proposed algorithm FPSO outperforms the RF [13] and ANN [14] with all the chosen performance metrics. It is because that FPSO classifies the records based on fuzzy technique by generating the rules in order to make effective classification, but the existing algorithms RF [13] and ANN [14] makes classification in the first in first out manner, which consumes more time also. Overall analysis of the research work founds that the proposed algorithm has its best performance in all the considered datasets.

VII. CONCLUSION

This paper has focused to make prediction towards the movement of stock and price indexes. The performance of two models namely ANN and RF is compared with the proposed algorithm F-PSO based on sixteen years (2003 to 2018) historical data of Indian stock market companies CNX Nifty, S&P BSE Sensex, Infosys, and Reliance. The experiment results indicate that ANN models has low performance of 82.41% of accuracy and F-PSO has high performance of 85.17% of accuracy

REFERENCES

1. Q. Wang, W. Xu, H. Zheng, "Combining the wisdom of crowds and technical analysis for financial market prediction using deep random subspace ensembles," Neurocomputing, Volume 299, Pages 51-61, 2018.
2. M. R. Hassan, K. Ramamohanarao, J. Kamruzzaman, M. Rahman, M. M. Hossain, "A HMM-based adaptive fuzzy inference system for stock market forecasting," Neurocomputing, Volume 104, Pages 10-25, 2013.
3. Y. Chen, Y. Hao, "A feature weighted support vector machine and K-nearest neighbor algorithm for stock market indices prediction," Expert Systems with Applications, Volume 80, Pages 340-355, 2017.
4. X. Zhong, D. Enke, "Forecasting daily stock market return using dimensionality reduction," Expert Systems with Applications, Volume 67, Pages 126-139, 2017.
5. D. Yadav, C. R. Chowdary, "OOIMASP: Origin based association rule mining with order independent mostly associated sequential patterns," Expert Systems with Applications, Volume 93, Pages 62-71, 2018.
6. S. Feuerriegel, J. Gordon, "Long-term stock index forecasting based on text mining of regulatory disclosures," Decision Support Systems, Volume 112, Pages 88-97, 2018.
7. I. Dutta, S. Dutta, B. Raahemi, "Detecting financial restatements using data mining techniques," Expert Systems with Applications, Volume 90, Pages 374-393, 2017.
8. R. C. Cavalcante, R. C. Brasileiro, V. L.F. Souza, J. P. Nobrega, A. L.I. Oliveira, "Computational Intelligence and Financial Markets: A Survey and Future Directions," Expert Systems with Applications, Volume 55, Pages 194-211, 2016.
9. T. H. Nguyen, K. Shirai, J. Velcin, "Sentiment analysis on social media for stock movement prediction," Expert Systems with Applications, Volume 42, Issue 24, Pages 9603-9611, 2015.
10. B. M. Henrique, V. A. Sobreiro, H. Kimura, "Literature review: Machine learning techniques applied to financial market prediction," Expert Systems with Applications, Volume 124, Pages 226-251, 2019.
11. J. R. Jang, ANFIS: Adaptive-Network-based Fuzzy Inference System, IEEE Transactions on Systems, Man and Cybernetics, Volume 23, Pages 665–685, 1993.
12. L. A. Zadeh, Fuzzy Sets, Information and Control, Volume 8, Pages 338–353. 1965.
13. J. Patel, S. Shah, P. Thakkar, K. Kotecha, "Predicting stock and stock price index movement using Trend Deterministic Data Preparation and machine learning techniques", Expert Systems with Applications, Volume 42, Pages 259-268, 2015.
14. Y. Kara, M. A. Boyacioglu, O. K. Baykan, "Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange", Expert Systems with Applications, Volume 38, Pages 5311-5319, 2011.
15. R. J. Vidmar. (1992, August). On the use of atmospheric plasmas as electromagnetic reflectors. IEEE Trans. Plasma Sci. [Online]. 21(3). pp.876—880. Available: http://www.halcyon.com/pub/journals/21ps03-vidmar

AUTHORS PROFILE

B Sharmila M.Sc., M.Phil, (Phd) scholar, has good interest in share market / mutual funds and investments.

Dr. R Khanchana Phd – working as an Asst. Professor, Sri Ramakrishna College of Arts & Science for Women, Coimbatore.