iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 **Issue:** XI **Month of publication:** November 2020

DOI: https://doi.org/10.22214/ijraset.2020.32218

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue XI Nov 2020- Available at www.ijraset.com

Comparative Analysis of Processing Time for Predicting Stock Market Movement by Hybrid Machine Learning Algorithms using Optimization Techniques

B. Sharmila[1], Dr. R. Khanchana[2]
[1]*Research scholar, Dept. of Computer Science, Sri Ramakrishna College of Arts and Science for Women, Coimbatore, India.*
[2]*Assistant Professor, Dept. of Computer Science, Sri Ramakrishna College of Arts and Science for Women, Coimbatore, India.*

***Abstract:** Highlighting Prediction of stock and its price index is always a challenging issue because of presence of uncertainty. Predicting the problem will give the major solution to the success of choosing the investment. Accuracy and precision is the root factor in stock market prediction. It utilizes the collected knowledge for predicting the stock market with increased accuracy. The research focus to provide a solution for stock prediction by optimization, data reduction and reduction in dimension techniques. This helps to determine the time needed to run the computational process which helps for the quick selection of purchasing of the stock. The comparison of running time is evaluated by the various algorithms projected in the exertion. 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 are taken for the experimental work. Performance metrics considered are sensitivity, specificity, precision, recall, accuracy and f-measure.*
***Keywords:** Prediction, Stock Price, Index, optimization, data reduction.*

I. INTRODUCTION

There exist two 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. 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. 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. Previous statistics oriented methods such as auto-regression and linear-regression were supportable for forecasting the finance time series due to their interpretation. The main factor behind the process of prediction was feature extraction and feature reduction. Features were subject to design the assumption that stock histories fully utilized to predict the future. The prediction model utilizes the features of technology-oriented analysis, where it is based on certain assumptions. The success of the prediction model fully depends on the increased correctness of assumptions.

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 has low accuracy, which will be impossible to accept by the investors to make investment. Adaptive Fuzzy Inference System [2] was proposed to enhance 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 is noted without disturbing the existing data.

The results indicate that accuracy of the prediction got decreased due to adding data gathered during the observations. Hybrid Clustering Technique [3] was proposed by combining the selection of portfolios. It is considered as necessary for predicting the price and volatility. Stock market of India was selected as the dataset for evaluating the proposed method. The results with increased false positive rate show the algorithms incompatibility towards different dataset. Investment Behavior based Classification [4] was presented to classify the investors behavior in the market of stock. It utilizes the cellular learning based automation model in deciding towards holding, buying or selling the stocks. It utilized evolution rules to trust the other investors. Macro factors, imitation and reliability were analyzed for the classification, but results shows that the algorithm needs better improvements towards improving the classification accuracy. Neural Network Classifier [5] proposed to classify the stock details to predict the stock prices. Further, it summarized the applications of artificial intelligence in business administration. But the classification accuracy is increased when comparing with traditional classifier. Financial Fraud Detection [6] proposed to detect the fraud transactions and details in stock markets. It used linguistic, vocal and cues of finance for classification. Unavailability of tools made the result to come with increased false positives. Fund Failure Prediction [7] proposed to predict the unexpected failure of funds in stock market. It calculates the chances of time window to get disappear. Surviving predictability was not focused and it results in low accuracy in stock market prediction. Stock repurchase prediction [8] proposed with the focusing on mechanism of earning management. Due to increased finance constraints, it has decreased the prediction accuracy. Time Series Support Vector Machine [9] proposed to analyze the finance time series for matching the patterns. In order to check the subsequence pattern, 2-phase method was used. Brain-inspired fuzzy neural network [10] proposed to learn and predict the trends in temporal series that are highly complex. Hebbian based weight learning concept in fuzzy logic replaced with synoptic metta-plasticy for learning weights. The results ended with decreased accuracy.

III. PROPOSED SYSTEM

A. Protractible Fuzzy Based PSO (PFPSO)

In PFPSO Common Wave Optimization Method (CWOM) is carried out for stock prediction in which each particle meant to the stock data is expected to have common discrete characteristics and tend to move towards the focal point of the possible field. Instead of making the common location and speed assignment, PFPSO assigns a common wave function to all the particles. The characteristics of all the particles in PFPSO are entirely different from traditional PSO. The likelihood of present particle in position *w* is determined from the likelihood function towards the population. Every particle makes a movement to the best location of the particles. The optimized solution is classified with the help of modified fuzzy adaptive fuzzy inference rule based system. the results are defuzzified and analyzed for the prediction of time series stock data in efficient way which throws an idea of purchasing the stock.

B. Adaptive Firefly Algorithm Based Association Rule Mining (AFA-ARM)

The major concept of AFA-ARM is to use the natural behavior of firefly and association rule mining to develop a better classification algorithm to predict the stock price index movements. AFA-ARM is proceeded by Cumulative Particle Filter Method (CPFM) which involves in three stages - Feature Selection Stage, Model Construction Stage, and Classification Stage. Two variables are considered essential in AFA-ARM, which are used to control the behavior of the algorithm, which are intensity of the light, attraction between fireflies. There will be a variation in light intensity based in the brightness. It is indicated and calculated using a fitness function. Attraction rate of individual firefly is calculated zero distance and for mathematical based calculation it is mostly assumed as one. The space that exists between the fireflies determines the attraction with each other. The fireflies tend to change their position continuously because of their flying characteristics. The attraction rate that exists between two fireflies is entirely dependent on the space or distance among them. Euclidean distance is used to calculate the distance exist between any 2 fireflies. After performing the calculation to find distance of two fireflies, brightness of the light is to be checked. The firefly having decreased brightness always moves to the better one with the consideration of three things. The first thing is to identify the position of firefly having decreased brightness. The second thing is to precede the current movement towards firefly having increased brightness by the attraction. The final thing is to have a random walk defined by the random generator. In Feature Selection Stage, Noisy and irrelevant data in dataset may lower the performance of classification to unexpected results. Hence, the feature selection was planned to utilize in AFA-ARM, which considers all the feature of firefly. Classification will be done only after filtrations are done. The final outcome of this feature selection stage ends with finding the relevant features (i.e.), the informative fireflies. The Model construction stage involves in finding the input for training samples and segregates the individual class into a group. Making assignment for the values indicates the firefly positions.

The intensity of the firefly is computed using the fitness function. In the classification stage, the processing of classification is done by distance based classification, intensity based classification and average intensity of class based classification.The above mentioned optimization is inherited as a framework for performing the prediction task for stock market.

C. Capricious Principal Component Analysis Based Feed Forward Neural Networks (CPCA-FFNN)

CPCA is specialized version of linear PCA, linear storing vectors are made to replace the functions related to non-linear functions. It is mathematically shown as

$$S = R(Z)$$
$$\hat{Z} = I(S) \quad (1)$$

Where R indicates the mapping function vectors and $\hat{I}$ represent the de-mapping function vectors, and $\hat{Z}$ denote the data that are reconstructed. The main intention of CPCA is to search R and I functions, which arises to lower the error while reconstructing the data.

In CPCA, two neural networks are trained to study the functions of mapping and de-mapping. But this method is not considered as a feasible option, due to, it needs the values from principal components as input, but it is not possible to calculate without R and I information. In this CPCA-FFNN, the mapping function G is approximated by two transformations and it is expressed as

$$y_l = d(zv_1 + c_1)$$
$$s = y_l v_2 + c_2 \quad (2)$$

De-mapping functions are adjusted with the use of last two transformations, which use Eq. (3)

$$y_e = d(sv_3 + c_3)$$
$$\hat{z} = y_e v_4 + c_4 \quad (3)$$

Where, Z denotes the row vector of matrix Z, y_e denotes the i^{th} dimension mapping layer vector, S represent the h^{th} dimension in principal component vector, $\hat{z}$ represent the data that are reconstructed, C and V indicates the vectors of bias and weight matrices.
While training the network, the parameter are made to adjust to decrease the error that arise during the reconstruction and it is expressed as

$$F = \frac{\sum_{u=1}^{k}\sum_{w=1}^{l}(z_{uw}-\hat{z}_{uw})}{kl} \quad (4)$$

IV. RESULTS AND DISCUSSIONS

Data of stock price indexes (CNX Nifty and S&P BSE Sensex) and two stocks (Infosys and Reliance) are taken for our experimental research. All the data are acquired from NSE & BSE INDIA web-portals. These data form the whole data set containing 3983 records in each stocks and prices.

A. Processing Time Analysis

The comparison of processing time is carried out in all the three above mentioned methods and the results are analyzed to evaluate the proposed work efficiently. The research is executed using MATLAB tool. Table 1 represents the average processing time (PT) of each module in the research. Sensitivity, Specificity, Precision, Recall, Accuracy and F-Measure are the metrics used to measure the performance of proposed algorithms.

Method	Processing Time (Seconds)
RF	121.6
PFPSO	109.2
AFA-ARM	91.8
CPCA-FFNN	62.4

Table 1 Processing time

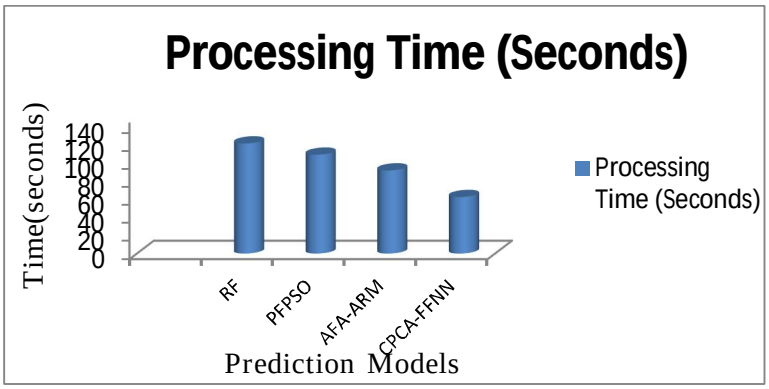


Figure 1 Comparison of processing time of Algorithms

PARAMETERS	RF	PFPSO	AFA-ARM	CPCA-FFNN
SENSITIVITY	87.94	89.95	91.98	96.59
SPECIFICITY	83.72	87.12	90.34	95.60
PRECISION	87.81	89.82	91.83	94.85
RECALL	86.86	88.99	91.76	96.63
ACCURACY	84.09	87.26	90.12	94.81
F-MEASURE	86.95	88.96	90.98	95.11

Table 2 Analysis of performance evaluation of all models on stock data.

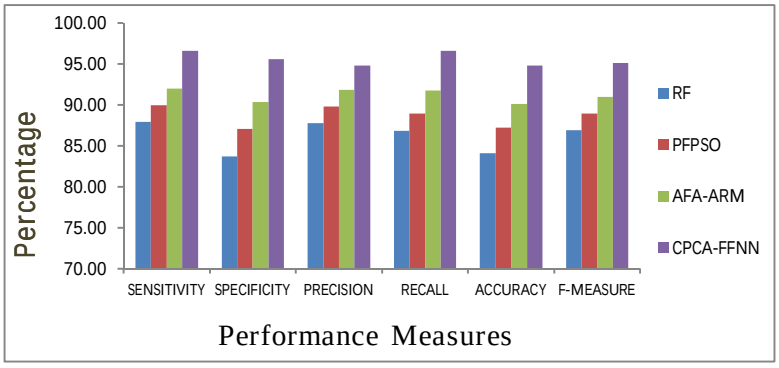
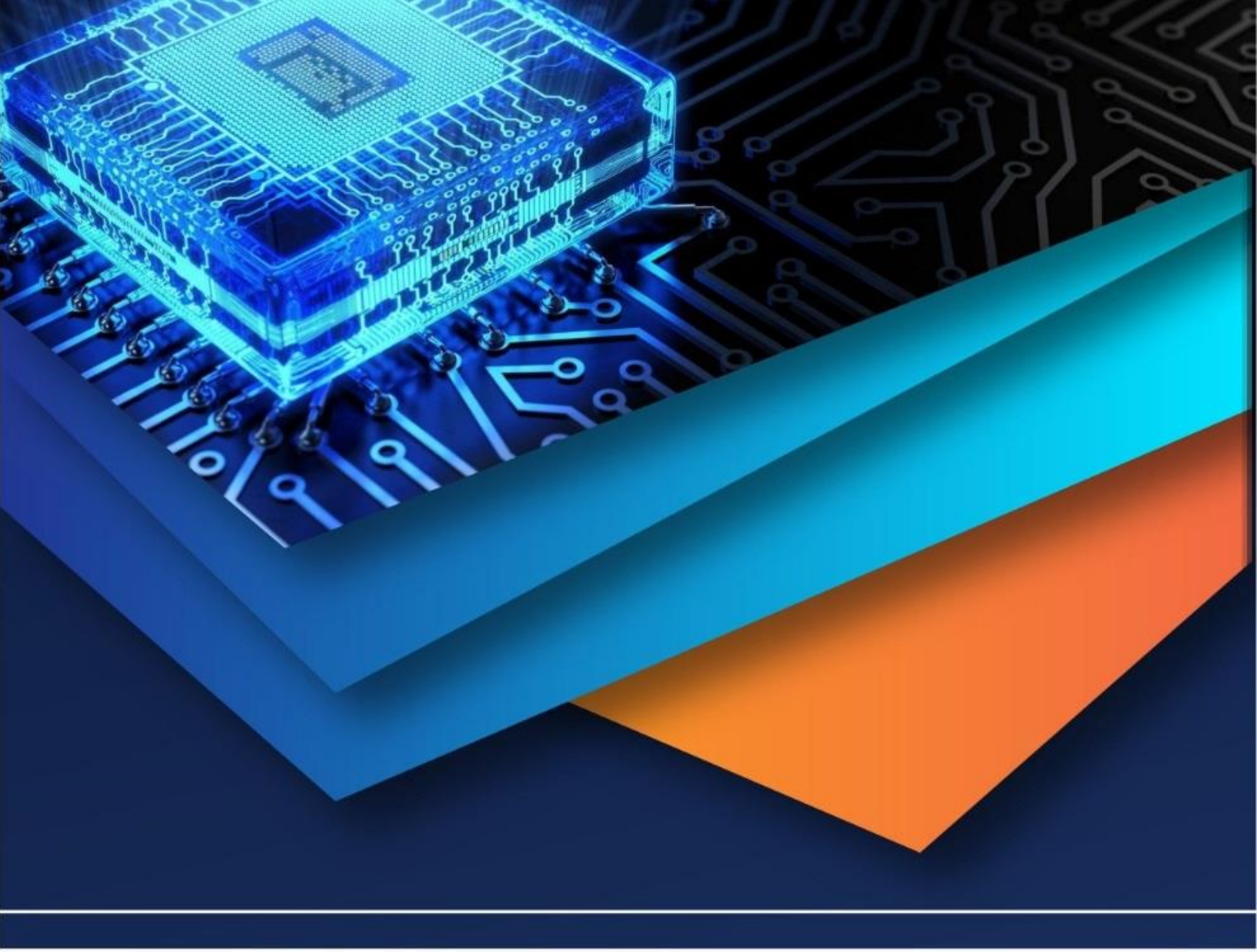


Figure 2 Comparisons of Results of Algorithms for Prediction of Stock Data

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] 3 S. Choudhury, S. Ghosh, A. Bhattacharya, K. J. Fernandes, M. K. Tiwari, "A real time clustering and SVM based price-volatility prediction for optimal trading strategy", Neurocomputing, Volume 131, Pages 419-426, 2014.
[4] M. Mozafari, R. Alizadeh, "A cellular learning automata model of investment behavior in the stock market," Neurocomputing, Volume 122, Pages 470-479, 2013
[5] Y. Li, W. Jiang, L. Yang, T. Wu, "On neural networks and learning systems for business computing," Neurocomputing, Volume 275, Pages 1150-1159, 2018.
[6] C. S. Throckmorton, W. J. Mayew, M. Venkatachalam, L. M. Collins, "Financial fraud detection using vocal, linguistic and financial cues," Decision Support Systems, Volume 74, Pages 78-87, 2015.
[7] P. Cogneau, G. Hübner, "The prediction of fund failure through performance diagnostics," Journal of Banking & Finance, Volume 50, Pages 224-241, 2015.
[8] K. Farrell, E. Unlu, J. Yu, "Stock repurchases as an earnings management mechanism: The impact of financing constraints," Journal of Corporate Finance, Volume 25, Pages 1-15, 2014.
[9] X. Gong, Y. Si, S. Fong, R. P. Biuk-Aghai, "Financial time series pattern matching with extended UCR Suite and Support Vector Machine," Expert Systems with Applications, Volume 55, Pages 284-296, 2016.
[10] A. R. Iyer, D. K. Prasad, C. H. Quek, "PIE-RSPOP: A brain-inspired pseudo-incremental ensemble rough set pseudo-outer product fuzzy neural network", Expert Systems with Applications, Volume 95, Pages 172-189, 2018.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY