Correlation Based ADALINE Neural Network for Commodity Trading

Chandra J¹, Dr Nachamai.M², Dr Anitha S Pillai³

¹ Christ University, Bangalore, Karnataka, India.
Email: Chandra.j@christuniversity.in
² Christ University, Bangalore, Karnataka, India.
Email: nachamai.m@christuniversity.in ;
³ Hindustan University, Chennai, India
Email: anithasp@hindustanuniv.ac.in

Abstract – Commodity trading is one of the most popular resources owing to its eminent predictable return on investment to earn money through trading. The trading includes all kinds of commodities like agricultural goods such as wheat, coffee, cocoa etc., and hard products like gold, rubber, crude oils etc.. The investment decision can be made very easily with the help of the proposed model. The proposed model correlation based multi layer perceptron feed forward adaline neural network is an integrated method to forecast the future values of all commodity trading. The correlation based adaline neuron is used as an optimized predictor in the multi layer perceptron feed forward neural network. The correlation is used for feature selection before building the predictive model. The main aim of the paper is to build the predictive model for commodity trading. The model is created using correlation based feature selection and adaline neural network to prognosticate all future values of commodities. The adaptive linear neuron is formed with the help of linear regression. To implement the proposed model the live data is captured from mcxindia. The mcxindia is considered as one the popular website for doing commodities and derivatives in India. To train the proposed model, few random samples are used and the model is evaluated with the help of few test samples from the same data set.

Keywords – Return on Investments (ROI); Artificial Neural Network (ANN); Multilayer Perceptron (MLP); Adaptive Linear Neuron (ADALINE); Correlation Based Feature Selection (CBFS); Mean of Magnitude Relative Error (MMRE).

1. Introduction

The artificial neural network was introduced by mcculloch and Pitts and it is considered as one of the logical simulation of a biological brain. The ANN has the ability to detect future values of any unseen data object in the dataset. ANN consists of an interdependent group of artificial neurons, and it processes information using a connection based approach for computation. The background knowledge is collected from its environment and is stored as interconnected neurons in the neural network. An artificial neural network is an essential method for financial services due to their ability to handle insufficient data which can process quickly. The predictive model is an important method to make variety of predictive decisions in the field of commodity trading. To determine the likelihood of the investment which is based on the current risk, the investor can decide whether to sell or buy the product. Generally, all marketers look at, how consumers have behaved to the overall economy when planning on a new campaign, and changes in demographics to determine if the current mix of the products will tempt the customers to make a purchase. Especially the trader who has a liking for short term investment wants to make more return’s on the investment will expect high optimality in the future prediction. The idea is to increase the ROI values with specific periods. So the proposed neural network model has the ability to learn by observation which is considered as one of the most powerful, feasible and very efficient method for doing predictive analytics. The predictive analytics is the process of extracting information from data to discover recurrent patterns.

For any investor, the main aim of the investment made would be maximizing the values of ROI and also to reduce the investment risks. The trader will be keen on taking a decision quickly like buy or sell is based on the commodity value. Performing future prediction with respect to commodity exchange is of varying nature.

2. Related Work

Purwanto et al., [1] has suggested three prediction models for an infant mortality rate. The prediction was done using Auto Regressive Integrated Moving Average (ARIMA), neural network and Linear Regression models. The performances of the prediction models are compared using Indonesia infant mortality rate data collected during 1995 – 2008. From the comparative study, it is found that NN model performs better than other methods and is evaluated with the help of mean absolute error, mean absolute percentage error and root mean square error. The MLP NN is used to predict the heart attack and also to diagnose cancer. The MLP NN is considered as a
popular or well known method for predictive analytics [2] [3]. Rohit R et al., [4] proposed the rain fall prediction using MLP NN. Arsad P M et al., [22] describes NN and linear regression prediction models to predict students’ performance based on multiple entry levels namely with respect to matriculation and diploma holders. This experiment was based on electrical engineering degree students at the department of electrical engineering UiTM. From the above findings, it was found that both prediction models indicated similar results as far as mean square error is concerned. The MLP NN is more optimized method than other methods like self organized feature map and recurrent neural network and the performance measure is evaluated with the help of mean MMRE. Mahdi Pakdaman Naeini et al., [5] has proposed the stock value prediction using MLP NN and the model is checked with Elmen network. It is found that the predicted values using MLP NN are very closer to the original values with higher accuracy compared to Elmen network.

3. Proposed Work

3.1. Predictive Model

The proposed predictive model is an integrated method for making effective and efficient prediction for all kinds of commodity trading. According to the investor, the integrated method works efficiently than the single method for forecasting the future values of any products in the investment. The proposed model uses Pearson’s Correlation Coefficient (PCC) for feature selection and the MLP adaline NN for predicting unseen data or future data. This integrated approach works well with respect to all kinds of investment in commodity trading. Especially the current model CBFS MLP adaline feed forward NN predicts the future values of all commodity indexes. The proposed predictive model performs feature selection as the first step, and then checks the efficiency of the feature selection with P-value summary. The selected variables are used in building neural network which user adaline neuron as a predictor. Due to the efficiency of the adaptive linear neuron, the model outperforms in prediction.

3.2 Data Set Description

The mcxindia is one of the popular webpage for the commodity trading in India. It is an ISO standard company for commodity exchange. It is a largest leading commodity futures exchange with a market share. The mcxindia data set contains more than thirteen features. The attribute of the data set are item name, ExpiryMonth(dd-mm-yy), Open(Rs), High(Rs), Low(Rs), LTP(RS), PCP(RS), Change(%), Buyqty, Sellqty, Sellprice(RS), etc,... It also describes many details like mcxagri, mcxenergy and mcxmetal. The historical data is used for investigation. The data set has multiple attributes but only four attributes are considered for building the adaline NN after feature selection. The four attributes are selected using correlation values.

3.3 Correlation Based Feature selection (CBFS)

In [8] Mark A Hall has proposed correlation based feature selection for machine learning in his PhD work. The thesis addresses the different problems associated with feature selection. The feature selection is compared with three machine learning algorithm such as C4.5, IB1 and Naïve Bayes. The correlation based feature selection is experimented with simulated data set as well as with UCI data set. The correlation based feature selection works quickly for identifying highly relevant features with better accuracy. In [20], PCC is used to test research hypotheses when there is a linear and bi-variate relationship between any two variables. In this paper, it is found that PCC based predictive model is used to remove highly irrelevant attributes from the data set. It is also suggested that the PCC removes all irrelevant attributes to do any task in computational intelligence. It gives highly relevant attributes than any other methods. The predictive model makes more intellect to the investor when all the irrelevant attributes are removed for doing data analysis. The attribute subset selection is a most challenging process faced in the field of data analysis. The real world data has multiple attributes in order to reduce irrelevant attribute, the attribute subset selection method needs to be used for reducing the irrelevant data with preserved content into the whole dataset. As in data mining, all prediction models need an efficient method for doing relevant analysis. So that the good predictor methods like neural network, support vector machine may work efficiently with relevant subset. The importance of the CBFS is to improve high relevancy in the data analysis. The CBFS method can also be used for supervised and unsupervised learning for the selection of highly relevant features for further processing in data analysis. There are many types of correlation coefficient methods available for feature selection, only PCC was used and the highly correlated values are considered for further data analysis. In [21], Jacek Biesiada et al., applied PCC method for finding non-redundant binned feature subset selection and compared with other
existing methods and it found that the PCC based feature selection works well with SVM classifier.

The current research work has implemented PCC as the feature subset selection method for selecting the relevant attributes. The PCC is defined as the variance of two variables divided by the product of their standard deviation. The PCC is represented by $P_{xy}$, where the formula is as follows:

$$P_{xy} = \frac{\text{cov}(xy)}{\sigma_x \sigma_y} = \sum \frac{(x_i - \mu_x)(y_i - \mu_y)}{\sigma_x \sigma_y} \quad (1)$$

Where cov(xy) is the covariance of x and y, $\mu_x$ is the arithmetic mean of x, $\mu_y$ is the arithmetic mean of y. Here the $\sigma_x \sigma_y$ is the standard deviation of x and y.

CBFS is mainly used as an efficient and effective feature selection method in computational intelligence.

| BuyPrice.Rs. | BuyQty | Change... | High.Rs. | Low.Rs. | Open.Rs. | LTP.Rs. | PCP.Rs. | Disc.Rs. | Sell.Rs. | SellQty |
|--------------|--------|-----------|---------|---------|---------|--------|---------|---------|---------|--------|
| 1.0          | 1.0    | 0.0       | 0.0     | 0.0     | 0.0     | 0.0    | 0.0     | 0.0     | 1.0     | 1.0    |
| BuyQty       | 1.0    | 1.0       | 1.0     | 1.0     | 1.0     | 1.0    | 1.0     | 1.0     | 1.0     | 1.0    |
| Change...    | 1.0    | 1.0       | 1.0     | 1.0     | 1.0     | 1.0    | 1.0     | 1.0     | 1.0     | 1.0    |
| High.Rs.     | 0.0    | 1.0       | 1.0     | 0.0     | 0.0     | 0.0    | 0.0     | 0.0     | 1.0     | 1.0    |
| Low.Rs.      | 0.0    | 1.0       | 1.0     | 0.0     | 0.0     | 0.0    | 0.0     | 0.0     | 1.0     | 1.0    |
| LTP.Rs.      | 0.0    | 1.0       | 1.0     | 0.0     | 0.0     | 0.0    | 0.0     | 0.0     | 1.0     | 1.0    |
| Open.Rs.     | 0.0    | 1.0       | 1.0     | 0.0     | 0.0     | 0.0    | 0.0     | 0.0     | 1.0     | 1.0    |
| PCP.Rs.      | 0.0    | 1.0       | 1.0     | 0.0     | 0.0     | 0.0    | 0.0     | 0.0     | 1.0     | 1.0    |
| Sell Price.Rs. | 0.0  | 1.0       | 1.0     | 0.0     | 0.0     | 0.0    | 0.0     | 0.0     | 1.0     | 1.0    |
| SellQty      | 1.0    | 1.0       | 1.0     | 1.0     | 1.0     | 1.0    | 1.0     | 1.0     | 1.0     | 1.0    |

The table 1 shows the pair wise correlation coefficient values using statistical summary in R tool. To select the relevant attributes, the highly correlated variables are considered for doing further analysis. To determine the variable as highly correlated, the correlated values which are near to one or equal to one is considered as the highly associated variables or related variables. As an average, any correlated value which is very closer to one is considered as highly correlated variables. From the fig (ii) the attributes ‘open’, ‘low’, ‘high-values’ and ‘selling-price’ are considered as highly correlated variables. The correlated variables are verified with the help of statistical parameter P-value as a performance measure, so the P-value calculation is considered as an evaluation parameter for verifying the correlated variables.

3.4. Performance Measures on CBFS

3.4.1 P-Value Selection: The P-values are considered as evaluation parameters to verify all associated variables. In data analysis, the P-value is considered as the possibility of receiving the same value for model building. If the P-value is less than or equal to the doorstep value before setting the data set, one discards the null hypothesis and accepts the trial hypothesis as applicable [16]. For easy understanding with an impact level of ten percentages, the possibility of having the significant P-values belongs to any one of the following criteria. The criteria for selecting the feature subset selection based on P value is as follows.

1) If P-value lies between the values 0.01 to 0.05 then it is assumed to be very strong presumption.
2) If P-value belongs to 0.05 to 0.1 then it is assumed to be low presumption.
3) If P-value is greater than 0.1 then it is considered as no presumption. Finally, the threshold values are too optimistic when P-values reaches to less than 0.001 or 0.0053. From the above criteria it is found that, when we try to verify the relevancy of the variable which is considered to be very strong, when the P-value reaches between 0.001 to 0.0053.

The P-values on PCC is implemented with the help of statistical summary in R tool. From table 2, the P values are shown, based on the criteria for selecting the variables, whichever variables scored between 0.01 to 0.005 are confirmed as highly correlated, which are fed as inputs for building adaptive linear regression NN.

4. Neural Network Implementation

Neural network is inspired by the computational structure found in the human brain and the neural network can imitate intelligence activities of the human brain like thinking, sensing, etc. It comprises of many processing elements called as units which are analogous to inter-neurons simply perform as intermediary processors [12]. The MLP feed forward neural network is used for multi class classification. The figure 2 shows the architecture of MLP feed forward network which uses the adaptive linear regression neurons.

| Buy.Price | BuyQty | Change... | High.Rs. | Low.Rs. | Open.Rs. | LTP.Rs. | PCP.Rs. | Disc.Rs. | Sell.Rs. | SellQty |
|-----------|--------|-----------|---------|---------|---------|--------|---------|---------|---------|--------|
| <0.001    | <0.001 | <0.001    | <0.001  | <0.001  | <0.001  | <0.001 | <0.001  | <0.001  | <0.001  | <0.001 |
| <0.001    | <0.001 | <0.001    | <0.001  | <0.001  | <0.001  | <0.001 | <0.001  | <0.001  | <0.001  | <0.001 |
| <0.001    | <0.001 | <0.001    | <0.001  | <0.001  | <0.001  | <0.001 | <0.001  | <0.001  | <0.001  | <0.001 |
| <0.001    | <0.001 | <0.001    | <0.001  | <0.001  | <0.001  | <0.001 | <0.001  | <0.001  | <0.001  | <0.001 |
| <0.001    | <0.001 | <0.001    | <0.001  | <0.001  | <0.001  | <0.001 | <0.001  | <0.001  | <0.001  | <0.001 |
| <0.001    | <0.001 | <0.001    | <0.001  | <0.001  | <0.001  | <0.001 | <0.001  | <0.001  | <0.001  | <0.001 |
| <0.001    | <0.001 | <0.001    | <0.001  | <0.001  | <0.001  | <0.001 | <0.001  | <0.001  | <0.001  | <0.001 |
| <0.001    | <0.001 | <0.001    | <0.001  | <0.001  | <0.001  | <0.001 | <0.001  | <0.001  | <0.001  | <0.001 |
| <0.001    | <0.001 | <0.001    | <0.001  | <0.001  | <0.001  | <0.001 | <0.001  | <0.001  | <0.001  | <0.001 |
| <0.001    | <0.001 | <0.001    | <0.001  | <0.001  | <0.001  | <0.001 | <0.001  | <0.001  | <0.001  | <0.001 |

Table 2. P-Values on Pearson’s correlation coefficient using R tool.
The process of training the MLP feed forward neural network classifier is a simple process in which the sum of weights is optimized on the basis of developing best prediction of the target variable. In order to obtain the best patterns, more computing power is needed, so to get more computing power, the current paradigm is processed quickly with in short period of time. The data set with no errors are considered for selecting both environmental and target variables. If the predicted value differs from an actual value then it considered as erroneous. The feed forward neural network is an artificial neural network where the connections between different units form a uni-directed cycle. In this neural network, the information moves only one direction forward from the input nodes, through the hidden nodes and to the output nodes. There are no cycles or loops in the network. The feed forward neural network uses the single layer as the hidden layer. The present model uses linear regression formula as the predictor in the NN.

4.1 Linear Regression

The linear regression formula is used to frame the adaptive neural network for making the prediction. Linear regression is used to predict the value of a variable based on other variables. The regression analysis is used to measure the strength of the relationship between two different sets X and Y. The data is modeled using linear predictor functions and the unseen variables are estimated from current variables with the linear regression function.

4.2 ADALINE Neural Network

The ADALINE neural network has many inputs and one output. The adaline neural network can be used to classify objects into two different categories. The adaptive linear element minimizes the sum squared error over the training data set. However, the adaline neural network can classify objects in three manners when the objects are linearly separable. The link between the input and the output neurons holds the weighted interconnections, which can change during the training phase of adaptive linear neuron.

4.3 Proposed algorithm

The proposed algorithm describes the different steps involved in building predictive model for making the numerical prediction is shown. Using the proposed algorithm, the future trade value for all commodity tradings are predicted and becomes useful for investment.

Predictive Algorithm

1. Load the optimized data set D.
2. Perform features selection using Pearson’s correlation coefficient.
3. Build neural network using adaptive linear neurons.
4. Compute \( h_1, h_2, ..., h_n \) then calculate \( \log(h_1), \log(h_2), ..., \log(h_n) \).
5. Perform predicted-value = \( i_n + d_1h_1 + d_2h_2 + d_nh_n \).
6. Evaluate the predicted-value with the help of MRE, MMRE and then compute
   \[ \text{Accuracy Rate} = 1 - \text{MMRE}. \]
7. Check If the Accuracy-rate is >99 then the predicted-value is used for future prediction
   Otherwise repeat step-1 to step-6 until the accuracy-rate is >99.

4.4 Performance Measures on NN.

The efficiency of NN method is checked with the help of MMRE and accuracy for the given data set. Using the training set to measure accuracy will typically provide an optimistically biased estimate, especially if the learning algorithm over fits the training data. Sample accuracy is an estimate of the true accuracy of the algorithm, that is, the probability of the algorithm will work correctly on classified instances drawn from the unknown distribution. Generally, the error analysis is done with the help of MRE and MMRE in data analytics. In this research
work, classification accuracy is the primary evaluation criteria for implementing the proposed predictive model. The classification accuracy is defined as the percentage of test data which is correctly classified by the algorithm

\[
MRE = \frac{\text{Estimated Value} - \text{Actual Value}}{\text{Actual Value}}
\]

(4)

\[
\text{MMRE} = \frac{1}{n} \times \text{Sum of MRE} \quad \text{Accuracy} = 100 - \text{MMRE}
\]

(5)

The MRE is an evaluation measure to check how large as the computation value and also to standardize the measurement. The relative error is performed as the absolute error divided by the magnitude of the exact value. The absolute error is the magnitude of different variance between the actual value and the estimated value. To calculate the value of MMRE measures, the model uses the difference between actual and estimated value. The average differences are considered as numerical value of every observation in the data distribution, and these values are considered as very sensitive to individual predictions. The predictive model is developed from a training sample where structures are randomly collected from the data set. Trade forecasting is the process of determining future values of trading or other financial trade on a commodity exchange. The successful prediction of all commodity future values could be important earnings about events whose actual outcomes have not yet been observed. The trader’s might estimate some variables of interest at some specified future date. Computable forecasting models are used to forecast future data with the help of historical data. These methods are usually implemented to reduce intermediate range decisions. The adaline NN forecasts are the most cost-effective methods to provide a benchmark against all other methods. Since the capabilities of such intelligence method would be difficult for an individual human mind to apprehend, the methodological singularity is seen as an occurrence beyond which events cannot be predicted.

5. Experimental Result

The figure 4 and table 3 describes the experimental result of predictor variable with help of all hidden variables. Once the high-values are estimated, the predictive high values are calculated. The predictive parameters for checking the accuracy of the high-values are performed with the help of the variables Error, MRE, MMRE and Accuracy. The MLP feed forward adaline NN is implemented using NNET package in R tool.
The estimated high-values are checked against the original values of the next day, and then the difference between the estimated high-value and the actual high value is found as 0.566.

![Figure 4. MLP feed forward Adaptive NN](image)

The same experiment is carried out randomly on another day in the same week, the difference is 0.766. So the approximate average error rate is 0.6 for all commodity trading. Based on CBFS, only 3 variables are considered as input variables to build MLP NN. From this experimental result, it is found that the estimated accuracy is measured as 99.5 with the error rate of 0.5. The same model can be used for making all kinds of items in commodity trading. To check the accurate value of the model, the MMR, MMRE is calculated. The total value of MRE is 0.084939 and MMRE is 0.566261. The accuracy is measured as 99.43374. Through the experimental result, it is found that the correlation based adaline MLP neural network is used as the predictor model for predicting the next high values of all commodity indexes. The short term investor who wants to buy or sell the product will benefit with this model based on the estimated high values. From this experimental work, it is also decided that the proposed predictive model can be used for making any kind of future values for different types of numeric prediction.

### 6. Conclusion

The high value of the next day commodity trading is estimated with the help of the current integrated predictive model. As an evaluation measure, the accuracy rate is calculated with the help of statistical parameter MRE, MMRE and the accuracy rate. From the current research, it is observed that the predictive high value is compared with the real-time data for the next day as final verification. The estimated error rate is very closer to zero value and the model is verified with all commodity indexes. So Correlation based adaline MLP NN can be applied for all kinds of commodity trading. The correlation based adaline MLP NN is capable of predicting exactly the same value as the next day high values for all commodities. As a further scope, the proposed predictive model can also be applied to other long term risk based investment.

### References

[1] Purwanto, Chikkannan Eswaran, Rajasvaran Logeswaran, A Comparison of ARIMA, Neural Network and Linear Regression Models for the Prediction of Infant Mortality Rate, Fourth Asia International Conference on Mathematical/Analytical Modeling and Computer Simulation, 2010.

[2] Yan.H.M, Jiang.Y-T, Zheng.Z, Peng, Q-H Li, A multilayer perceptron based medical decision support system for heart disease diagnosis, Expert Syst Application, pp: 727-81, 2006.

[3] Artificial Neural Network, International Journal of Emerging Technology and Advanced Engineering (IJETAE), Vol.No.2, issue-1, Jan-2012.

[4] Necdet SUT, Yahya CELIK, Prediction of mortality in stroke patients using multilayer perceptron neural networks, Turk J Medical Sciences, Volume No-42 (5), pp: 886-893, 2012.

[5] Mahdi Pakdaman Naeini, Hamidreza Taremian, Homa Baradaran Hashemi, Stock Market Value Prediction Using Neural Networks, International Conference on Computer Information Systems and Industrial Management Applications (CISIM), 978-1-4244-7818-7/10.2010@IEEE.

[6] Sut.N, Şenocak.M. Assessment of the Performances of Multilayer Perceptron Neural Networks in Comparison with Recurrent Neural Networks and two statistical methods for diagnosing...
coronary artery disease. Expert System, Volume no: 24, pp: 131-42, 2007.
[7] Savita, Divneet Singh Kapoor and Amit Kumar Kohli “Comparative Evaluation of Generalized ADALINE using Variable Learning Rate Parameter”, IJCEM International Journal of Computational Engineering & Management, Vol. 16 Issue 4, July 2013, ISSN (Online): 2230-7893.
[8] Mark A. Hall, PhD Thesis, submitted to Waikato University, New Zealand, Correlation-based Feature Selection for Machine Learning, April 1999.
[9] P.Langley and S.sage .Induction of selective Bayesian classifier. Proceedings of 10th conference on uncertainty in artificial intelligence,  Seattle, Margan Kaufmann, W.A, 1994.
[10] Yardimci A. Soft computing in medicine, Applied Soft Computing. Volume No: 9, pp: 1029 – 1043, 2009.
[11] D.Hammerstram, “Working with neural network”, IEEE spectrum, 1993, pp: 46-53.
[12] Hecht-Nielsen.R. Neuro Computing. Addison Wesley, 1990.
[13] Hertz. J., Krogh. A and Palmer.R. Introduction to the Theory of Neural Networks, Addison Wesley, 1991.
[14] Hopfield. J, Neural networks and physical systems with emergent collective computational abilities. Proceeding National Academics Sciences (USA), Volume No: 79, pp: 2554-2558, 1982.
[15] Rubinstein ,R.Y and Kroese. D.P, The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte Carlo Simulation and Machine Learning, Springer-Verlag,NewYork, 2004.
[16] www.mcxindia.com
[17] Nuzzo, R. (2014). "Scientific method: Statistical errors". Nature 506 (7487): 150. DOI: 10.1038/506150.
[18] Thawornwong, S, Enke, D. Forecasting Stock Returns with Artificial Neural Network, G.P. IRM Press, 2004.
[19] Ravindran Ramasamy α & Tan Chee Siang “Convergence of Actual and Predicted Share Prices – An ADALINE Neural Network Approach”
[20] M.A.Hall. Correlation-based Feature Subset Selection for Machine Learning PhD thesis, Department of Computer Science, University of Waikato, Waikato, N.Z, 1999.
[21] L.Yu and H.Liu. Feature selection for high-dimensional data: A fast correlation-based filter solution. 12th International Conference On Machine Learning (ICML-03), Washington, D.C., pp.856–863, Morgan Kaufmann, CA 2003.

Vitae

Chandra.J is an Associate Professor in Computer Science Department, Christ University. Her research interests include Artificial Neural Network, Data Mining, Genetic algorithm, big data analytics and Predictive analytics.

Dr Nachamai M, is an Associate Professor in Computer Science Department, Christ University. Her research interests include Perception, Computer Vision, Image Processing, Digital Signal Processing and Artificial Intelligence.

Dr Anitha S Pillai is working as a Professor and Head in department of MCA, Hindustan University, Chennai, Tamil Nadu, India. Her research interests include Software Engineering, Data Mining, Natural Language Processing, Soft Computing and Artificial Intelligence.