The Pharma Innovation Journal 2021; 10(10): 1220-1222

www.ThePharmaJournal.com

The Pharma Innovation

ISSN (E): 2277- 7695
ISSN (P): 2349-8242
NAAS Rating: 5.23
TPI 2021; 10(10): 1220-1222

www.thepharmajournal.com
Received: 28-08-2021
Accepted: 30-09-2021

Mohanapriya M
Department of Physical Sciences & Information Technology, AEC&RI, TNAU, Coimbatore, Tamil Nadu, India

R Vasanthi
Department of Physical Sciences & Information Technology, AEC&RI, TNAU, Coimbatore, Tamil Nadu, India

Patil Santhosh Ganapati
Department of Physical Sciences and Information Technology, AEC&RI, Tamilnadu Agricultural University, Coimbatore, Tamil Nadu, India

D Muruganandhi
Department of Agriculture and Rural Management, Centre for Agriculture and Rural Development Studies, AC&RI, TamilNadu Agricultural University, Coimbatore, Tamil Nadu, India

Corresponding Author:
Mohanapriya M
Department of Physical Sciences & Information Technology, AEC&RI, TNAU, Coimbatore, Tamil Nadu, India

Estimating forecasting performance of auto regressive integrated moving average model and artificial neural network for futures trading volume of maize

Mohanapriya M, R Vasanthi, Patil Santosh Ganapati and D Muruganandhi

Abstract
This study involves comparing the predictive performance of the ARIMA model and ANN model. Monthly data of open price, close price, mean cash price, open interests and trading volume of maize futures are used for analysis. Trading volume is taken as dependent variable and all other variables are taken as independent variables. Data are checked for stationarity using Augmented Dickey-Fuller test based on the p-value. Trading volume and open interest are stationary at level and all other variables are stationary at first difference. The optimal ARIMA model was selected based on ACF plot, PACF plot and lower AIC, BIC values. The optimal model was found to be ARIMA (3, 1, 0). For ANN data set are divided into train and test set in the ratio of 7:3. Weights and bias are obtained for the built model. The forecast is obtained based on the coefficients of the model. The forecasted values are compared using RMSE, MAPE and MAE. Based on the results, ANN is better in forecasting the futures trading volume of maize than ARIMA.

Keywords: futures, trading volume, ARIMA, ANN and maize

Introduction

Maize, sometimes known as corn (*Zea mays* L.), is a major grain crop grown around the world. After rice and wheat, maize is India's third most significant food crop. In India, maize makes up almost 9% of the country's food supply. Its significance comes from the fact that it is utilized not only for human food and animal feed, but also for corn starch manufacturing, corn oil production, and the generation of baby corns.

Forecasting is a technique for producing well-informed forecasts by using past data as the primary source of information for predicting future trends. When the environment is not rapidly changing, forecasting techniques are more accurate.

A futures market is an auction market where people purchase and sell commodity and futures contracts for delivery at a later date. Futures are exchange-traded derivatives contracts that guarantee the delivery of a commodity or security in the future at a certain price. Producers and suppliers of commodities use futures contracts to try to avoid market volatility. These producers and suppliers enter into agreements with an investor who undertakes to bear both the risk and the reward of a volatile market.

Agricultural commodity futures price prediction through long and short term time series network (Hong bing *et al*, 2019). back propagation neural network model is used to forecast monthly futures trading volume for barley, canola, flax, oats, rye and wheat traded on the Winnipeg Commodity Exchange (WCE). The accuracy of forecasting is compared by commodity and forecast horizon. The results show that, with the exception of barley and rye, neural networks can anticipate up to nine months ahead of time and outperform the naive model for all commodities (Kaastra; 1994)

We can explain the usage of order execution procedures and the function of algorithmic trading, the design of trading hours, and the expanding significance of global information flows in driving trade and price by predicting trading volume at specific times of day and using specific information. Large hedgers and speculators looking to lower transaction costs may be able to enhance order execution by forecasting volume better. So it is important to forecast the futures trading volume of maize.

Methodology

Data

Secondary data was used to conduct this study. Cotton futures trading volume is predicted using monthly data from the National Commodity Derivatives Exchange Limited (NCDEX) and the Multi Commodity Exchange (MCX) on open price, close price, open interest, and mean cash price. From October 2011 through December 2020, data will be collected

Subheadings should be as the above heading “2.1 Subheadings”. They should start at the left-hand margin on a separate line.

Artificial Neural Network

An artificial neural network (ANN) is a component of a computer system that mimics how the human brain analyses and processes data. It is the cornerstone of artificial intelligence (AI) and is used to solve issues that would be impossible or difficult to solve by human or statistical standards. ANNs have self-learning abilities, which allows them to improve their performance as more data becomes available.

A neural network is made up of a large number of artificial neurons called units that are stacked in layers. The Artificial Neural Network is made up of three layers.

The input layer accepts data in a variety of forms specified by the programmer. The hidden layer appears in between the input and output layers. It does all the calculations to uncover hidden features and patterns. To produce the output, the weighted total is provided as an input to an activation function.

$$\sum w_i \times x_i + b$$

Activation functions decide whether or not a node should fire. Those who are fired are the only ones who make it to the output layer. There are several activation functions that can be used depending on what type of task we're performing.

Steps

- Input neurons in the input layer transfer information to the hidden layer.
- Data is sent from the hidden layer to the output layer.
- Each neuron has a single output, a weighted input (synapses), and an activation function (which determines the output given an input).
- Synapses are the flexible components that make a neural network into a parameterized system.
- The activation signal is generated by the weighted sum of the inputs and given to the activation function to obtain one output from the neuron.

ARIMA

Auto Regressive Integrated Moving Average (ARIMA) is an acronym for Auto Regressive Integrated Moving Average. ARIMA models are a type of statistical model that can be used to analyze and forecast time series data. It's a more complex version of the auto regressive moving average, with the addition of integration notation.

Assumptions

- Data should be stationary - This indicates that the properties of the series are independent of the time at which it is recorded. A stationary series can also be defined as a white noise series or a series with cyclic pattern.
- ARIMA typically works with a single variable, so data should be univariate. Auto-regression is the process of regressing historical values.

Steps

Step 1: Exploratory analysis: Autocorrelation analysis is performed to determine which historical value has a correlation with the current value, resulting in the p, d, and q estimates for ARIMA models. To investigate cyclic behaviour, use spectral analysis.

Step 2: Make the model fit: Three variables are needed to specify the order in which the model should be fitted to the data: p, d, and q, which are non-negative integers that relate to the order of the autoregressive, integrated, and moving average sections of the model, respectively. The values of p and q are determined using PACF and ACF, respectively. The ARIMA model is estimated using maximum likelihood estimation (MLE). Compute the Akaike's Information Criterion (AIC) and the Schwarz Bayesian Information Criterion (BIC) for a set of models, then explore the models with the lowest AIC and BIC values. Decomposition of patterns. Along with AIC and BIC, we must keep a close watch on those coefficient values and decide whether or not to incorporate them depending on their significance level.

Step 3: Diagnostic measures

Step 4: Forecasting using an ARIMA model: Once the best-suited model for time series data is determined, the parameters of that ARIMA model can be employed as a predictive model for creating forecasts for future values of the time series.

The data are first tested for stationarity using the Augmented Dickey-Fuller test at a significance level of 5%. We take the first difference of the data to make the series stationary if it is non-stationary. Models are estimated equation by equation using the concept of least squares in both circumstances. The forecast is obtained when the coefficients are estimated

Results and Discussion

For forecasting of futures trading volume the forecasting methodologies ARIMA and ANN were used.

ANN

ANN are forecasting methods are based on the simple mathematical models of the brain as they allow complex non linear relationships between the response variables and the predictors. ANN analyze as the brain does so we first need the ANN with our data set by dividing the data into train and test set. Here we divided data into train and test of 70% and 30% respectively. Then the data are scaled and the model is built and the weights are obtained. Here we took open price, close price, mean cash price and Open interests as the independent variables and the trading volume as a dependent variable. The weights obtained from the built model are 0.80196, 0.28269, 0.30229 and -3.82758 for open price, close price, mean cash price and open interest from input layer and bias for the built model are 0.33911 and -0.08402 from the hidden layer 1 and hidden layer 2 respectively. Forecasts are obtained based on these weights and bias and the forecast are compared using Root Mean Square Error, Mean Absolute Percentage Error and Mean Absolute Error.

ARIMA

ARIMA is a method for understanding historical data and forecasting future data in a series. ARIMA uses univariate past trading volume factors to predict futures trading volume. The data for ARIMA should be stationary. ARIMA makes forecasts about the future based on stationary data. The ADF test is used to determine stationarity. The data appears to be stationary based on the ADF test, with a p value of 0.001 at the first difference. So, at initial difference, the data is already stationary, and d=1. p is the order of the "Auto Regressive" to (AR) term. It is the number of y lags that will be utilized as predictors. The amount of AR terms required can be determined using the Partial Auto-Correlation Function (PACF) graphic. In the PACF plot, the order of the AR term will be equal to the number of delays that cross the significance limit. Technically, MA words refer to the error of the lagged forecast. The Auto Correlation Function indicates how many MA terms are needed to remove an auto correlation from a stationarized series. The model is identified as ARIMA based on the ACF and PACF plots, AIC, and BIC (3, 1, 0). The best model's parameters were calculated. The model is written using the following significant parameters:

$$y_t = -0.8101y_{t-1} - 0.7272y_{t-2} - 0.5151y_{t-3}$$

The forecast is obtained by back forecasting (are presented in table).

Table 1: ADF test results

Variables	p-value at level	p-value at first difference
Volume	0.1059*	0.01*
Open price	0.07	0.01*
Close price	0.098	0.01*
Mean cash price	0.1034	0.01*
Open interests	0.04*	0.01*

Table 2: ARIMA model

Aspects	Estimates	Standard error
AR1	-0.8010	0.0776
AR2	-0.7272	0.0824
AR3	-0.5155	0.0724

Table 3: Model performance metrics

Metrics	ARIMA	ANN
RMSE	2116.257	1557.7493
MAPE	80.35	36.06
MAE	1383.261	683.185

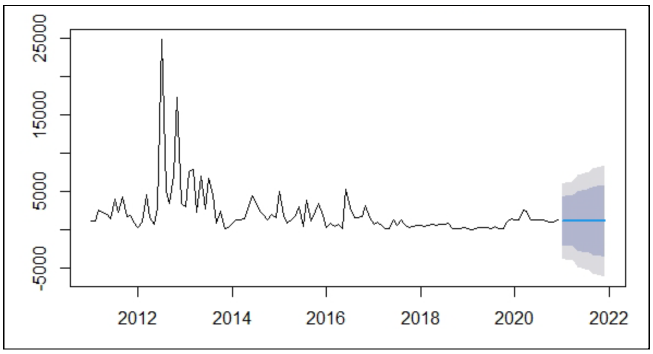

Fig 1: ARIMA forecast

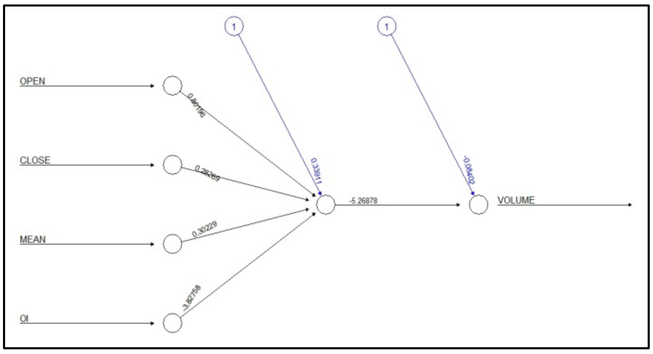

Fig 2: ANN architecture

Conclusion

Forecasted values are compared using model performance metrics such as Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE) and Mean Absolute Error (MAE). Metrics values are compared for ARIMA and ANN using table and it shows that ANN is better than ARIMA as the metrics of ANN are lower than ARIMA. Therefore, the ANN is better in forecasting futures trading volume of cotton than ARIMA.

References

1. Guha B, Bandyopadhyay G. Gold price forecasting using ARIMA model. Journal of Advanced Management Science 2016,4(2).
2. Lasheras FS, de Cos Juez FJ, Sánchez AS, Krzemień A, Fernández PR. Forecasting the COMEX copper spot price by means of neural networks and ARIMA models. Resources Policy 2015;45:37-43.
3. Li GQ, Xu SW, Li ZM. Short-term price forecasting for agro-products using artificial neural networks. Agriculture and Agricultural Science Procedia 2010;1:278-87
4. Merh N, Saxena VP, Pardasani KR. A comparison between hybrid approaches of ANN and ARIMA for Indian stock trend forecasting. Business Intelligence Journal 2010;3(2):23-43
5. Ntungo C, Boyd M. Commodity futures trading performance using neural network models versus ARIMA models. Journal of Futures Markets: Futures, Options, and Other Derivative Products 1998;18(8):965-83.
6. Wang L, Zou H, Su J, Li L, Chaudhry S. An ARIMA-ANN hybrid model for time series forecasting. Systems Research and Behavioral science 2013;30(3):244-59.
7. Zhao H. Futures price prediction of agricultural products based on machine learning. Neural Computing and Applications 2021;33(3):837-50.