Forecasting Gold Prices in India using Time series and Deep Learning Algorithms

P. Sai Shankar, M. Krishna Reddy

Abstract: The primary object of this paper is to compare the traditional time series models with deep learning algorithm. The ARIMA model is developed to forecast Indian Gold prices using daily data for the period from 2016 to 2020 obtained from World Gold Council. We fitted the ARIMA (2,1,2) model which exhibited the least AIC values. In the meanwhile, MLP, CNN, and LSTM models are also examined to forecast the gold prices in India. Mean absolute error, mean absolute percentage error and root mean squared errors were used to evaluate the forecasting performance of the models. Hence, LSTM model superior than that of the other three models for forecasting the gold prices in India.

Keywords: Gold Prices, Box-Jenkins Methodology, ARIMA, Lag Variables, MLP, CNN and LSTM Models

I. INTRODUCTION

Gold is a precious metal and it is completely different from other assets and metals. It is highly liquid and sensitive to price changes (Ranson and Wainwright, 2005). A majority of gold is bought as jewelry items. In the gold consumption, India and China together account almost 60% of worldwide gold jewelry. Gold is consider as a traditional gift in many weddings and often given away on the occasion of festivals, such as Diwali, Akshaya Tritiya and Diwali. Gold plays a unique role as a store of value and hedge risks (Taylor, 1998; Hammoudeh et al., 2010). People investing in gold have mainly two primary objectives, one being hedge against inflation as over a period of time (Baur and McDermott (2010) and Baur and Lucey (2010) and next is to mix your investment basket and hence diversify the risk and will help you reduce the overall volatility of your portfolio. Typically the return on gold investment is almost in line with the inflation rate. In modern days, people has to choose different ways to invest the gold by buying jewellery which is safer way or by purchasing gold coins and bars which is available in public sector banks nowadays or by investing in Gold Exchange traded fund (Gold ETF). Gold ETF is in financial instrument of mutual fund in nature which in turn invests in gold and these are listed in a stock index.

Time series can be defined as a sequence of data points which is ordered in sequence and collected at regular time intervals. Time series approach can be used on any data which is changes over time. Time series are used in statistics, weather forecasting, sales forecasting, stock market predictions etc. In decisions, that involve factor of uncertainty of the future, time series models found one of the most effective methods of forecasting. The classical time series analysis procedure decomposes the time series data into four components: Trend, Seasonal, Cyclic and residual component.

There are a range of techniques to forecast the gold prices. There are classic econometric and statistical techniques like the exponential smoothing method (ESM), autoregressive moving averages(ARMA), auto regressive integrated moving averages (ARIMA) etc., Although these methods are useful to capture the linear relationship and they fail to capture the non-linear characteristics of gold prices. To tackle this problem, artificial intelligence (AI) models are implemented to forecast the gold prices. The common methods are Feedforward neural networks, Multilayer perception (MLP), Recurrent neural networks (RNN), Convolutional neural networks (CNN) and Long short term memory (LSTM).

II. MATERIALS AND METHODS

The data collected from World Gold Council of gold prices in Rupees per gram, daily frequency ranging from January 2016 to December 2020 consisting a total of 1304 observations. The train and test datasets consists of 1277 and 27 respectively. We built the model on train dataset and predictions on test dataset.

Table 1: Average Daily Gold Prices (Rs/gram) in India

<table>
<thead>
<tr>
<th>Average Daily Gold Prices (Rs/gram)</th>
<th>Mon</th>
<th>Tue</th>
<th>Wed</th>
<th>Thu</th>
<th>Fri</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2755.606</td>
<td>2756.874</td>
<td>2755.702</td>
<td>2756.907</td>
<td>2755.783</td>
</tr>
</tbody>
</table>

A. Review of Box and Jenkins Approach

The Box–Jenkins methodology is one of the well known method in the field of time series. This method applies ARIMA models to find the best fit of a time series to past values of this time series, in order to make accurate future forecasts. These methods are also commonly used when the data is not-stationary.

Let \( \{Z_t\} \) be the time series. Then \( \{Z_t\} \) is stationary if \( E(Z_t) = \mu \) and \( V(Z_t) = \sigma^2 \) for all \( t \). Otherwise it is non –stationary. Let \( Z_1,Z_2,\ldots, Z_N \) be an observed sample. If there is no trend line and constant variance for all values of \( t \) in the time series plot, then the time series is said to be stationary. Alternatively, if the ACF of sample dies out slowly is an indication for stationary.
The Box – Jenkins Methodology is valid for only stationary time series data. If the data is non – stationary, we convert it into stationary by stabilizing mean using successive differencing and stabilizing variance using logarithmic transformation. The Auto Regressive Integrated Moving Average model for the time series is denoted by ARIMA(p, d, q) and is defined by $\phi(B) \, \theta(B) \, z_t = \theta(B) \, \theta(B) \, \alpha_t$, where $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$ is a polynomial in B of order p and is known as Auto Regressive (AR) operator, $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q$ is a polynomial in B of order q and is known as Moving Average (MA), operator. $\nabla = 1-B$, B is the Backward shift operator $B^k z_t = z_{t-k}$ and d is the number of differences required to arrive stationarity timeseries. AR(p), MA(q) and ARMA(p, q) may be obtained as particular case of it with parameter values (p, 0, 0), (0, 0, q) and (p, 0, q), respectively.

The four steps procedure for Box-Jenkins method: Identification of model, Parameter estimation, diagnostic checking and finally model forecast.

B. Review on Deep Learning Algorithms

In the study of ANN, the most popular algorithm is a feed-forward neural network (FFNN). It has only forward connections in between the neurons. ANNs are commonly used for time series forecasting. ANNs can model any form of unknown relationship in the data with few assumptions. ANNs models trained on train dataset and it can generalise on unseen data. Traditional time series are capable of modelling only linear relationships whereas ANNs are capable of modelling any form of relationship in the data, especially non-linear relationships (Hornik et al., 1989). The forecasting accuracies are also better than other traditional time series models.

Tang et al. (1991) compared ARIMA model with deep learning model for forecasting and results show that deep learning models are better than the ARIMA. Clavera and Torra (2014) studied performance of ANNs and traditional time series ARIMA and finally showed that ANNs are higher accuratethan ARIMA model to a tourism-demand forecasting problem. Zhang & Kline (2007) paper revealed that the accuracy of the model is depends on the input variables, best network structure and also training procedure.

Multilayer Perceptron (MLP)

A multilayer perceptron(MLP) is a class of feedforward artificial neural network(ANN). MLP Neural Networks can be formed a relation between the input and output data and it can be modelled by neural networks (Bildirici et al., 2010). It is an extension of feed forward neural network. A multilayer perceptron neural network with two hidden layers is depicted in Fig. 1 which consists of three types of layers—the first one input layer and output layer these two connected with hidden layers. In MLP, the input nodes are connected to the output nodes with the help of hidden layers in between the two layers. The input will be changed under the effect of the hidden layers, which behave in nonlinear way and the results of the changes multiply the weights will be transferred to the output nodes. Unlike to a feed forward network in a MLP the data flows in the forward direction from input to output layer. The number of neurons are trained with the back propagation learning algorithm. The major use cases of MLP are forecasting, recognition, prediction, pattern classification, and approximation.

![Figure 1. A simple schema of a three layer perceptron neural network](image)

In Eq. (1), the variable, x, is the input vector, $\theta$ is called the activation function and the output result of the MLP could be a number or a vector of zero numbers. The variable, w, is the weights and f is the output function. Log-sigmoid, ReLU, and hyperbolic tangent are the different activation functions, which can be found in the neural networks literature. Activation function is the core of any node in the hidden layers of MLP changes the input of the node to the output. ReLU activation function is used in this research which is presented in Eq. (2).

In deep learning models, the most commonly used activation function is Rectified Linear Unit. The function returns for any positive value x it returns that value back and 0 if it receives any negative input. It can be written as $f(x) = max(0, x)$.

![Figure 2. ReLU activation function](image)

Long Short Term Memory (LSTM)

RNNs are a most powerful type of artificial neural network that can internally maintain memory of the input. In Recurrent Neural networks(RNN) studies, one of the most common algorithm is LSTM. The LSTM network is first developed by Hochreiter & Schmidhuber. Since RNN has maintain memory gate, this enables them particularly used for solving the time series problems.
One of the drawbacks of RNNs is to suffer from a problem called vanishing gradient which leads to the model learning becoming too slow or stopping altogether. LSTM models are developed such that to avoid the vanishing gradient problem. LSTMs have longer memories and can learn from inputs that are separated from each other by long time lags. An LSTM has three gates: an input gate which determines whether or not to let the new input in, an output gate which decides what information to output and finally the import gate which deletes information that is not useful. These three gates are analog gates based on the sigmoid function which works on the range 0 to 1 which is shown in Fig. 3 below. LSTM component is a end-to-end feature (enclosed in green dash box) with a three-layer model (consisted of input layer, LSTM unit layer and output layer).

Convolution Neural Networks (CNN)

In deep learning, a CNN is a class of Deep learning neural network, most commonly used for vision and image processing-based classification problems (object detection, image classification, image segmentation, etc.). CNN are regularized versions of MLP. In MLP, each neuron in one layer connected with all then neurons of next layer. The “full connectivity” of these networks make them prone overfitting of the data. Regularization is one way to deal the overfitting in the data. CNN take a different approach towards regularization.

The backpropagation process has been used to train the model in CNN model, The tuning parameters of the CNN:
- Number of hidden layers
- Number of neurons in each layer
- Learning rate
- Activation functions: ReLu, Sigmoid
- Epochs,

Batch size, Optimization algorithms: are SGD, ADAM and RMSProp

Measures:
- If \( Z_t \) is the actual price for period \( t \) and \( \hat{Z}_t \) is the forecast, then the error is defined as \( e = Z_t - \hat{Z}_t \). The following measures may be considered
  - Root Mean Square Error (RMSE) = \( \sqrt{\frac{1}{n} \sum_{t=1}^{n} e_t^2} \)
  - Mean Absolute Percent Error (MAPE) = \( \frac{1}{n} \sum_{t=1}^{n} \left| \frac{e_t}{z_t} \right| \times 100 \)
  - Mean Absolute Error (MAE) = \( \frac{1}{n} \sum_{t=1}^{n} |e_t| \)

III. ARIMA MODEL

Identification of Model

To fit any traditional time series model, the initial step is to check the stationarity condition. Stationarity can be determined from a time series plot which should show constant mean and variance. It can also be checked from an ACF plot. Specifically, very slow decay in the lags in ACF plots indicates a non-stationarity.

Testing Stationary of Time Series

The stationary condition can also be determined by the test called Augmented Dickey Fuller (ADF) unit root test. The hypothesis of the test are:
- The null Hypothesis: \( H_0: X_t \) is non–stationary
- Alternative hypothesis: \( H_1: X_t \) is stationary

The p-value of the Augmented Dickey-Fuller (ADF) test equals 0.996 and it is larger than the value of \( \alpha = 0.05 \). This result indicates that the time series of daily Gold Prices in India is not stationary.
All the above plots and results confirm that the original time series data is non-stationary, and need to apply some transformations to convert it into the stationary series. The differencing method is used to convert non-stationary time series into stationary time series of the data. In this paper, non-seasonal difference of order 1 (i.e. d=1) is sufficient to achieve stationary in mean and variance. The derived variable $W_t = \nabla^1 \tilde{z}_t$ can now be examined for stationarity.

Figure (7) displays the time series plot of the data after first differencing the series and indicates that the time series is a stationary series. To make sure of that, we conduct the unit roots test (Augmented Dickey-Fuller) for the transformed series $Z_t$.

The p-value of the ADF test equals 0.00 which is less than the value of alpha=0.05 and this indicates that the non-stationary hypotheses of the differenced daily Gold Prices data is rejected and this demonstrates the success of difference transformation for the time series data of daily Gold Prices data. Thus, the series became stationary.

The panel of plots-Time series, ACF and PACF displays the time series plot of the data after first differencing the series and indicates that the time series is stationary.

The ARIMA models have been examined and estimated as shown in table (2) below. The best ARIMA model is chosen through the AIC criteria if it shows the lowest values of these criteria.

It is shown in table (2) that the ARIMA (2, 1, 2) is significant with respect to parameters as well as adequacy of the model. This means that the ARIMA(2, 1, 2) model is superior among all the other models.

<table>
<thead>
<tr>
<th>S.No</th>
<th>Parameters</th>
<th>AIC</th>
<th>S.No</th>
<th>Parameters</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(2, 1, 2)</td>
<td>12,359.2</td>
<td>6</td>
<td>(2, 2, 1)</td>
<td>12,370.9</td>
</tr>
<tr>
<td>2</td>
<td>(4, 2, 1)</td>
<td>12,361.7</td>
<td>7</td>
<td>(1, 2, 1)</td>
<td>12,371.2</td>
</tr>
<tr>
<td>3</td>
<td>(4, 1, 1)</td>
<td>12,364.5</td>
<td>8</td>
<td>(3, 2, 2)</td>
<td>12,372.5</td>
</tr>
<tr>
<td>4</td>
<td>(4, 1, 2)</td>
<td>12,366.4</td>
<td>9</td>
<td>(3, 2, 1)</td>
<td>12,372.8</td>
</tr>
<tr>
<td>5</td>
<td>(3, 1, 2)</td>
<td>12,368.1</td>
<td>10</td>
<td>(1, 2, 2)</td>
<td>12,373.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The fitted ARIMA model for the daily gold price in India is $(1 + 0.52B + 0.93B^2) y_t = (1 − 0.50B − 0.97B^2)e_t$.
Diagnostic Tests:
The autocorrelations and partial auto correlations of the residuals plots are used for diagnostic checking.

Analysis of Residuals

Figure (8) show the estimated autocorrelation function for the residuals of the ARIMA (2, 1, 2) model for the time series of daily Gold prices.

Box-Ljung Q-test statistic is used to test the adequacy of the model. The hypothesis on the model is below:

H₀: The ARIMA (2,1,2) model is adequate.
H₁: The model is ARIMA(2,1,2) inadequate.

Table (4): Ljung-Box Test Statistic

<table>
<thead>
<tr>
<th>Ljung-Box (Q)</th>
<th>Prob(Q)</th>
<th>Jarque-Bera (JB)</th>
<th>Prob(JB)</th>
<th>Heteroskedasticity</th>
<th>Skew</th>
<th>Prob(H) (two-sided)</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.06</td>
<td>0.81</td>
<td>7091.2</td>
<td>0</td>
<td>3.21</td>
<td>-0.67</td>
<td>0.00</td>
<td>14.47</td>
</tr>
</tbody>
</table>

From the Table(4), the probability value of the test statistic is 0.81 which is greater than 0.05, therefore, we fail to reject H₀. Hence, we conclude that the ARIMA (2,1,2) model is an adequate model for the given time series data.

IV. DEEP LEARNING MODELS

A. Multilayer Perceptron (MLP)

Python software has been used to train the MLP model on evaluate the model on test dataset.

Structure of the Network:
The MLP model consists of three: input layer, a hidden layer and an output layer. Two input neurons considered in this model, each representing the values of lag1 (previous day price in the same week i.e. lag1 gold price) and lag2 (lag 2 gold price). One output unit is needed in this model which indicates the forecasts of daily gold price. Trial and error approach can be used to find the optimal number of hidden units in the neural network. Typically we can use either the forward selection method or backward selection to arrive the optimum the hidden layer units. In the forward selection method, we choose a small number of hidden neurons then compute the network performance using RMSE, MAE and MAPE values. In the next step increment the hidden neurons by one until train and test error is acceptably small or no further improvement is noted.

B. Long Short-Term Memory (LSTM)

Python software has been used to train the LSTM model on evaluate the model on test dataset.

Structure of the Network:

C. Convolutional neural networks (CNNs)

Python software has been used to train the CNN model on evaluate the model on test dataset.

Structure of the Network:

V. RESULTS AND DISCUSSION

The train dataset is used to train the models. Several traditional and deep learning models ARIMA, MLP, CNN and LSTM are examined. Once we trained the dataset using training dataset then we test the model on test dataset. The performance measures are showed in the Table 5. Figures 9–12 represents the line charts of actual gold price and predicted gold prices.
The main object of the paper is to compare the traditional time series models with deep learning algorithms. The historical data is collected from the World Gold Council of Blue Eyes Intelligence Engineering, a leading engineering and IT company.

The study is based on secondary analysis of data collected from the World Gold Council of Blue Eyes Intelligence Engineering, a leading engineering and IT company.

The experimental results revealed that the LSTM model is the best among the four methods in terms of forecasting gold prices in India. The LSTM model has the lowest MAPE, RMSE and MAE compared to the other three comparative models in terms of forecasting gold prices in India.

<table>
<thead>
<tr>
<th>LSTM</th>
<th>CNN</th>
<th>MLP</th>
<th>ARIMA</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>0.65</td>
<td>0.71</td>
<td>0.72</td>
</tr>
<tr>
<td>RMSE</td>
<td>33.28</td>
<td>32.23</td>
<td>31.46</td>
</tr>
<tr>
<td>MAPE</td>
<td>29.31</td>
<td>29.7</td>
<td>25.46</td>
</tr>
</tbody>
</table>

The LSTM model is superior in incorporating some explanatory variables which are influence the gold price. The LSTM model can be improved further by incorporating other factors such as the US dollar, crude oil, inflation and bank rates into the forecast model. Our future research work is mainly to incorporate the gold prices to ensure the accuracy of gold prices in India using historical data. This study is based on secondary data collected from January 2016 to December 2020. The LSTM model has the lowest MAPE, RMSE and MAE compared to the other three comparative models in terms of forecasting gold prices in India.

The main object of the paper is to compare the traditional time series models with deep learning algorithms. The historical data is collected from the World Gold Council of Blue Eyes Intelligence Engineering, a leading engineering and IT company.

The experimental results revealed that the LSTM model is the best among the four methods in terms of forecasting gold prices in India. The LSTM model has the lowest MAPE, RMSE and MAE compared to the other three comparative models in terms of forecasting gold prices in India.

<table>
<thead>
<tr>
<th>LSTM</th>
<th>CNN</th>
<th>MLP</th>
<th>ARIMA</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>0.65</td>
<td>0.71</td>
<td>0.72</td>
</tr>
<tr>
<td>RMSE</td>
<td>33.28</td>
<td>32.23</td>
<td>31.46</td>
</tr>
<tr>
<td>MAPE</td>
<td>29.31</td>
<td>29.7</td>
<td>25.46</td>
</tr>
</tbody>
</table>

The LSTM model is superior in incorporating some explanatory variables which are influence the gold price. The LSTM model can be improved further by incorporating other factors such as the US dollar, crude oil, inflation and bank rates into the forecast model. Our future research work is mainly to incorporate the gold prices to ensure the accuracy of gold prices in India using historical data. This study is based on secondary data collected from January 2016 to December 2020. The LSTM model has the lowest MAPE, RMSE and MAE compared to the other three comparative models in terms of forecasting gold prices in India.
<table>
<thead>
<tr>
<th>Date</th>
<th>Price</th>
<th>Pred_Price</th>
<th>Date</th>
<th>Price</th>
<th>Pred_Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>25-11-20</td>
<td>4268.6</td>
<td>4328.7</td>
<td>15-12-20</td>
<td>4383.6</td>
<td>4339.7</td>
</tr>
<tr>
<td>26-11-20</td>
<td>4298.6</td>
<td>4295.0</td>
<td>16-12-20</td>
<td>4414.2</td>
<td>4374.3</td>
</tr>
<tr>
<td>27-11-20</td>
<td>4249.6</td>
<td>4299.6</td>
<td>17-12-20</td>
<td>4458.8</td>
<td>4410.9</td>
</tr>
<tr>
<td>30-11-20</td>
<td>4230.0</td>
<td>4264.7</td>
<td>18-12-20</td>
<td>4450.0</td>
<td>4452.2</td>
</tr>
<tr>
<td>01-12-20</td>
<td>4289.8</td>
<td>4282.8</td>
<td>21-12-20</td>
<td>4463.5</td>
<td>4459.9</td>
</tr>
<tr>
<td>02-12-20</td>
<td>4342.1</td>
<td>4279.5</td>
<td>22-12-20</td>
<td>4416.7</td>
<td>4465.5</td>
</tr>
<tr>
<td>03-12-20</td>
<td>4372.4</td>
<td>4333.6</td>
<td>23-12-20</td>
<td>4446.3</td>
<td>4431.8</td>
</tr>
<tr>
<td>04-12-20</td>
<td>4362.2</td>
<td>4369.1</td>
<td>24-12-20</td>
<td>4437.4</td>
<td>4443.2</td>
</tr>
<tr>
<td>07-12-20</td>
<td>4421.0</td>
<td>4368.3</td>
<td>25-12-20</td>
<td>4411.9</td>
<td>4443.3</td>
</tr>
<tr>
<td>08-12-20</td>
<td>4434.9</td>
<td>4411.0</td>
<td>28-12-20</td>
<td>4428.8</td>
<td>4421.6</td>
</tr>
<tr>
<td>09-12-20</td>
<td>4360.8</td>
<td>4435.4</td>
<td>29-12-20</td>
<td>4431.8</td>
<td>4426.6</td>
</tr>
<tr>
<td>10-12-20</td>
<td>4352.9</td>
<td>4381.8</td>
<td>30-12-20</td>
<td>4456.3</td>
<td>4434.9</td>
</tr>
<tr>
<td>11-12-20</td>
<td>4362.6</td>
<td>4358.5</td>
<td>31-12-20</td>
<td>4460.0</td>
<td>4454.4</td>
</tr>
<tr>
<td>14-12-20</td>
<td>4327.8</td>
<td>4364.1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

REFERENCES


AUTHORS PROFILE

P Sai Shankar, is a data science enthusiast with a strong inclination towards problem-solving and propelling data-driven decisions. He is currently pursuing his Ph.D in Statistics under Dr. M Krishna Reddy from University college of Science (Osmania a University). He has worked on multiple projects in the field of machine learning.He is a keen learner of new Technologies. His field of research is Time series, Machine learning and Deep learning, E-mail: saishankar26@gmail.com

Prof. M. Krishna Reddy, has more than 42 years of teaching and research experience. He served as Professor, Head and Chairperson, Board of Studies at Department of Statistics, Osmania University, Hyderabad, Telangana. Presently, he is Professor of statistics at CVR college of Engineering, Hyderabad. He guided 15 Ph.Ds in Statistics and more than 40 publications in national and international journals. E-mail: reddympk54@gmail.com