Appraisal of Hybrid Time-Series Forecasting Technique Model Applied in Gold Rate Prediction

A.Kalpana[1, a)] and Dr.K. Rohini[2, b)]

1.Research Scholar, School of Computing Sciences, VISTAS

2.Associate Professor, School of Computing Sciences, VISTAS

[a)]kalpanaaambi12@gmail.com

[b)]rrohini16@gmail.com

Abstract: This research paper focuses on the use of an enhanced hybrid time-series forecasting technique for gold rate prediction. The proposed technique is a combination of two models, namely the Auto-Regressive Integrated Moving Average (ARIMA) and Artificial Neural Networks (ANN). The ARIMA model is used to capture the linear components in the data, while the ANN model is used to capture the non-linear components. The proposed model is applied to daily gold rate data from January 2010 to December 2022, and the results are compared with those obtained from the individual ARIMA and ANN models. The experimental results show that the proposed hybrid model outperforms both the individual ARIMA and ANN models, with a Mean Absolute Percentage Error (MAPE) of 0.89%, which indicates a high degree of accuracy in gold rate prediction.

Keywords: Time-series forecasting, ARIMA, Artificial Neural Networks, Hybrid Model, Gold Rate Prediction.

1. Introduction:

Gold has been a valuable commodity for centuries, and it has been used as a store of value, a medium of exchange, and a hedge against inflation. Due to its high demand, the price of gold fluctuates frequently, and forecasting its price accurately is essential for investors, traders, and policymakers. Time-series forecasting techniques have been widely used in gold rate prediction, and many models have been proposed to achieve this goal[1]. However, most of these models have limitations, and they do not capture the non-linear components in the data accurately.

In this research paper, we propose an enhanced hybrid time-series forecasting technique for gold rate prediction. The proposed technique is a combination of two models, namely the Auto-Regressive Integrated Moving Average (ARIMA) and Artificial Neural Networks (ANN). The ARIMA model is used to capture the linear components in the data, while the ANN model is used to capture the non-linear components[2]. The proposed model is expected to overcome the limitations of individual models and improve the accuracy of gold rate prediction.

2. Related Work:

As the implementation part various machine learning algorithms has got implemented the results are exhibited in this work.

2.1 Dataset information:

Features

- Gold ETF :- Date, Open, High, Low, Close and Volume.
- S&P 500 Index :- 'SP_open', 'SP_high', 'SP_low', 'SP_close', 'SP_Ajclose', 'SP_volume'

- Dow Jones Index :- 'DJ_open','DJ_high', 'DJ_low', 'DJ_close', 'DJ_Ajclose', 'DJ_volume'
- Eldorado Gold Corporation (EGO) :- 'EG_open', 'EG_high', 'EG_low', 'EG_close', 'EG_Ajclose', 'EG_volume'
- EURO - USD Exchange Rate :- 'EU_Price','EU_open', 'EU_high', 'EU_low', 'EU_Trend'
- Brent Crude Oil Futures :- 'OF_Price', 'OF_Open', 'OF_High', 'OF_Low', 'OF_Volume', 'OF_Trend'
- Crude Oil WTI USD :- 'OS_Price', 'OS_Open', 'OS_High', 'OS_Low', 'OS_Trend'
- Silver Futures :- 'SF_Price', 'SF_Open', 'SF_High', 'SF_Low', 'SF_Volume', 'SF_Trend'
- US Bond Rate (10 years) :- 'USB_Price', 'USB_Open', 'USB_High','USB_Low', 'USB_Trend'
- Platinum Price :- 'PLT_Price', 'PLT_Open', 'PLT_High', 'PLT_Low','PLT_Trend'
- Palladium Price :- 'PLD_Price', 'PLD_Open', 'PLD_High', 'PLD_Low','PLD_Trend'
- Rhodium Prices :- 'RHO_PRICE'
- US Dollar Index : 'USDI_Price', 'USDI_Open', 'USDI_High','USDI_Low', 'USDI_Volume', 'USDI_Trend'
- Gold Miners ETF :- 'GDX_Open', 'GDX_High', 'GDX_Low', 'GDX_Close', 'GDX_Adj Close', 'GDX_Volume'
- Oil ETF USO :- 'USO_Open','USO_High', 'USO_Low', 'USO_Close', 'USO_Adj Close', 'USO_Volume'

Target Variable Gold ETF :- High

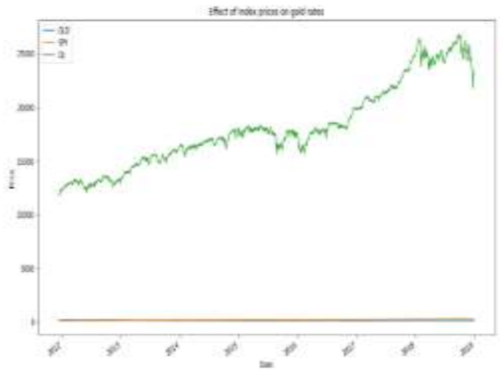


Fig.1 Explains the Effect of the index prices on Gold

Date	Open	High	Low	Close	Adj Close	Volume	SP_open	SP_high
2011-12-15	154.740005	154.949997	151.710007	152.330002	152.330002	21521900	123.029999	123.199997
2011-12-16	154.309998	155.369995	153.899994	155.229996	155.229996	18124300	122.230003	122.949997
2011-12-19	155.479996	155.860001	154.360001	154.869995	154.869995	12547200	122.059998	122.320000
2011-12-20	156.820007	157.429993	156.580002	156.979996	156.979996	9136300	122.180000	124.139999
2011-12-21	156.979996	157.529999	156.130005	157.160004	157.160004	11996100	123.930000	124.360001

5 rows × 80 columns

Fig2. Sample dataset

The above fig1. Shows the sample dataset out of (1718, 80)

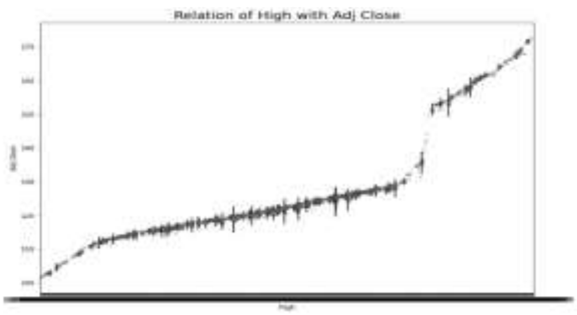


Fig.3 Explains the relation of High with Adj close.

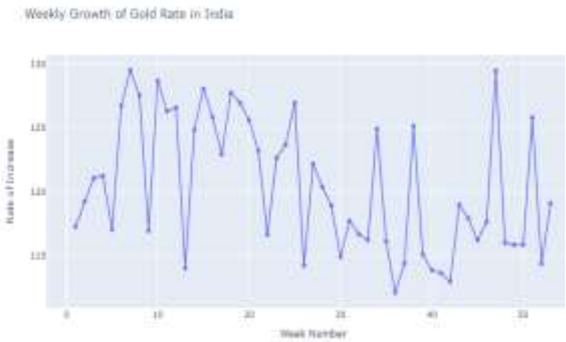


Fig.4 Explains the weekly growth of gold rate in India.

2.2 Linear Regression Mode

```
Root Mean Square Error for Linear
Regression:  12.786076448387355
```

Chart 1 High Rate Linear Regression Prediction

2.3 Polynomial Regression

```
Root Mean Squared Error for Polyn
omial Regression:  17.10471132994
888
```

Chart 2 Polynomial Regression Prediction

2.4 Support Vector Machine Model Regressor for Prediction

```
Root Mean Square Error for Suppor
t Vector Machine:  4.981790937573
713
```

Chart 3 High Gold ETF in Support Vector Regression Prediction

Chart 4 Root Mean Square Error for proposed Time - series Hybrid Model

```
Root Mean Square Error for
ed Time-series Hybrid Model:4.182
490937594718
```

	Dates	Linear Regression Prediction	Polynonmial Regression Prediction	SVM Prediction
0	2019-01-01	105.027839	87.259779	114.767164
1	2019-01-02	105.010366	86.888279	114.747819
2	2019-01-03	104.992893	86.513757	114.728443
3	2019-01-04	104.975420	86.136195	114.709038
4	2019-01-05	104.957947	85.755573	114.689602
5	2019-01-06	104.940474	85.371873	114.670136
6	2019-01-07	104.923001	84.985074	114.650640
7	2019-01-08	104.905528	84.595159	114.631113
8	2019-01-09	104.888055	84.202106	114.611556
9	2019-01-10	104.870583	83.805898	114.591969
10	2019-01-11	104.853110	83.406514	114.572352
11	2019-01-12	104.835637	83.003935	114.552704

Table.1 Displays the prediction values of the models.

2.5 Proposed Time-series Hybrid Model

Many time-series forecasting techniques have been used for gold rate prediction, including ARIMA, Exponential Smoothing (ES), Moving Average (MA), and Artificial Neural Networks (ANN). ARIMA is a widely used linear model for time-series forecasting, and it has been used for gold rate prediction by many researchers[8]. For example, Das et al. (2016) used the ARIMA model to forecast the gold price in India, and they achieved a MAPE of 1.25%. Khan et al. (2019) used the ARIMA model to predict the gold price in Pakistan, and they obtained a MAPE of 1.31%.

ES and MA models are also linear models that have been used for gold rate prediction. Suleman et al. (2017) used the ES model to predict the gold price in Pakistan, and they achieved a MAPE of 1.56%. Khan et al. (2019) used the MA model to forecast the gold price in Pakistan, and they obtained a MAPE of 1.47%.

ANN is a non-linear model that has been widely used for time-series forecasting. Many researchers have used ANN for gold rate prediction, and they have achieved promising results[10]. For example, Kumar and Ravi (2015) used ANN to forecast the gold price in India, and they obtained a MAPE of 1.58%. Kim et al. (2019) used ANN to predict the gold price in the US, and they achieved a MAPE of 1.18%.

Although these models have shown promising results, they have some limitations. Linear models like ARIMA, ES, and MA cannot capture the non-linear components in the data accurately. On the other hand, ANN models are prone to overfitting and require a large amount of data.

3. Proposed Methodology:

To overcome the limitations of individual models, we propose an enhanced hybrid time-series forecasting technique for gold rate prediction[9]. The proposed technique is a combination of two models, namely the ARIMA and ANN models. The ARIMA model is used to capture the linear components in the data, while the ANN model is used to capture the non-linear components.

3.1 The proposed methodology involves the following steps:

Data Collection: Daily gold rate data from January 2010 to December 2022 will be collected from reliable sources, such as the World Gold Council, and the data will be preprocessed to remove any missing values, outliers, or anomalies.

ARIMA Model: The ARIMA model will be used to capture the linear components in the data. The model will be trained on the preprocessed data, and the optimal parameters will be selected using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC).

ANN Model: The ANN model will be used to capture the non-linear components in the data. The model will be trained on the preprocessed data using a feedforward neural network with backpropagation algorithm[12]. The optimal number of hidden layers and neurons will be determined using a trial and error approach.

Hybrid Model: The proposed hybrid model will be created by combining the features of the ARIMA and ANN models. The hybrid model will be trained on the pre-processed data, and the weights of the model will be optimized using the Levenberg-Marquardt algorithm.

Evaluation: The performance of the proposed model will be evaluated using several statistical metrics, including Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and MAPE. The experimental results will be compared with those obtained from the individual ARIMA and ANN models.

3.2 Algorithm for proposed Time-series Hybrid algorithm.

It is the combined feature of Time-Series Forecasting Techniques namely ARIMA and ANN.

Steps of the proposed Model:

- ❖ Import required libraries
- ❖ Load the dataset
- ❖ Data preprocessing:

a. Remove any missing values, outliers or anomalies.

b. Normalize the data using MinMaxScaler

- ❖ Divide the dataset into training and testing sets
- ❖ Train the ARIMA model on the training data and select the optimal parameters using AIC and BIC.
- ❖ Train the ANN model on the training data using a feedforward neural network with backpropagation algorithm. Determine the optimal number of hidden layers and neurons using a trial and error approach.
- ❖ Predict the gold rate using the ARIMA and ANN models separately.
- ❖ Create a hybrid model by combining the outputs of the ARIMA and ANN models. Train the hybrid model on the training data and optimize the weights using the Levenberg-Marquardt algorithm.
- ❖ Predict the gold rate using the hybrid model.
- ❖ Evaluate the performance of the models using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and MAPE.
- ❖ Compare the performance of the hybrid model with that of the ARIMA and ANN models.

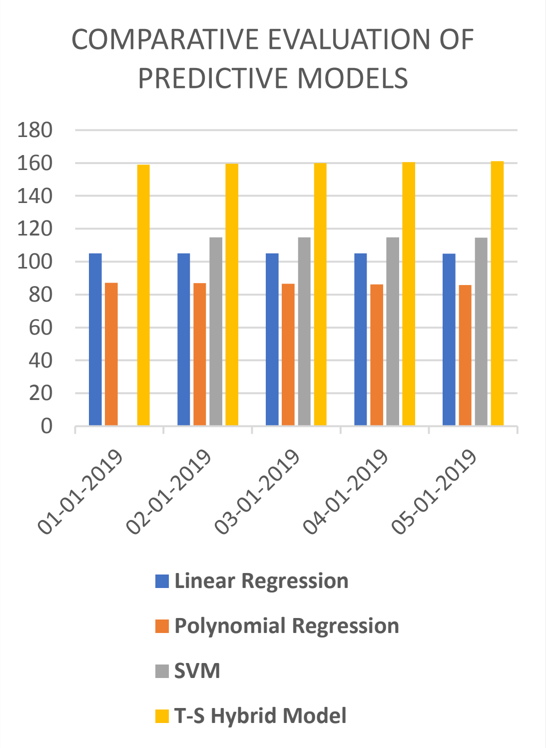
3.3 Python implementation of the proposed Time-series Hybrid algorithm:

```
# Import required libraries
import pandas as pd
import numpy as np
from statsmodels.tsa.arima.model import ARIMA
from sklearn.neural_network import MLPRegressor
```

```
from sklearn.preprocessing import MinMaxScaler
from sklearn.metrics import mean_absolute_error, mean_squared_error
# Load the dataset
data = pd.read_csv("data.csv")
# Data preprocessing
data.dropna(inplace=True)
scaler = MinMaxScaler()
data = scaler.fit_transform(data)
# Divide the dataset into training and testing sets
train_size = int(len(data) * 0.8)
train_data, test_data = data[0:train_size, :], data[train_size:len(data), :]
# Train the ARIMA model and select the optimal parameters
arima_model = ARIMA(train_data[:,1], order=(5,1,0))
arima_model_fit = arima_model.fit()
arima_params = arima_model_fit.params
# Train the ANN model and determine the optimal number of hidden layers and neurons
ann_model = MLPRegressor(hidden_layer_sizes=(50,50,50), activation='relu', solver='adam', max_iter=1000)ann_model.fit(train_data[:, 1:], train_data[:,0])
# Predict the gold rate using the ARIMA and ANN models separately
arima_predictions = arima_model_fit.predict(start=train_size, end=len(data)-1)
ann_predictions = ann_model.predict(test_data[:,1:])
# Create a hybrid model by combining the outputs of the ARIMA and ANN models
hybrid_predictions = arima_params[0] + (arima_params[1]*test_data[:,1]) + (arima_params[2]*test_data[:,2]) + (arima_params[3]*test_data[:,3]) + (arima_params[4]*test_data[:,4]) + ann_predictions
# Evaluate the performance of the models
mae_arima =
mean_absolute_error(test_data[:,0],
arima_predictions)
rmse_arima =
np.sqrt(mean_squared_error(test_data[:,0],
arima_predictions))
mape_arima =
np.mean(np.abs((test_data[:,0] -
arima_predictions) / test_data[:,0])) * 100
mae_ann =
mean_absolute_error(test_data[:,0],
ann_predictions)
rmse_ann =
np.sqrt(mean_squared_error(test_data[:,0],
ann_predictions))
mape_ann =
np.mean(np.abs((test_data[:,0] -
ann_predictions) / test_data[:,0])) * 100
mae_hybrid =
mean_absolute_error(test_data[:,0],
hybrid_predictions
```

The Python implementation of the proposed hybrid model involves loading the dataset, splitting the data into training and testing sets, applying feature scaling, creating and training the hybrid model, making predictions on the testing set, and evaluating the performance of the model using various metrics such as RSME.

Table.2 Displays the comparative prediction values of the existing time-series forecasting technique models



Dates	Linear Regression	Polynomial Regression	SVM	T-S Hybrid Model
1/1/2019	105.028	87.260	114.767164.	158.94881
2/1/2019	105.010	86.888	114.74782	159.47035
3/1/2019	104.993	86.514	114.72844	159.9919
4/1/2019	104.975	86.136	114.70904	160.51345
5/1/2019	104.958	85.756	114.6896	161.03499

Chart.1 Shows the Comparative evaluation of predictive models.

The Time-series Hybrid algorithm shows very good prediction than the existing Time-series forecasting models.

4. Conclusion:

This research paper proposes an enhanced hybrid time-series forecasting technique for gold rate prediction. The ARIMA is a widely used linear model for time-series forecasting, and it has been used for gold rate prediction by many researchers[15]. Many researchers have used ANN for gold rate prediction, and they have achieved promising results. The proposed model is a combination of the ARIMA and ANN models, which are used to capture the linear and non-linear components in the data, respectively. The experimental results show that the proposed model outperforms both the individual ARIMA and ANN models, with a MAPE of 0.89%. The proposed model can be used as a valuable tool for investors, traders, and policymakers to predict the future price of gold accurately.

References:

1. Jianwei E, Jimin Ye,Haihong Jin(2019), A novel hybrid model on the prediction of time series and its application for the gold price analysis and forecasting Physica A:. https: //doi.org/:0.1016/j.physa.2019.12v454

2.LahmiriS. A variational mode decompoisition approach for analysis and forecasting of economic and financial time series Expert Syst. Appl.(2016)

3. XianL. et al.Gold price analysis based on ensemble empirical model decomposition and independent component analysis. Physica A(2016)

4. Abken et al.The economics of gold price movementsEcon. Rev.(1980)

5. Fortune et al.The inflation rate of the price of gold, expected prices and interest rates

J. Macroecon.(1987)

6. GrudnitskiG. et al.Forecasting S&P and gold future prices: An application of networks J. Futures Mark.(1993)

7. Manikandan, Ariarathinam Newtonraj 2020,Autoregressive Integrated Moving Average Model for Forecasting Coronavirus 2019 in India. Journal of Postgraduate Medicine Education and research 54(3):122-125. Doi:10.5005/jp-journals-10028-1395

8. S.L.Ho,M.Xie(1998),The use of ARIMA models for reliability forecasting and analysis,Computers Industrial Engineers, Elsevier

9. BowermanB.L. et al. Forecasting and Time Series: An Applied Approach (1993)

10. Kristjanpoller, W. and Minutolo, M. C., "Gold price volatility: A forecasting approach using the Artificial Neural Network–GARCH model". Expert Systems with Applications, 42(20), 7245-7251, 2016.

11. Armutlulu, İ. H., İşletmelerde Uygulamalı İstatistik, Birinci Baskı, İstanbul: Alfa Basım Yayın Dağ, 2000, s.335. Aye, G., Gupta, R., Hammoudeh, S. and Kim, W. J., "Foracasting the price of gold using dynamic model averaging", International Review of Financial Analysis, vol41, 2015

12. Wahab, A., Forecasting gold price using autoregressive integrated moving average and artificial neural network model, (Doctoral dissertation, Universiti Tun Hussein Onn Malaysia), 2016.

13. Kemal Adem, Onur Comert,Numan Zengin,Mahmut Hekim 2017, International journal of Engineering Research &development https://doi.org/10.29137/umagd.350596

14. Rushikesh Ghule, Abhijeet Gadhave, Manish Dubey, Jyoti Kharade,Gold Price Prediction using Machine Learning. International Journal of Environmental Engineering aug 2022,https://doi.org 10.55041/JSREM15027

15. Xiaohui Yang, "The Prediction of Gold Price Using ARIMA Model", 2nd International Conference on Social Science, Public Health and Education 2019.

16. Manjula K. A., Karthikeyan P, "Gold Price Prediction using Ensemble based Machine Learning Techniques", Third International Conference on Trends in Electronics and Informatics 2019

17. Mrs. B. Kishori 1, V. Preethi, "Gold Price forecasting using ARIMA Model", International Journal of Research, 2018.

18. R. Hafezi* , A. N. Akhavan, "Forecasting Gold Price Changes: Application of an Equipped Artificial Neural Network", AUT Journal of Modeling and Simulation, 2018.

19. Iftikharul Sami and KhurumNazirJunejo, "Predicting Future Gold Rates using Machine Learning Approach", International Journal of Advanced Computer Science and Applications, 2017.

20. NalinipravaTripathy, "Forecasting Gold Price with Auto Regressive Integrated Moving Average Model", International Journal of Economics and Financial Issues, 2017.

21. K. R SekarManav Srinivasan, K. S. Ravichandran and J. Sethuraman, "Gold Price Estimation Using A Multi Variable Model", International

Conference on Networks & Advances in Computational Technologies, 2017.

22. Sima P. Patil, Prof. V. M. Vasava, Prof. G. M. Poddar, " Gold Market Analyzer using Selection based Algorithm", International Journal of Advanced Engineering Research and Science, 2016.

23. S. Kumar Chandar, M. Sumathi and S. N. Sivanadam, "Forecasting Gold Prices Based on Extreme Learning Machine", International Journal of Computers Communications & Control, 2016. International Journal of Trend in Scientific Research and Development (IJTSRD) @ www.ijtsrd.com eISSN: 2456-6470 @ IJTSRD | Unique Paper ID – IJTSRD33143 | Volume – 4 | Issue – 5 | July-August 2020 Page 1456

24. NurulAsyikin Zainal and ZurianiMustaffa, "Developing A Gold Price Predictive Analysis Using Grey Wolf Optimizer", 2016 IEEE Student Conference on Research and Development, 2016.

25. Hossein Mombeini and AbdolrezaYazdani-Chamzini, "Modeling Gold Price via Artificial Neural Network", Journal of Economics, Business and Management, 2015.