



Daily Price Prediction of Gold and Silver in India: Statistical and Machine Learning Approach

Karanam Lakshmi

Assistant Professor, Department of Statistics, BBCIT, Kachiguda, Hyderabad, Telangana.
500027. karanam.lakshmi2@gmail.com

Dr. N. Konda Reddy

Associate Professor, Department of Engineering Mathematics, K L E F (Deemed to be University), Green Fields, Vaddeswaram, Guntur District, Andhra Pradesh.
kondareddymamatha@gmail.com

Abstract

The integration of multiple forecasting models has been widely recognized in both theoretical and empirical research as an effective approach for improving the predictive performance of individual models in time series analysis. This improvement becomes particularly significant when the combined models capture distinct characteristics of the underlying data. Hybrid forecasting models, which decompose a time series into its linear and nonlinear components, have emerged as one of the most popular and effective approaches for handling complex and volatile datasets. In this study, the daily prices of gold and silver are forecast using a hybrid modeling framework, artificial neural networks (ANN), and conventional Box-Jenkins time series models. To rigorously evaluate and compare the forecasting performance of these models, Friedman's test and the Morgan-Granger-Newbold (MGN) test are employed. The empirical results demonstrate that the hybrid model consistently outperforms both the traditional Box-Jenkins and feed-forward neural network (FFNN) models in terms of forecasting accuracy across different datasets.

Key Words: Forecasting, Time Series, Machine Learning, Error Metrics, Friedman's test, MGN test

Introduction

Middle-class Indian households feel that investing in gold and silver metals makes sense because they can't afford to buy real estate with just one salary, so they choose gold or silver instead. Gold jewellery and silver household equipment are highly favoured by the vast majority of Indians. Indians still place a great value on valuable metals like gold and silver. Among all the valuable metals, silver has been a popular choice for bullion investments and decorations among the general population due to its remarkable quality and rising affordability. Fair value transparency and reasonable liquidity are provided by a reputable market. These elements make this precious metal a wise speculative asset for investors.

Every household in India makes considerable use of gold and silver in their cultural and traditional rituals. The ability of the unique time arrangement models to predict is a search issue since the cost components are constantly changing. The goal of this study is to forecast metal prices using standard time series models, hybrid models, and artificial neural networks (ANN). Friedman's and MGN tests are used to assess the forecasting model's performance.

2. Methodology

2.1 Auto Regressive Integrate Moving Average

In statistics and econometrics, an autoregressive moving average (ARMA) model is generalized into an autoregressive integrated moving average (ARIMA) model for time series analysis. These models are fitted to time series data to either better comprehend the data or forecast future points in the series. They are employed when data show non-stationary behaviour; in these cases, the "integrated" part of the model corresponds to the initial differencing step that can be utilized to eradicate the non-stationary behaviour.

The non-negative integer parameters p , d , and q , respectively, specify the order of the model's autoregressive, integrated, and moving average components. The model is frequently called an ARIMA (p , d , q) model. An essential part of the Box-Jenkins time-series modelling method is the ARIMA model.

General formula of ARIMA (p , d , q)

$$w_t = \phi_1 w_{t-1} + \dots + \phi_p w_{t-p} + \delta + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$$

Where

$$w_t = \Delta^d y_t$$

The formula for the model may be shortened as follows:

$$\phi(B)\Delta^d y_t = \delta + \theta(B) \varepsilon_t$$

where

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \text{ and } \theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$$

2.2 Feed Forward Neural Network

Feed Forward Neural Networks (FFNN) are the most used Feed Forward Neural Network model for temporal order forecasting applications. While the input nodes show the previous lagged data, the output provides a forecast for the future values. Hidden nodes use appropriate nonlinear transfer functions to handle information received by input nodes.

The FFNN model may be expressed as

$$Z_t = \beta_0 + \sum_{j=1}^q \beta_j f \left(\sum_{i=1}^p \gamma_{ij} Z_{t-1} + \gamma_{0j} \right) + \varepsilon_t$$

Let f be a hyperbolic tangent function, and let q be the number of veiled cables and p be the number of information cables. Weights from the input to hidden nodes are represented as $\{\gamma_{ij}, i=0,1\dots p; j=1, 2, \dots, q\}$, whereas weights from the hidden to output nodes are vectorized as $\{\beta_j=1,2,\dots,q\}$.

The easiest way to conceptualize a multilayer network is as a series of cascading single-layer networks. The complexity of computers is demonstrated by the fact that this multilayer network is composed of multiple single-layer networks. The number of hidden layers in an artificial neural network should be considered by the network designer based on the required computing complexity.

2.3 Hybrid Methodology

It seems that ARIMA and ANN are not always useful for time series. This is because almost all real-world time series exhibit both linear and nonlinear correlation patterns between the observations. Zhang has recognized this important problem and developed a hybrid approach that uses ARIMA and ANN separately to describe the linear and nonlinear components of a time series. According to Zhang, we have

$$Y_t = L_t + N_t$$

where Y_t denotes the observation at time t , and L_t and N_t stand for the linear and nonlinear components, respectively. Initially, the linear component is fitted with ARIMA, yielding the equivalent forecast \hat{L}_t at time t . The residual at time t is thus given by $e_t = Y_t - \hat{L}_t$ gives the residual at time t . Zhang claims that an ANN can effectively predict the residuals dataset after fitting ARIMA since it only includes nonlinear components. The ANN for residuals takes the following structure with p input nodes:

$$e_t = f(e_{t-1}, e_{t-2}, \dots, e_{t-p}) + \varepsilon_t$$

where f is a nonlinear function computed by the ANN and ε_t is the white noise. If \hat{N}_t is the ANN's forecast, then the final hybrid forecast at time t may be found as follows:

$$\hat{Y}_t = \hat{L}_t + \hat{N}_t$$

Zhang discovered that his hybrid ARIMA-ANN approach significantly outperforms both ARIMA and ANN models in predicting accuracy through empirical analysis using three real-world time series.

3. Forecasting Accuracy

This section provides a quick overview of the prophetic performance tests that have been published in the literature. A Friedman's test and the Morgan-Granger-Newbold (MGN) test are used to compare the prediction abilities of two models. This section adopts the test methods to determine if the expectation intensity of two models corresponds.

3.1 Morgan-Granger-Newbold (MGN) Test

The null hypothesis for the MGN test is that the accuracy of the forecasted values is zero. The s_t and d_t correlation, i.e. $\rho = 0$. We employ the following test statistic to evaluate the hypothesis:

$$\text{MGN} = \frac{r_{sd}}{\sqrt{(1-r_{sd}^2)/(m-1)}}$$

This test statistic has $m-1$ degrees of freedom and follows a t -distribution, where r_{sd} is the sample correlation between s_t and d_t . If the forecasts are equally accurate, there will be no link between s_t and d_t . This test must be passed by errors that are devoid of sequential correlation and deviations from normal.

3.2 Friedman's Test

The performance's statistical significance is assessed using Friedman's test. Friedman's test for out-of-sample squared errors is used to compare the many forecasting models in an effort to ascertain whether the models' results differ significantly from one another. The following

statistic is examined under the null hypothesis, which asserts that every model performs equally well, using Friedman's test, which is based on the sum of the ranks R_j :

$$\chi^2 = \frac{12}{nk(k+1)} \sum_{j=1}^k R_j^2 - 3n(k+1)$$

is often distributed with a degree of freedom of $(k-1)$ where n is the number of observations in a model and k is the number of models.

4. Results

The historical prices of silver and gold were collected from January 1st 2022 to 30 July 2025 (1307 observations) from www.goldpricesindia.com and the same data is divided into training sample from 1st January 2022 to 30th June 2025 and testing sample from 1st July 2025 to 30th July 2025.

4.1. Forecasts of Daily prices of Silver in India

The best ARIMA model for silver price prediction was found using Expert Modeler, which looks for and estimates the best-fitting ARIMA for one or more dependent variable series automatically. Finding the ideal model through trial and error is eliminated by this method. The ARIMA (0, 1, 1) model fits the data rather well, according to the validation set tests.

The model that fits the data to predict India's daily silver prices is

$$\nabla^1 \ln Z_t = (1-0.153) + 0.00061 \Delta Y_{t-1}$$

The neural network model is constructed, with three covered neurons in the hidden layer being the optimal number. The best network is (1, 1, 1) since the selected network has the lowest

MAE, MAPE, and RMSE.

The hidden activation function is

$$H_{(1,1)} = \text{Tanh} (-0.066-0.276 z_{t-1})$$

Where z_{t-1} is the rescaled input variable and the forecasting model is

$$\hat{Z} = \mu_2 + \sigma_2 (0.241-3.791 H_{(1,1)})$$

The residuals from the linear model will only contain the nonlinearity relationship if the ARIMA model is applied to the linear component in the creation of the hybrid model. After applying ARIMA to the linear component values of (0, 1, 1) and obtaining the predicted values, the FFNN model is utilized to examine the residual values using SPSS. (1, 2, 1) is the ideal model.

The hidden activation function is

$$H_{(1,1)} = \text{Tanh} (0.488-0.133 e_{t-1})$$

Where \bar{e}_{t-1} is the rescaled input variable and forecasting model is

$$\bar{e}_{t-1} = \mu_2 + \sigma_2 (-0.244+0.779 H_{(1,1)})$$

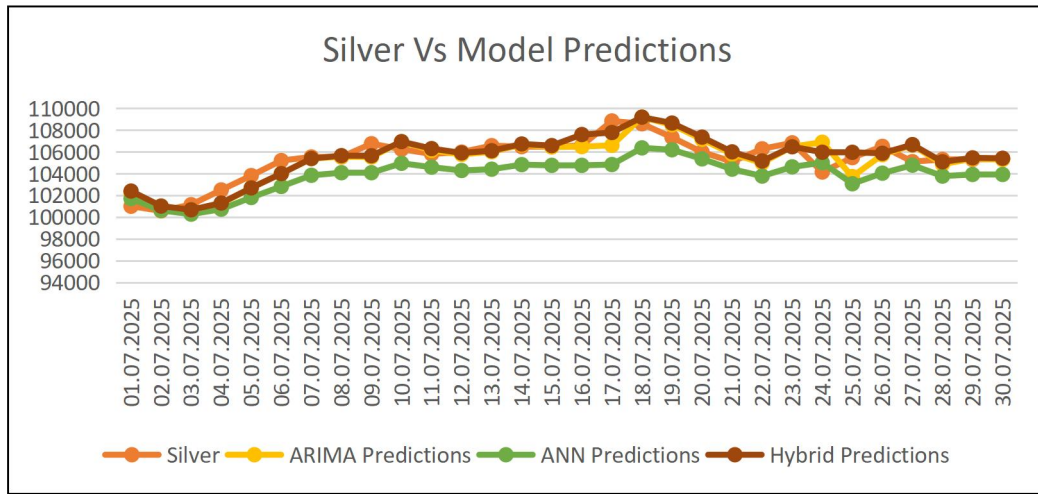
Hence $Z_t = \hat{L}_t + \hat{e}_t$

The following is a graphical representation of the expected daily silver prices in India using ARIMA, FFNN, and hybrid models.

Table 4.1.1. Forecasts using ARIMA, FFNN AND Hybrid Models for daily silver prices

Date	Silver	ARIMA Predicted	ANN Predicted	Hybrid Predicted
01.07.2025	100996	100810	101692	100805
02.07.2025	100573	100537	100611	100531
03.07.2025	101137	101229	100266	101226
04.07.2025	102492	102686	100725	102685
05.07.2025	103814	103988	101801	103987
06.07.2025	105195	105381	102816	105380
07.07.2025	105525	105547	103840	105547
08.07.2025	105525	105522	104079	105519
09.07.2025	106746	106934	104079	106931
10.07.2025	106236	106129	104945	106128
11.07.2025	105808	105759	104587	105754
12.07.2025	105974	106007	104283	106004
13.07.2025	106567	106653	104401	106651
14.07.2025	106474	106446	104820	106445
15.07.2025	106474	106478	104755	106475
16.07.2025	106578	106593	104755	106590
17.07.2025	108839	109185	104828	109182
18.07.2025	108577	108483	106361	108485
19.07.2025	107330	107152	106189	107148
20.07.2025	105976	105795	105349	105790
21.07.2025	105096	104988	104403	104983
22.07.2025	106275	106473	103768	106469
23.07.2025	106824	106878	104615	106878
24.07.2025	104142	103721	105000	103719
25.07.2025	105463	105731	103063	105723
26.07.2025	106499	106617	104035	106618
27.07.2025	105095	104861	104772	104859
28.07.2025	105300	105368	103767	105362
29.07.2025	105300	105368	103916	105366
30.07.2025	105353	105368	103916	105366

Figure 4.1.1 Forecasts of out sample using ARIMA, FFNN and HYBRID



4.1.2. Testing Equal Forecasting Accuracy with respect to Absolute Errors w.r.t to MGN test

The MGN test results are presented in this section to assess the forecasting models' equal prediction accuracy with respect to the absolute errors that happen outside of the sample. In order to ascertain whether the prediction accuracy of the two models is equivalent in terms of absolute errors, the MGN test is employed. The findings are shown in the corresponding table.

Table 4.1.2 Results of MGN test

Models	MGN Test Statistics	p Value	Remarks
ARIMA Vs FFNN	0.4836	0.6324	Insignificant
FFNN Vs HYBRID	1.8685	0.0722	Insignificant
ARIMA Vs HYBRID	0.9456	0.3524	Insignificant

Table 4.1.2 demonstrates that the ARIMA model's prediction performance is negligible in comparison to that of the FFNN model. The prediction performance of the HYBRID model differs significantly from that of the FFNN model. Similarly, there are notable differences between the FFNN and HYBRID models' prediction performance.

4.1.3 Testing Equal Forecasting Accuracy with respect to Squared Errors w.r.t MGN test

This section displays the outcomes of the MGN and Bootstrap tests, which evaluate the models' equal prediction accuracy in respect to out-of-sample squared errors. According to squared errors testing using the MGN test, the prediction accuracy of the two models is equal, as shown in the accompanying table.

Table 4.1.3 Results of MGN test

Models	MGN Test Statistics	p Value	Remarks
--------	---------------------	---------	---------

ARIMA Vs FFNN	-3.8947	0.0006	Significant
FFNN Vs HYBRID	5.4464	0	Significant
ARIMA Vs HYBRID	9.041	0	Significant

Table 4.1.3 shows that the forecast from the ARIMA model is much greater than the prediction from the FFNN model. Similar to the HYBRID model, the FFNN model is important. When used with a hybrid model, the accuracy forecast of the ARIMA model is negligible.

4.1.4 Testing Equal Forecasting Accuracy with respect to Absolute Errors w.r.t Friedman's test

The forecasting models' mean ranks in terms of squared errors are shown in the table below, along with the results of Friedman's test statistic.

Table 4.1.4 Friedman's Test

Sample	Model	Mean Rank	n	χ^2	p value
Out of Sample	ARIMA	1.867	30	85.84	0.002
	FFNN	1.7			
	HYBRID	1.77			

Table 4.1.4, which displays no appreciable difference in the models' performance, supports the acceptance of the null hypothesis. This is because the significant probability is greater than 0.05. However, according to the models' average rankings, the FFNN model is ranked top, followed by the HYBRID model in second and the ARIMA model in third. The FFNN model outperforms the ARIMA and HYBRID models in predicting silver prices in India according to Friedman's test.

4.2. Forecasts of Daily prices of Gold in India

Similarly, it has been determined that ARIMA (0, 1, 1) is significant in terms of the limits and model fit for daily gold prices in India. Given that this model has the lowest SBC values among the selected models, it is the best appropriate ARIMA model (0, 1, 1). The neural network model is created, with the ideal number of covered neurons being 2 in the hidden layer and the optimal network is 1-1-1 since it has the lowest MAE, MAPE, and RMSE. Similarly, by considering the ARIMA (0,1,1) to linear component values obtained the hybrid predicting values and model proposed for hybrid is (1, 1, 1).

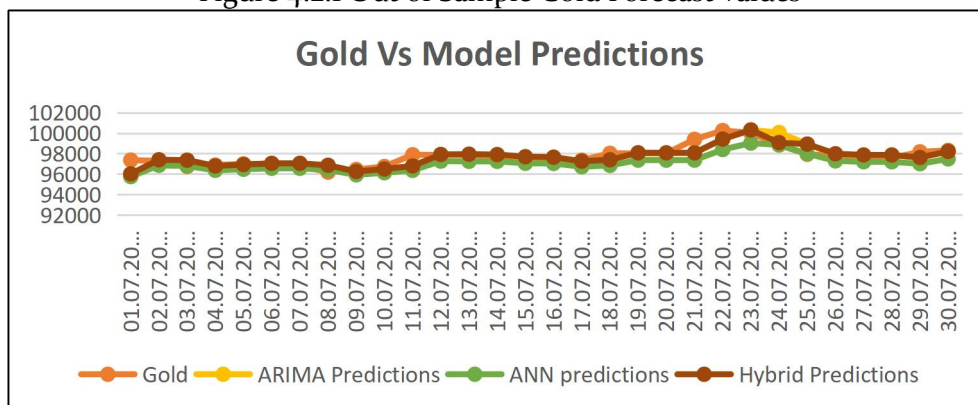
The following are the estimated values of the daily gold prices in India based on the FFNN, ARIMA and hybrid models.

Table 4.2.1 Gold Forecasts values by ARIMA, FFNN and Hybrid Models

Date	Gold	ARIMA Predictions	ANN predictions	Hybrid Predictions
01.07.2025	97360	95999	95745	96017

02.07.2025	97275	97374	96850	97393
03.07.2025	96737	97335	96784	97352
04.07.2025	96870	96773	96360	96789
05.07.2025	96988	96908	96465	96925
06.07.2025	96988	97025	96558	97042
07.07.2025	96794	97029	96558	97046
08.07.2025	96183	96836	96405	96853
09.07.2025	96445	96234	95918	96250
10.07.2025	96732	96469	96128	96486
11.07.2025	97884	96775	96356	96792
12.07.2025	97859	97893	97257	97912
13.07.2025	97859	97921	97237	97938
14.07.2025	97616	97881	97237	97898
15.07.2025	97600	97673	97050	97690
16.07.2025	97182	97626	97037	97643
17.07.2025	97352	97240	96711	97256
18.07.2025	98020	97373	96844	97390
19.07.2025	98015	98057	97361	98075
20.07.2025	98015	98055	97357	98072
21.07.2025	99388	98053	97357	98070
22.07.2025	100248	99396	98393	99415
23.07.2025	100014	100290	99022	100308
24.07.2025	98846	100056	98853	99073
25.07.2025	97922	98910	97989	98925
26.07.2025	97806	97964	97286	97979
27.07.2025	97806	97845	97197	97862
28.07.2025	97556	97845	97197	97862
29.07.2025	98174	97601	97003	97617
30.07.2025	98303	98194	97479	98212

Figure 4.2.1 Out of Sample Gold Forecast values



4.2.1. Testing Equal Forecasting Accuracy with respect to Absolute Errors of MGN test

The results of the MGN tests are shown in this section in order to assess if forecasting models have equivalent prediction accuracy in relation to the absolute mistakes that occur outside of the sample. The MGN test is used to determine if the two models' prediction accuracy is equal in terms of absolute errors, and the results are shown in the table that goes with it.

Table 4.2.2: Results of MGN test

Models	MGN Test Statistics	p Value	Remarks
ARIMA Vs FFNN	1.2802	0.1053	Insignificant
FFNN Vs HYBRID	1.8862	0.0846	Insignificant
ARIMA Vs HYBRID	2.7648	0.0048	Significant

Table 4.2.2 illustrates that the FFNN model's expected performance is also in line with the hybrid model's expectation, indicating that the ARIMA model's prediction performance is irrelevant when compared to both assumptions. However, the hybrid model's prediction performance and the ARIMA model's prediction performance diverge considerably. Compared to ARIMA and FFNN models, the hybrid model is hence more accurate.

4.2.3 Testing Equal Forecasting Accuracy with respect to Squared Errors

The results of the MGN tests, which assess the models' equal prediction accuracy in relation to out-of-sample squared errors, are shown in this section. The following table shows the equality of the two models' prediction accuracy as measured by squared errors using the MGN test.

Table 4.2.3: Results of MGN test

Models	MGN Test Statistics	p Value	Remarks
ARIMA Vs FFNN	0.2989	0.3835	Insignificant
FFNN Vs HYBRID	1.8862	1.1E-06	Significant
ARIMA Vs HYBRID	2.7648	1.74E-05	Significant

Table 4.2.3 reveals that there is a considerable difference in the ARIMA model's prediction accuracy compared to the FFNN model's prediction accuracy. When compared to the hybrid model's prediction precision, the FFNN model's prediction accuracy is negligible. Based on empirical data, the prediction accuracy of the hybrid model is assumed to be higher than that of the ARIMA and FFNN models, with a considerable difference in accuracy between the two models.

4.2.4 Testing of Equality of Forecasting Accuracy of Performance of the Models

The average rankings of the forecasting models in relation to squared errors are shown in the following tables, and the results of Friedman's test statistic are provided in the table that follows.

Table 4.2.4: Results of Friedman's test

Sample	Model	Mean Rank	n	χ^2	p value
Out of Sample	ARIMA	2.63	30	20.067	0.00
	FFNN	2.63			
	HYBRID	1.87			

Table 4.2.4. indicates that there is a significant difference in the models' performance, and as the significant probability is smaller than 0.05, the null hypothesis is rejected. It is noted that the Hybrid model receives first rank and that the ARIMA and FFNN models share second rank equally in light of the models' mean positions. As a result, the hybrid model outperforms the FFNN and ARIMA models in terms of predicting gold prices in India.

5. Comparison of the Models with Error to Measures

The forecasts from the FFNN, Hybrid, and ARIMA models for both in- and out-of-sample periods are summarized here. The models are compared using MAE, MAPE, and RMSE, as shown in the table.

Table 5.1. Performance of the Models with respect to error for Silver Prices

Model	Performance Metric					
	In Sample			Out Sample		
	MAE	MAPE	RMSE	MAE	MAPE	RMSE
ARIMA	574.31	0.758	790.80	799.80	0.76	1044.66
ANN	579.3985	0.8054	786.4863	693.4981	0.7742	986.4863
HYBRID	565.454	0.7521	764.125	731.434	0.695	886.813

Table 5.1 compares the forecasting performance of the ARIMA, ANN and Hybrid models for silver prices using MAE, MAPE and RMSE. In the in-sample period, the Hybrid model shows the lowest error values, indicating the best model fit.

For the out-of-sample period, the ANN model records a slightly lower MAE, while the Hybrid model achieves the lowest MAPE and RMSE. Overall, the results suggest that the Hybrid model provides more reliable and stable forecasts, whereas the ANN model performs

marginally better in terms of absolute error. Both models outperform the ARIMA model in forecasting daily silver prices.

Table 5.2 Performance of the Models with respect to error for Gold Prices

Model	Performance Metrics					
	In Sample			Out Sample		
	MAE	MAPE	RMSE	MAE	MAPE	RMSE
ARIMA	302.125	0.456	491.684	409.756	0.418	584.526
ANN	496.814	0.782	641.436	684.223	0.697	853.124
HYBRID	286.935	0.425	451.175	379.745	0.387	540.14

Table 5.2 shows that the Hybrid model achieves the lowest error values in both the in-sample and out-of-sample data, indicating superior forecasting performance of gold prices. The ANN model records the highest errors, while the ARIMA model performs moderately. Overall, the Hybrid model outperforms both the neural network and the Box-Jenkins(ARIMA) approach in forecasting daily gold prices in India.

6. Conclusion Based on Prediction Accuracy tests

Friedman’s test indicates significant differences in the predictive accuracy of the models for the daily silver prices in India. The MGN test further confirms that both absolute and squared errors differ across the models. While ARIMA and FFNN exhibit comparable performance in terms of absolute errors, significant differences are observed between the Hybrid and ARIMA models show similar accuracy, whereas both differ from the FFNN model.

For daily gold price forecasting, Friedman’s test shows that the Hybrid model outperforms both the ARIMA and FFNN models, while the latter two perform similarly. According to the MGN test, absolute errors across the three models are comparable; however, the Hybrid model demonstrates superior overall performance. Similar conclusions are obtained when squared errors are considered.

Overall, the error metrics (MAE, MAPE, RMSE) clearly highlight the advantage of the hybrid approach for forecasting daily silver and gold prices. These results suggest that combining linear and nonlinear models significantly reduces forecasting errors compared to using individual models alone.

References

1. Amjad A. Alsuwaylimi ,2023, Comparison of ARIMA, ANN and Hybrid ARIMA - ANN
2. Models for Time series Forecasting, An International Journal of Information Sciences
3. Letters,12, No. 2,1003-1016.
4. Becher, H., Hall, P. and Wilson, S. R. (1993). Bootstrap hypothesis testing procedures, Biometrics. Vol. 9(4). 1268-1272.
5. Bing Hong Uh & Noriza Majid, (2021), Comparison Of ARIMA Model and Artificial Neutral Network in Forecasting Gold Price, 17(2) 31-39
6. Chaovalit, A. Gangopadhyay, G. Karabatis, Z. Chen, (2011) Discrete wavelet transform Based time series analysis and mining, ACM Computing Surveys, 43(2)



7. Efron, B. (1979). Bootstrap methods: Another look at the jackknife, *Annals of Statistics*, Vol. 7. 1-26
8. Hornik, K. (1993), Some new results on neural network approximation, *Neural Networks*, Vol. 6, 1069-1072.
9. Khandelwal Ina, Adhikari A, Ghanshyam Verma, (2015), Time Series Forecasting Using Hybrid ARIMA and ANN Models Based on DWT Decomposition, *Procedia Computer Science*, Volume 48, 173-179
10. Mohammad Alavi, Mohammad Albaji, Mona Golabi, Abd Ali Nasei (2024), Estimation of
11. sugarcane evapotranspiration from remote sensing and limited meteorological variables
12. using machine learning models, *Journal of hydrology*, Volume 629, 130605
13. Murali Krishna, M., Raghavender Sharma, M., Konda Reddy. N, 2021 Accuracy of
14. forecasting
15. models using Bootstrap test and Friedman's test, *AIP Conference Proceedings*, Volume
16. 2375, (1):020005.
17. Naveen Kumar, B., Ramu, Y., Venugopala Rao, M. and Krishna Reddy, M. (2011a). A
18. bootstrap
19. test for equality of mean absolute errors. *ARPN Journal of engineering and applied*
20. *sciences*, Vol.
21. 6(5). 153-161.
22. Ramu, Y., Naveen Kumar, B and Krishna Reddy, M. (2011). An approximation to the cdf of
23. standard Normal distribution, *International Journal of Mathematical Archive*, Vol. 2(7).
24. 1077-
25. 1079.
26. Zhang, G. and Hu, M. Y. (1998). Neural Network forecasting of the British pound/US dollar
27. exchange rate, *Journal of Management science*. Vol. 26(4). 495-506. 020005-18
28. Zhang, G., Patuwa, E.B., and Hu, M.Y. (1998). Forecasting with artificial neural networks;
29. The
30. state of the art. *International Journal of Forecasting*, Vol. 14. 35-62.
31. Zhang, G. P. (2001). An investigation of Neural networks for Linear Time Series
32. Forecasting,
33. computers and operations Research. Vol. 28. 1183-120