VIT®

Vellore Institute of Technology

(Deemed to be University under section 3 of UGC Act, 1956)

School of Computer Science and Engineering

J Component report

Programme : **B.Tech**

Course Title : **Essentials of Data Analytics**

Course Code : **CSE3506**

Slot : **G1/G2**

Title: **GOLD PRICE PREDICTION USING TIME SERIES MODELS**

Team Members: **Surya S | 20BCE1071**

RS Jyothish | 20BCE1040

Siddharth Bhuthapuri | 20BCE1465

Faculty: Dr. Rajalakshmi R **Sign:**

Date: 08/04/2023

ACKNOWLEDGEMENT

We wish to express our sincere thanks and deep sense of gratitude to our project guide, **Dr. R. Rajalakshmi,** School of Computer Science and Engineering for her consistent encouragement and valuable guidance offered to us throughout the course of the project work.

We are extremely grateful to **Dr. R. Ganesan, Dean,** School of Computer Science and Engineering (SCOPE), Vellore Institute of Technology, Chennai, for extending the facilities of the School towards our project and for his unstinting support.

We express our thanks to our **Head of the Department** for his support throughout the course of this project.

We also take this opportunity to thank all the faculty of the School for their support and their wisdom imparted to us throughout the courses.

We thank our parents, family, and friends for bearing with us throughout the course of our project and for the opportunity they provided us in undergoing this course in such a prestigious institution.

BONAFIDE CERTIFICATE

Certified that this project report entitled "**Gold Price Prediction"** is a bona-fide work of **RS Jyothish (20BCE1040), Surya S (20BCE1071) Siddharth Bhuthapuri (20BCE1465)** carried out the "Gold Price Prediction"-Project work under my supervision and guidance for **CSE3506 – Essentials of Data Analytics**.

Dr. R. Rajalakshmi

SCOPE

ABSTRACT

This work aims to predict the gold prices using time series models. The analysis is based on the historical data of the gold prices from December 1978 to December 2022. This work considers 5 time series models to forecast the future prices of gold. The performance of the models that forecast gold price for the next 15 months is evaluated based on various visualization techniques. The results indicate that ARIMA, SARIMA and LSTM models perform well for the data collected, with the ARIMA model performing slightly better than the other models. The study concludes that time series models can provide useful insights into the future trends of the gold prices, which can be valuable for investors and policymakers in making informed decisions.

TABLE OF CONTENTS

INTRODUCTION

Gold price prediction using time series models is an area of analysis in finance and economics. Time series models are statistical models that are used to analyze and forecast time series data, which is a sequence of data points collected at regular intervals over time. In this paper where gold price in the future is predicted, time series models can be used to analyze historical gold prices and predict future gold prices.

This work includes various time series models for gold price prediction: autoregressive (AR), autoregressive integrated moving average (ARIMA), seasonal ARIMA (SARIMA), exponential smoothing (ETS), and long short-term memory (LSTM) models. These models are used to analyze the trends, seasonality, and other patterns in historical gold prices and make predictions about gold prices for the next 15 months.

One of the main challenges in gold price prediction using time series models is the volatility and unpredictability of the gold market. Gold prices are influenced by a variety of factors, including global economic conditions, geopolitical events, inflation, and supply and demand. Therefore, it can be difficult to accurately predict gold prices over long periods of time.

Despite these challenges, time series models are widely used for gold price prediction by financial analysts, investors, and traders. These models can provide valuable insights into the trends and patterns of gold prices and help stakeholders make informed decisions about buying, selling, or holding gold.

RELATED WORKS

[1] It is a review of literature in the field of using machine learning techniques to predict future gold prices. The authors provide an overview of the various machine learning algorithms used for this purpose and their results, and then present their own approach and findings. The machine learning algorithm used in this paper is artificial neural network(ANN) and have used ARMA model for time series analysis. The authors conclude that machine learning methods have shown to be effective in predicting future gold prices, but further research is needed to improve the accuracy of these predictions.

[2]This paper is a study on predicting gold prices using machine learning techniques. The authors compare the performance of different machine learning algorithms including multiple linear regression, decision tree, and artificial neural networks. They evaluate the accuracy of these models on historical gold price data. The results show that artificial neural networks provide the best performance in terms of accuracy and the authors conclude that machine learning techniques can effectively be used for predicting gold prices.

[3]This is a paper in the field of using machine learning algorithms to predict gold prices. The authors present a new approach to forecasting gold prices using the XGBoost algorithm and SHAP interaction values, which are measures of the impact of each feature on the prediction. The authors evaluate the performance of their model on historical gold price data and compare the results with other machine learning algorithms. The results show that their approach provides improved accuracy compared to other methods and is a promising tool for predicting gold prices

[4]This presents a study on predicting onion prices using machine learning techniques. The authors compare the performance of different machine learning algorithms including KNN, Naive Bayes, Decision tree, SVM, and Neural Network. They evaluate the accuracy of these models on historical onion price data and compare the results with traditional time-series forecasting methods. The results show that the artificial neural network model outperforms the other methods and provides a better prediction of onion prices.

[5]This paper presents a study on predicting onion price volatility in Indian wholesale markets using hybrid deep learning models. The authors use a combination of Generalized Autoregressive Conditional Heteroscedasticity (GARCH) and Long Short-Term Memory (LSTM) models to make predictions. They evaluate the performance of the model on historical onion price data and compare it with traditional time-series forecasting methods. The results show that the hybrid GARCH-LSTM deep learning model outperforms the other methods in terms of accuracy and provides a better prediction of onion price volatility.

[6] This paper proposes recurrent neural networks-based on long short-term memory (LSTM) and gated recurrent unit (GRU) networks. The accuracy of the proposed techniques is measured using Mean absolute percentage error (MAPE). Using this, it has been determined that the LSTM and GRU networks are very useful and successful in predicting the nickel price owing to having average MAPE values of 7.060% and 6.986%, respectively.

[7] The original gold prices are decomposed into sub-layers with different frequencies by the improved complete ensemble empirical mode decomposition with adaptive noise (ICEEMDAN). The long short-term memory, convolutional neural networks,

and convolutional block attention module (LSTM-CNN-CBAM) joint forecasting all sub-layers from the previous step. The prediction of the sub-layers with different models is reconstructed as the final predicted results with the summation method. Among them, the proposed model could extract the long-term effect of the gold price by LSTM, mining the deep features of gold price data with CNN, and improving the feature extraction ability of the network through CBAM. It is proved by experiment that the cooperation among LSTM, CNN and CBAM can strengthen the modeling ability and improve the prediction accuracy.

[8]This paper seeks to forecast of gold volatility by combining two deep learning methodologies: short-term memory networks (LSTM) added to convolutional neural networks (CNN). The CNN-LSTM hybrid model is capable of including images as input which provides a wide variety of information associated with both static and dynamic characteristics of the series. In parallel, different lags of profitability of the series are entered as input, which allows it to learn from the temporal structure. The results show a substantial improvement when this hybrid model is compared to the GARCH and LSTM models. A 37% reduction in MSE is observed compared to the classic GARCH model, and 18% compared to the LSTM model. Finally, the Model Confidence Model (MCS) determines a significant improvement in the prediction of the hybrid model.

[9]This paper proposes a model based on short and long-run decomposition of input variables using the Ensemble Empirical Mode Decomposition algorithm and forecasting each component separately based on the Support Vector Regression method. The empirical

findings suggest that the proposed method significantly increases the economic returns of a trading strategy based on the forecasts of the proposed scheme.

[10]This paper puts forth a novel architecture based on long short-term memory model optimized by genetic algorithm (GA-LSTM) and error correction strategy for multi-step-ahead copper price forecasting. The proposed architecture includes the following two sub-phases: (1) initial forecasting of copper price; (2) error correction. In the first phase, to improve the forecast accuracy, a number of recent copper price data and some selected influencing factors are combined to construct a hybrid input of the GA-LSTM model. The hybrid input strategy simultaneously considers the historical price data and the causal relationship between the influencing factors. One real copper price data series is used for validating the forecasting ability of the proposed architecture and a same length of iron ore price data series is also used to test the generalization ability and robustness of the proposed architecture. The experiment results illustrate that the proposed architecture outperforms the benchmark models considered in this study.

[11]In this study, the closing prices of five companies from various industry sectors were predicted using Artificial Neural Network and Random Forest techniques. The stock's open, high, low, and close prices are used to create new variables that are used as model inputs. RMSE and MAPE, two common strategic indicators, are used to evaluate the models. The models are effective at forecasting stock closing price, as evidenced by the low values of these two indicators.

[12]This study examines housing data from 5359 townhouses in Fairfax County, Virginia, obtained by the Multiple Listing Service (MLS) of the Metropolitan Regional Information Systems to increase the precision of housing price prediction (MRIS). A housing price prediction model is created, and its classification accuracy performance is compared, using machine learning algorithms like C4.5, RIPPER, Naive Bayesian, and AdaBoost. A more accurate housing price prediction model is created to help home sellers and real estate agents make more informed choices based on the valuation of a home. The experiments show that the RIPPER algorithm consistently outperforms the other models in the performance of housing price prediction, measured in terms of accuracy.

[13]Predicting a real estate property's market value is the paper's main goal. This system aids in determining a property's starting price based on geographic factors. Future costs will be predicted by dissecting past market trends, price ranges, and upcoming technological advancements. With the help of a decision tree regressor, this investigation aims to forecast house prices in Mumbai. Clients will benefit from being able to invest money in a bequest without turning to a broker. The findings of this study demonstrated that the Decision Tree Regressor has an 89% accuracy rate.

[14]For comparison and analysis in this survey, papers were chosen from the Digital Bibliography & Library Project (DBLP) database. Convolutional neural network (CNN), Long Short-Term Memory (LSTM), Deep neural network (DNN), Recurrent neural network (RNN), Reinforcement Learning, and other deep learning techniques like Hybrid Attention Networks (HAN), self-paced learning mechanism (NLP), and Wavenet were among the

techniques we used to classify papers. The paper also examines each article's dataset, variable, model, and findings. The survey in use presents the results using the most popular performance metrics, including accuracy, Sharpe ratio, mean square error (MSE), mean absolute percent error (MAPE), and return rate.It is concluded that in recent years, the trend of using deep-learning-based methods for financial modeling is rising exponentially.

[15]Establishing a model that can forecast the near future in relation to the variable collection of features is the aim of this study.

This study is based on a preliminary examination of features that can improve forecasting models' resistance to market swings, such as COVID-19 cases and other market tickers. The predictions are based on market value data and daily time-series data from the COVID-19 epidemic (i.e. the number of new cases).

The dataset has been processed using convolutional neural networks (CNN) LSTM, vector sequence output LSTM, bidirectional LSTM, and encoder-decoder LSTM. The vector sequence output LSTM's findings have outperformed other methods that have been suggested in the literature.

PROPOSED METHODOLOGY

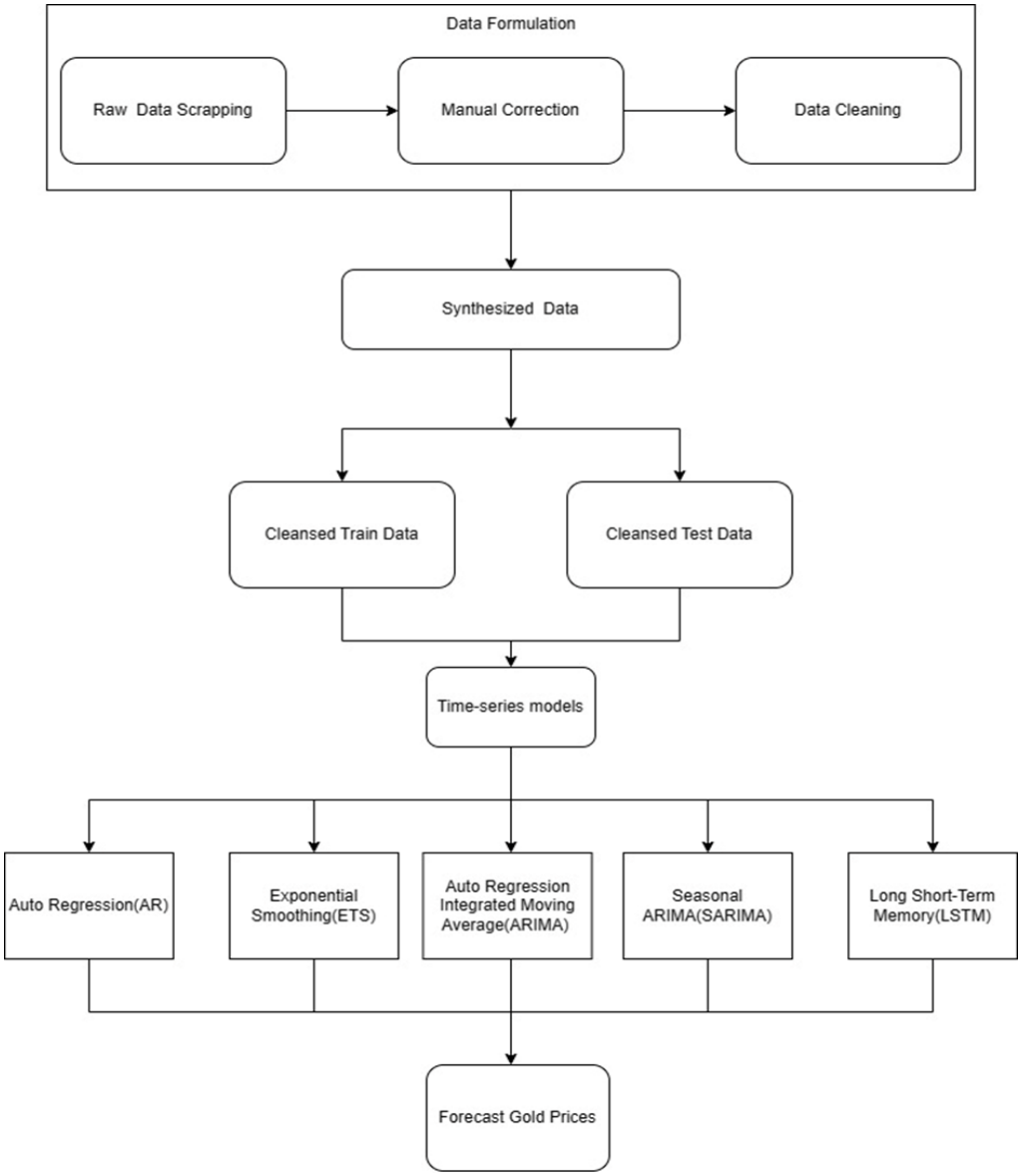

Figure 1 Proposed Framework for Gold Price Forecasting

The proposed framework as mentioned in Figure 1 is implemented in this paper considers different time-series models to predict future gold prices and conclude which model gives the best results. The time-series models implemented in this paper are- Auto Regression (AR), Exponential Smoothing (ETS) , Autoregressive Integrated Moving Average (ARIMA) , Seasonal ARIMA (SARIMA), and Long short-term memory(LSTM). This paper includes visualization with different graphs in order to understand trends in Gold prices and to forecast Gold Prices for the upcoming 15 months.

RESULTS & DISCUSSION

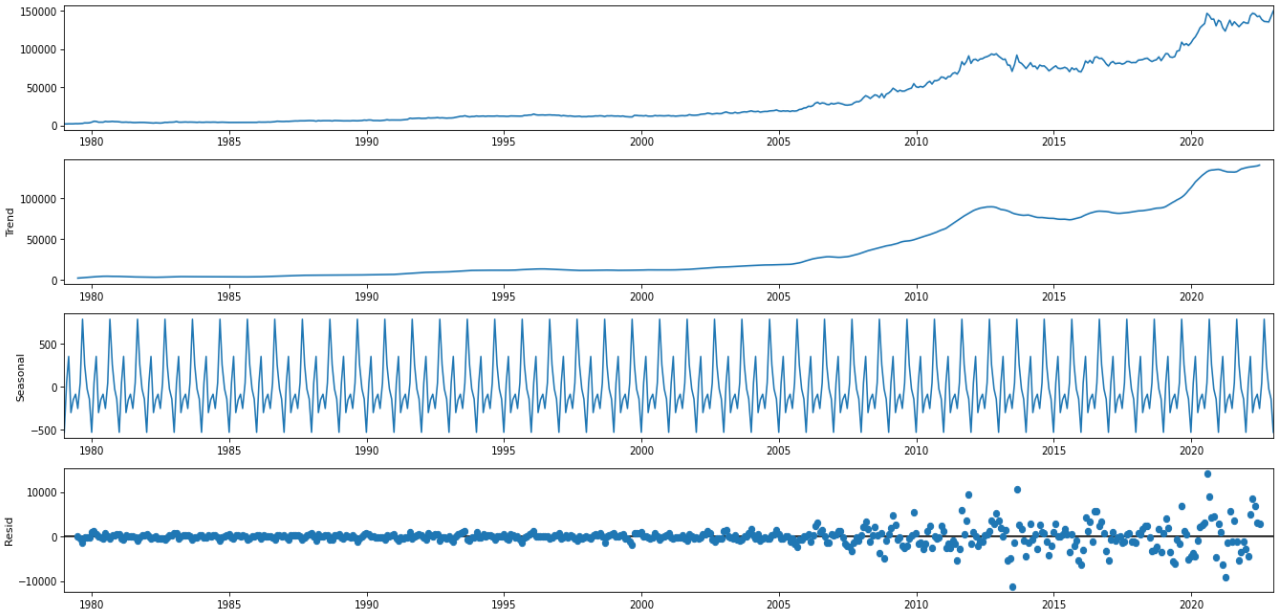

Figure 2 Decomposition Graph for Gold Prices in India

From Figure-2, decomposition graph can analysed as follows:

Trend: Overall,the Gold Price has increased over the years in India. In the period 1978-2005, the rate of increase has been minimal. After 2005, the rate of increase of prices has been steep.

Seasonality: On observing the pattern, it can be concluded that price has been least during the start of the year. During the first quarter, after falling to a minimum value, price increases. In contrast, during the second quarter, it is seen that price decreases. In this quarter it reaches the maximum value with increase in Gold prices and again experiences a decrease in prices in the fourth quarter.

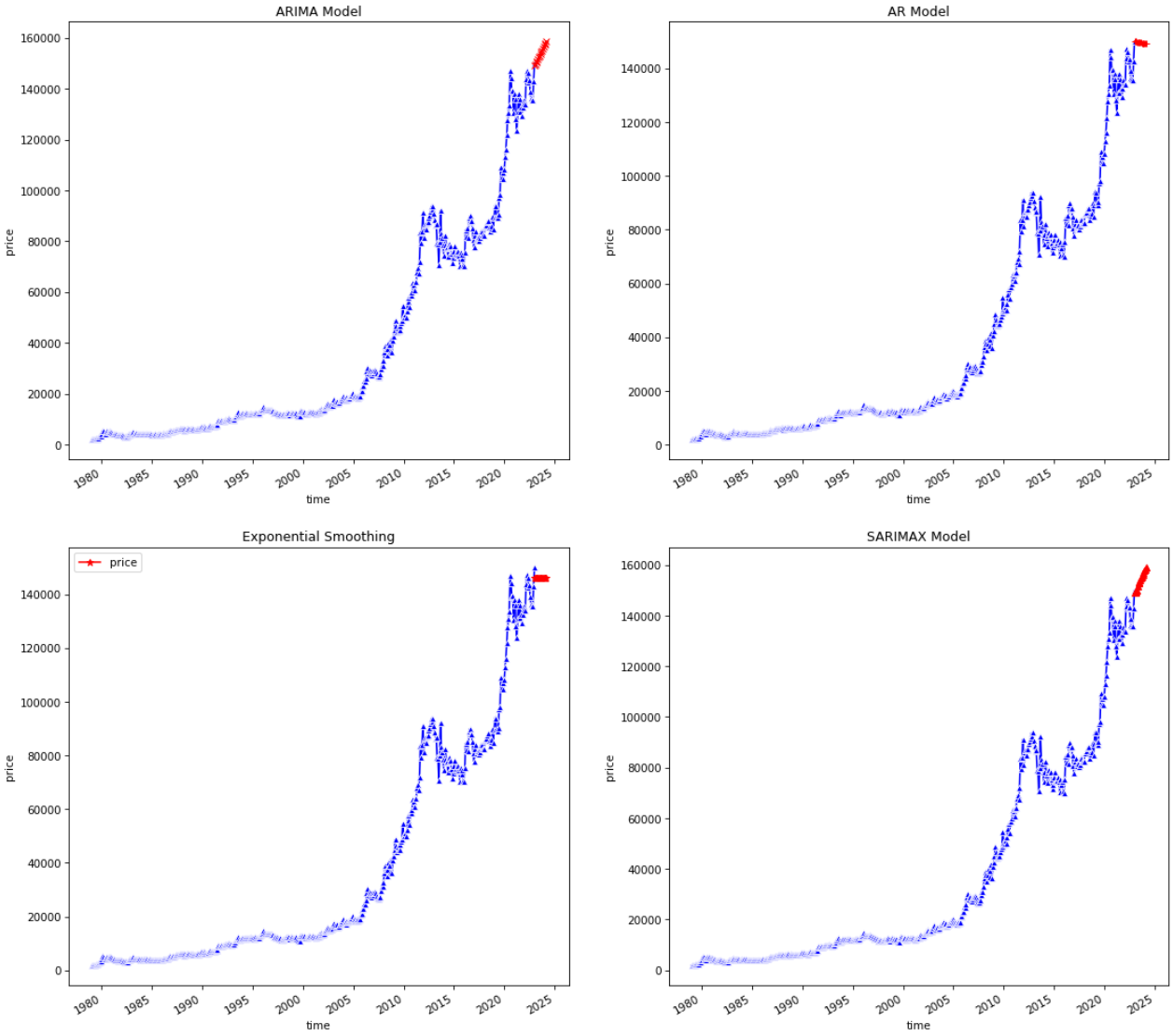

Figure 3 Comparative visualization: Forecasting Gold Prices

In the Figure-3, four time-series models are plotted with their forecasted prices in red colour. The ARIMA and the SARIMAX model show an increase in Gold prices in India with the SARIMAX model showing a slightly steeper rate of increase. The Auto-Regression (AR) model forecasts the Gold price to decrease in the upcoming 15 months. The Exponential smoothing model forecasts the price to be constant.

The ARIMA and SARIMAX provide a better forecast than the other models.

Figure 4 Predicted v/s Actual using LSTM model

Figure-4 shows the actual value v/s predicted value from the LSTM model. Trend and seasonality for the predicted is the same as the actual value whereas the predicted value is slightly lower than the actual value.

CONCLUSION

This work is a forecasting of Gold prices in India. The results of this work indicates that Gold price has been increasing steeply from 2005 and is forecasted to increase in the next 15 months in India. From the five used time-series models ARIMA and SARIMAX have yielded good results for our dataset. Auto-regression and Exponential Smoothing models haven't been up to the mark in this work. For getting a better understanding of the dataset, decomposition graph and time series plot have been visualized for the monthly Gold prices in India. This methodology can be used to forecast Gold prices in any other country.

Along with the time series data, other factors that affect the price of Gold such as Inflation rate, Geo-political factors, Rupee-dollar impact, etc. can be integrated into this methodology to provide a more accurate forecaster.

References

[1]ul Sami, I., & Junejo, K. N. (2017). Predicting future gold rates using machine learning approach. *International Journal of Advanced Computer Science and Applications*, *8*(12).

[2]Ghule, Rushikesh & Gadhave, Abhijeet & Dubey, Manish & Kharade, Jyoti & Mumbai, Navi. (2022). Gold Price Prediction using Machine Learning. International Journal of Environmental Engineering. 6. 10.55041/IJSREM15027.

[3]Jabeur, S. B., Mefteh-Wali, S., & Viviani, J. L. (2021). Forecasting gold price with the XGBoost algorithm and SHAP interaction values. Annals of Operations Research, 1-21.

[4]Hasan, Md. Mehedi & Hafiz, Rubaiya & Tuz, Muslima & Saifuzzaman, Mohd & Sykot, Md. (2020). Solving Onion Market Instability by Forecasting Onion Price Using Machine Learning Approach. 10.1109/ComPE49325.2020.9200033.

[5]Swarup, S., & Kushwaha, G. S. (2022). An approach for forecasting the Onion Price volatility in Indian Wholesale Markets using hybrid GARCH-LSTM Deep learning models.

[6]A. C. Ozdemir, K. Buluş, and K. Zor, "Medium- to long-term nickel price forecasting using LSTM and GRU networks," *Resources Policy*, vol. 78, p. 102906, Sep. 2022, doi: 10.1016/j.resourpol.2022.102906.

[7]Y. Liang, Y. Lin, and Q. Lu, "Forecasting gold price using a novel hybrid model with ICEEMDAN and LSTM-CNN-CBAM," *Expert Systems with Applications*, vol. 206, p. 117847, Nov. 2022, doi: 10.1016/j.eswa.2022.117847.

[8]A. Vidal and W. Kristjanpoller, "Gold volatility prediction using a CNN-LSTM approach," *Expert Systems with Applications*, vol. 157, p. 113481, Nov. 2020, doi: 10.1016/j.eswa.2020.113481.

[9]V. Plakandaras and Q. Ji, "Intrinsic decompositions in gold forecasting," *Journal of Commodity Markets*, p. 100245, Jan. 2022, doi: 10.1016/j.jcomm.2022.100245.

[10]H. Luo, D. Wang, J. Cheng, and Q. Wu, "Multi-step-ahead copper price forecasting using a two-phase architecture based on an improved LSTM with novel input strategy and error correction," *Resources Policy*, vol. 79, p. 102962, Dec. 2022, doi: 10.1016/j.resourpol.2022.102962.

[11]Stock Closing Price Prediction using Machine Learning Techniques

Mehar Vijh , Deeksha Chandola , Vinay Anand Tikkiwal , Arun Kumar

[12]Using machine learning algorithms for housing price prediction: The case of Fairfax County, Virginia housing data

Byeonghwa Park , Jae Kwon Bae

[13]House Price Forecasting Using Machine Learning

Alisha Kuvalekar,Shivani Manchewar,Sidhika Mahadik,Shila Jawale

[14]A Survey of Forex and Stock Price Prediction Using Deep Learning

Zexin Hu ,Yiqi Zhao and Matloob Khushi

[15]A Deep Learning-Based Method for Forecasting Gold Price with Respect to Pandemics

Mahtab Mohtasham Khani, Sahand Vahidnia & Alireza Abbasi