2022 5th International Conference on Contemporary Computing and Informatics (IC3I)

Commodity Price Prediction for making informed Decisions while trading using Long Short-Term Memory (LSTM) Algorithm

Shashikant Suman
Dept. of Computer Science
Amity School of Engineering and Technology
Greater Noida, India
sshashikant18@outlook.com

Prakarsh Kaushik
Dept. of Computer Science
Amity School of Engineering and Technology
Greater Noida, India
prakarsh.kr@gmail.com

Sai Sri Nandan Challapalli
Dept. of Computer Science
Amity School of Engineering and Technology
Greater Noida, India
ch.saisrinandan854@gmail.com

Bhanu Prakash Lohani
Dept. of Computer Science
Amity School of Engineering and Technology
Greater Noida, India
bhanuplohani@gmail.com

Pradeep Kushwaha
Dept. of Computer Science
Amity School of Engineering and Technology
Greater Noida, India
pkap1984@gmail.com

Amar Deep Gupta
Dept. of Computer Science
Amity School of Engineering and Technology
Greater Noida, India
amardeep.ip@gmail.com

2022 5th International Conference on Contemporary Computing and Informatics (IC3I) | 979-8-3503-9826-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/IC3I56241.2022.10072626

***Abstract*— Commodity markets are physical or virtual marketplaces where market players meet to buy or sell positions in commodities such as crude oil, gold, copper, silver, cotton, and wheat. People invest their hard-earned money based on some predictions to gain some profit from commodity market. Although, traditional methods such as technical analysis & fundamental analysis are very popular among traders, they are not as accurate as analysis by long short-term memory (LSTM) algorithm. In this paper, we have developed a model of well-known efficient LSTM algorithm to predict the commodity market price by utilizing a freely accessible dataset for commodity markets having open, high, low, and closing prices from historical data.**

Keywords—Commodity, LSTM, Future Price, Deep Learning, RNN, Commodity Price Prediction

I. INTRODUCTION

A commodity is a basic good that can be exchanged with some other items similar type in trade. Commodities are generally used as inputs in manufacturing of other commodities or services. Thus, a commodity is typically defined as a raw resource used in the production of consumer items [1]. They could also be fundamental necessities like some of the agricultural items. The interesting component of a commodity is that there is little, if any, difference between it originating from one source and the same item coming from another. Regardless of the producer, a bag of rice is essentially the same product. The same can be said for a crude oil barrels or a tonne of ore. In contrast, the quality and characteristics of a given consumer product will frequently fluctuate significantly based on the manufacturer. One example of this could be Fanta & Mirinda. Both of them are fruit flavored carbonated drinks, have used the same material to products, the only difference is in the composition of these materials and the manufacturing company [2]. Commodities include wheat, silver, pork, petroleum, and natural gas. Recently, the concept has been broadened to add financial instruments such as foreign currencies and indices. Technology have also resulted in the trading of new commodities in the commodity market. For example, cell phone minutes and bandwidth. Commodities can be traded and purchased as financial assets on some marketplaces. There are also well-developed derivatives markets where contracts on such commodities can be purchased (e.g., forwards, futures, and options). Commodity futures are traded by two categories of traders.

The first are commodity buyers and sellers who use commodity futures contracts for the safeguarding reasons that they were designed for. Once the futures contract expires, these traders make or receive delivery of the actual commodity. The speculator is the second category of commodities trader. These are traders who operate in the commodities markets solely to profit from dramatic price changes. When the futures contract expires, these traders have no intention of making or taking delivery of the physical commodity [20].

A. Types of Commodities

Based on the commodity, the market can be categorized into two types – Hard Commodity and Soft Commodity.

1) Hard Commodity - Hard commodities are those that are required by the manufacturing industry. These need to be mined and extracted manually from the earth or the sea. They have finite reserves and are particularly vulnerable to economic and geopolitical events. Oil, Gold, silver, rubber, copper, and other commodities are examples of hard commodities. The extraction process accounts for the majority of the pricing.

2) Soft Commodity - Soft commodities are primarily agricultural or livestock goods. In contrast to hard commodities, they are produced rather than extracted or mined. They have practically infinite supplies and are influenced by weather or natural events rather than geopolitical circumstances. Corn, wheat, barley, sugar, coffee, tea, and other such commodities are examples.

B. Working of Commodity Market and Trading

The commodities market is essentially governed by the principles of demand and supply. When demand equals supply, the market has attained equilibrium [20]. The process of trading commodities is divided into four parts. The commodity market commences with commodity production. This is referred to as primary production. Cultivators, animal rearers, miners, and other primary producers bring their products to markets to trade. The following stage is the transformation of raw materials into finished goods, such as cotton into yarn or cloth, wheat into flour, or mustard into oil. This is referred to as secondary production. The next stage entails traders, wholesalers, and retailers selling final items to consumers. This is known as the distribution commerce stage. The commodities market in India comes to an end at this stage, which is defined as the consumption or use of products and services by individuals and institutions for their personal purposes or for use in further processing or manufacturing [21].

For Commodity Trading, the commodities exchange provides data on the current bid and offer rates of the commodity being sold. The dealers who post such bids and offers provide these details [22]. The commodity market in general is divided into three major segments, which are as follows:

1) Commodity exchanges serve as a meeting place for sellers and buyers. These exchanges will keep a list of commodities that they will add to on a regular basis based on demand and supply patterns. These commodities can be traded on the exchange, at your broker's office, or online from the comfort of your own home.

2) Brokers are also active commodity market participants. They handle all transactions between buyers and sellers at the risk of their money under a contract with their clients.

3) Forward contracts are also used to trade commodities between farmers and exporters/importers who want to hedge against price volatility.

C. Advantages of Commodity Trading

The commodity market is older and more sophisticated than the financial stock market. Barter trading was the first form of trading known to humankind, in which goods such as food grains were swapped between farmers and customers [11]. There are many advantages of commodity trading which includes:

1) Immunity against Inflation - Stock prices fall during inflationary times. In contrast, as demand grows, the prices of commodities used in the production of finished goods rise significantly, resulting in higher final-goods costs. As a result, investors flock to commodity futures to protect their capital from the consequences of inflation and preserve its value.

2) Hedge against geopolitical conflicts - When there is unrest, war, or conflict, the supply of raw resources is disrupted, resulting in a mismatch between demand and supply, causing commodity prices to rise dramatically. During such circumstances, the market is pessimistic, causing stock prices to plummet dramatically. As a result, investing in commodities can help mitigate portfolio losses.

3) Huge leverage facility - Any slight change in commodity prices can result in exponential benefits. As a result, traders use leverage in commodity trading to generate the chance of massive gains.

4) Diversification - Stocks do well when inflation is constant or slowing. Commodities, on the other hand, do better when the inflation rate rises. Because of this negative connection, where an increase in commodity prices pulls down stock prices, losses in stocks might be offset by gains in commodity derivatives. Adding commodities to your portfolio therefore diversifies it.

5) Transparency - Digital trading platform that allows for broad-scale participation without the interference of the buyer and seller aids in fair price determination. Price determination is influenced by supply and demand, removing the possibility of price manipulation. Price discovery occurs when the price and quantity quoted by the supplier and customer precisely match. The buyer and seller remain anonymous throughout the transaction, allowing for honest price discovery with no room for deception.

II. RECURRENT NEURAL NETWORK

A recurrent neural network (RNN) is a kind of artificial neural network that is designed to work with time series data or information that contains sequences. A traditional neural network can't address a certain point in a series of events, but Recurrent Neural Networks (RNN) can. RNN have loops in them which allow the information to keep going. In Figure 1, A is a RNN with input value xt and ht are input and output respectively. The loop is useful in passing information from one step to the next of the network [19].

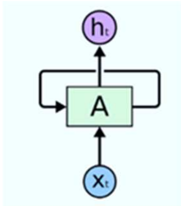

Figure 1. RNN have Loops

A recurrent neural network is composed of several reprint of the same net, where each reprint sends a signal to a successor. RNN have been used in solving many of the modern-day problems, be it speech recognition, image processing or translation of language. The major benefit of using RNN is that they might be able to connect the historical events to the current situation.

Assume a learning algorithm that is attempting to guess the upcoming word by analyzing its preceding words. If we are trying to guess the ending word in "the sun rises in the east," we do not require any more context because it is very evident that the following word will be east. RNNs can learn to utilize previous data in similar accounts, where the gap between both the useful information and the location at which it is required is minimum. However, there are times when further context is required. Try attempting to guess the very last word in the phrase "I grew up in India... I speak fluent Hindi." Above data implies that the following word is most likely some language, but if we want to pinpoint the exact language, we need to go back farther in time. It is perfectly feasible for the gap between useful knowledge and the moment where it is required to grow significantly. Unfortunately, as the gap increases, RNNs mislay their capability to grasp and connect the dots. This problem is resolved by Long Short-Term Memory (LSTM) algorithm.

III. Long Short-Term Memory Algorithm

Long Short-Term Memory algorithm, commonly known as "LSTMs," are a type of RNN which can understand long-term dependencies [10]. They work extremely well on a large type of issues and are now frequently being employed. LSTMs are expressly designed to circumvent the problem of long-term dependency [4]. All recurrent neural networks take the formation of a chain of repeating neural network modules. This chain like structure in common RNN will have a only one layer of tan h.

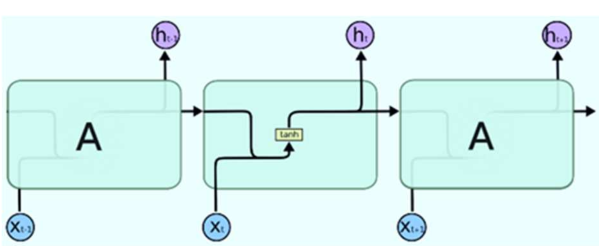

Figure 2. Single layer in RNN repeating module.

Since LSTMs is a chain-like structure, the recurrent module has different arrangement. In place of a one neural network layer, there are 4 layer that interact in a unique way. Each line in the diagram below transports a full vector from one node's output to the inputs of others. Pink nodes representing point wise actions such as vector addition, whereas Yellow Square represent trained neural network layers. Lines that amalgamate suggest concatenation, but lines that fork denote their content being replicated and the copies being sent to distinct destinations [15].

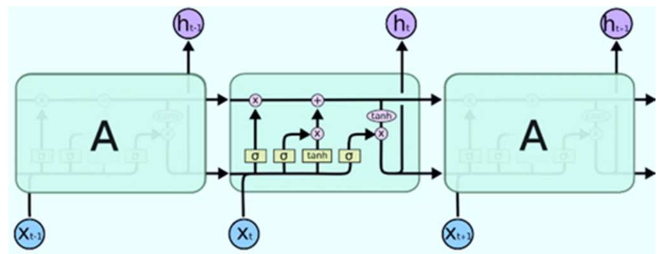

Figure 3. Four layers in LSTM repeating module

IV. Workflow

Commodities, like stocks, are linked to changes in global economic trends and factors like the supply-demand ratio. These dependencies shape price movements all the time. Our Machine Learning model express the future value of the commodity as a function of the past prices by learning the trend in price movements.

A. Data Collection

Around 500 months historical price list of various commodities like Gold, Aluminium, Copper, Crude oil was collected through various sources. The data set for each commodity contained 2 features i.e. number of columns. The first column is year and month in the form YYYYMM, where YYYY represents year and MM is the month number in that particular year. The second column is the price of the commodity in USD.

Commodities	Value Taken	No. of Months	Source
Aluminum, Copper, Crude Oil, Lead, Soybean, Tin and Zinc	Monthly average	500	https://bit.ly/3CoQIqF https://bit.ly/3E9SmxU

The Figure 4 represents the workflow of LSTM model to train and forecast the future values of commodities. Pandas and Keras libraries are used in Step-1 to put the data in a structured format and rapidly build and iterate through the model. The second steps involve data preprocessing which ensures & enhances the performance of model. Third step is of LSTM components, where an input is taken and processed in LSTM to generate an output. Matplotlib library, which is a python library to visualize data is used in the final step to generate prediction and loss graphs. Then these predicted output graphs can be used to make informed decision while trading commodities. The loss graph is used to optimize the performance of the model and configure it for more accurate performance.

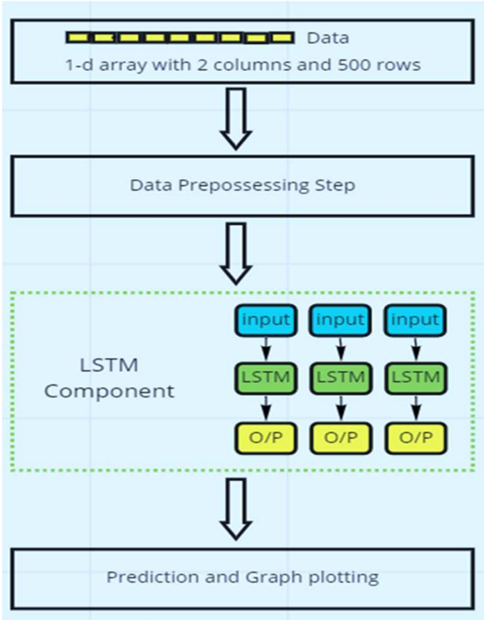

Figure 4. The steps involved in using LSTM model to train and predict the future prices

B. *Data Preprocessing*

The industry norm is to use 80% of the available data for training and rest for testing. The data needs to be divided and then for training the supervised model, the data is split into two arrays- one each for set of inputs and outputs. It is important to scale the data so that all the columns have values in the range [0,1]. Data preprocessing also involved removing rows that contain null or incorrect data. Some irrelevant text was also removed to make the data uniform for the model to read. After this, the data was normalized.

C. *Implementation*

Some factors like epochs and number of nodes effect the loss and accuracy of the model. Higher accuracy can be achieved by tuning parameters like look_back, batch size, LSTM units and num_epochs during the training of the model. For this research, an LSTM with 85 nodes was used. The dropout value is set to 0.2 and dense to 1. Using dropout layer stops over fitting by turning off few random nodes during the training phase.

Figures 5 to 7 are the graphs generated for various commodities like Aluminium, Crude Oil and Soybean with price in US$ on y-axis and day of forecast on x-axis. By changing the value of number of LSTM nodes, initial difference of 12-15% in the predicted and actual prices came down to 5-6%.

V. Results

With the help of this model, we were able to develop a simple forecasting tool. Although the model is not perfect, we do have one that can pretty closely mimic the historical data. However, we would need to fine-tune the parameters for new data. The machine learning model created by the authors for this experiment was able to predict the prices of various commodities for a period of 60 days with convincing accuracy. The model generated prediction curves and loss curves for 8 commodities and 3 of them are represented here.

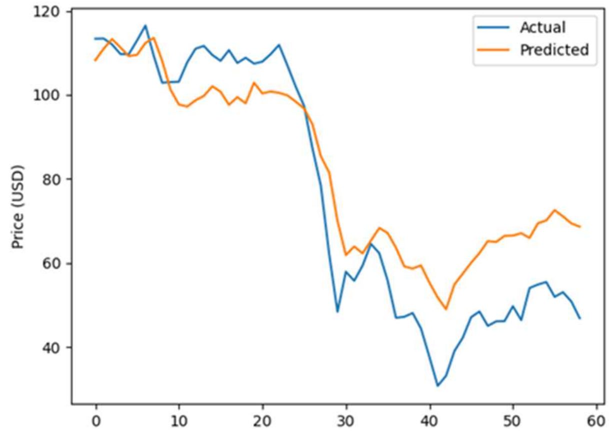

Figure 5. LSTM predicting Crude oil price

The model is able to predict long term trends in market prices with some accuracy rather than the day to day prices. In Figure 5, we can see that the values of hyper parameters such as number of neurons in the LSTM layer directly affect the accuracy of the predictions. Such hyper parameters can be further tweaked to achieve a model that works well for daily predictions as well.

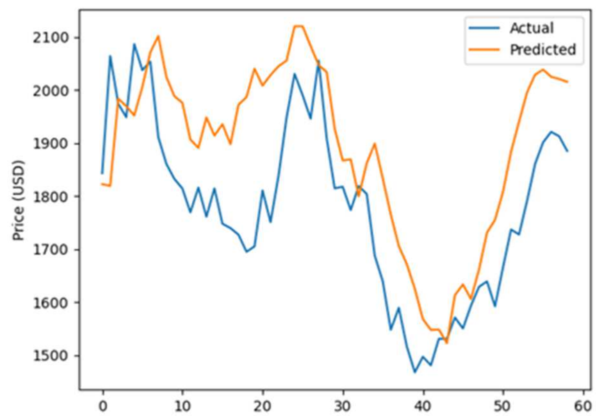

Figure 6. Prediction curve generated for Aluminium prices.

In Figure 6, the predicted prices have a slight deviation from the actual prices (as indicated in the testing data) but the trend curve is traced perfectly.

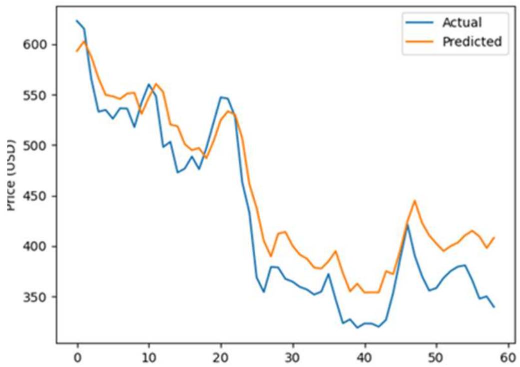

Figure 7. Prediction curve for Soybean

As for the Figure 7, it has been observed that the farther the forecast, the less accuracy the model provides. We can see that the predicted prices for 60 days in future are following the highs and lows of actual prices of soybean.

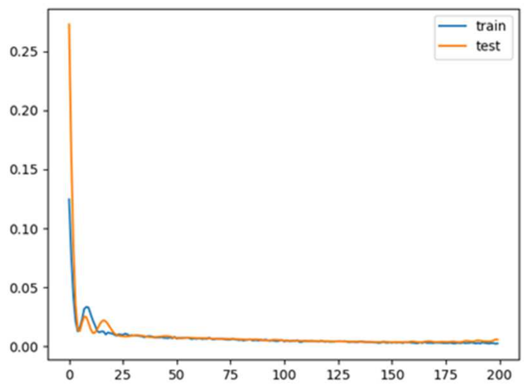

Figure 8. Loss curve for Soybean prediction

The graph in figure 8 is plotted with number of epochs on the x-axis and loss value (out of 1) on the y-axis. Number of epochs represents the number of times the model is trained on the data during the training phase. By increasing the number of epochs, training loss will tend to zero but due to over fitting, loss curve for testing will increase. Lower the value on the y-axis, the lower is the loss and hence better is the model.

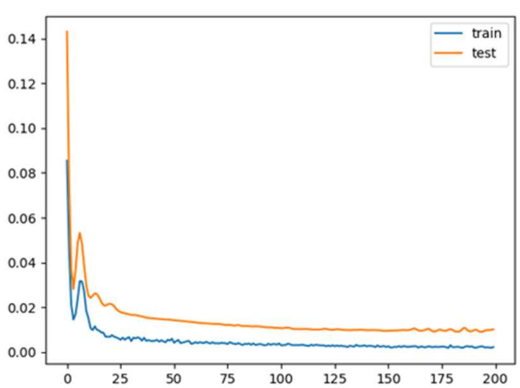

Figure 9. Loss curve for Crude Oil prediction

In Figure 9, At 200 epochs, test loss value is around 0.15. Since the plot of the training curve decreased to a point of stability and has only a small difference with the test curve, we can conclude that it is good-fit learning curve. In Figure 10, At 200 epochs, test loss value is around 0.075 for Aluminium.

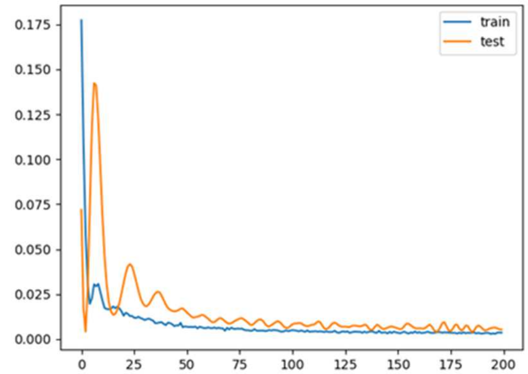

Figure 10. Loss Curve for Aluminium

VI. Conclusion

The financial and trading market has seen one of the most enormous and fastest growth that began during COVID lockdown. Through this paper we made an attempt to consolidate the major milestones and breakthroughs in financial world, especially commodity market. In this paper, we projected our finding through result of a machine learning model that uses Long Short-Term Memory algorithm (LSTM). The output consists of graphs which depicts the comparison between the actual prices of commodities that were gathered through various reliable data sources and the price predicted by the machine learning model developed by us.

In this project-based research paper, we tried to develop a machine learning model that predicts prices of commodity. Through our model we generated the prices of some commodities such as Gold, Aluminium, Soybean etc. The graphs and output data show that our model has a convincing accuracy percentage and an optimal test train loss percentage. As a result, individuals and businesses can use our suggested LSTM prediction model for commodities market forecast. This can assist investors in gaining significant monetary gain while maintaining a sustainable environment in the commodity market.

References

[1] Saydam, S., Coulton, J., & Sammut, C. (2018, March). Alternative techniques for forecasting mineral commodity prices. *International Journal of Mining Science and Technology*, *28*(2), 309–322. https://doi.org/10.1016/j.ijmst.2017.09.001

[2] Cortazar, G., Kovacevic, I., & Schwartz, E. (2013, June). Commodity and Asset Pricing Models: An Integration. *National Bureau of Economic Research*. https://doi.org/10.3386/w19167

[3] Sharma, A., Bhuriya, D., & Singh, U. (2017, April). Survey of stock market prediction using machine learning approach. *2017 International Conference of Electronics, Communication and Aerospace Technology (ICECA)*. https://doi.org/10.1109/iceca.2017.8212715

[4] Istiake Sunny, M. A., Maswood, M. M. S., & Alharbi, A.G. (2020, October 24). Deep Learning-Based Stock Price Prediction Using LSTM and Bi-Directional LSTM Model. *2020 2nd Novel Intelligent and Leading Emerging Sciences Conference (NILES)*. https://doi.org/10.1109/niles-50944.2020.9257950

[5] Eugene F. Fama "The Behavior of Stock Market Prices", the Journal of Business, Vol 2, No. 2, pp. 7–26, January 1965.

[6] Stock market challenges from http://www.google.com.

[7] A Sheta, "Software Effort Estimation and Stock Market Prediction Using Takagi-Sugeno Fuzzy Models", InProceedings of The IEEE International Conference on Fuzzy Systems, pp.171-178, Vancouver, BC, 2006.

[8] S. Selvin, R. Vinayakumar, E. Gopalakrishnan, V. K. Menon, and K. Soman, "Stock price prediction using lstm, rnn and cnn-sliding window model," in 2017 international conference on advances in computing, communications and informatics (icacci). IEEE, 2017, pp. 1643–1647.

[9] K. Chen, Y. Zhou, and F. Dai, "A lstm-based method for stock returns prediction: A case study of china stock market," in 2015 IEEE international conference on big data (big data). IEEE, 2015, pp. 2823–2824.

[10] M. Roondiwala, H. Patel, and S. Varma, "Predicting stock prices using lstm," International Journal of Science and Research (IJSR), vol. 6, no. 4, pp. 1754–1756, 2017.

[11] *What is MCX Trading? Meaning, Factors & Advantages*. (n.d.). Groww. Retrieved September 19, 2022, from https://groww.in/p/what-is-mcx-trading

[12] *Commodity News*. (n.d.). moneycontrol.com. https://www.moneycontrol.com/commodity/

[13] *Commodity Forecast 2022/2023*. (n.d.). Retrieved September 19, 2022, from https://tradingeconomics.com/forecast/commodity

[14] *Forecasting Commodity Prices: Futures Versus Judgment*. (n.d.). International Monetary Fund. https://www.imf.org/external/pubs/ft/wp/2004/wp0441.pdf

[15] *Understanding LSTM Networks -- colah's blog*. (n.d.-b). Retrieved September 19, 2022, from https://colah.github.io/posts/2015-08-Understanding-LSTMs/

[16] Jose, G. V. (2021, December 11). *Predicting Sequential Data using LSTM: An Introduction*. Medium. Retrieved September 19, 2022, from https://towardsdatascience.com/time-series-forecasting-with-recurrent-neural-networks-74674e289816

[17] Club, S. M. L. (2022, January 7). *Using LSTM's to Predict Future Stock Prices - Visionary Hub*. Medium. Retrieved September 19, 2022, from https://medium.com/visionary-hub/using-lstms-to-predict-future-stock-prices-61f4458fc860

[18] *Time series forecasting | TensorFlow Core*. (n.d.). TensorFlow. Retrieved September 19, 2022, from https://www.tensorflow.org/tutorials/structured_data/time_series

[19] Kuguoglu, K. (2022, January 22). *Predicting future values with RNN, LSTM, and GRU using PyTorch*. Medium. Retrieved September 19, 2022,from https://towardsdatascience.com/predicting-future-values-with-rnn-lstm-and-gru-using-pytorch-d9ef50991ec7

[20] *What Is a Commodity and Understanding Its Role in the Stock Market*. (2022, August 1). Investopedia. Retrieved September 26, 2022, from https://www.investopedia.com/terms/c/commodity.asp

[21] Shenmare, K. (2022, September 12). *Commodity Market*. WallStreetMojo. Retrieved September 26,2022,from https://www.wallstreetmojo.com/commodity-market/

[22] How Commodity Market Works in India? (n.d.). *5paise.Com*. Retrieved September 27, 2022, from,https://www.5paisa.com/stock-market-guide/commodity-trading/how-commodity-market-works-in-india

[23] Jose, G. V. (2021, December 11). *Predicting Sequential Data using LSTM: An Introduction*. Medium. Retrieved September 27, 2022, from https://towardsdatascience.com/time-series-forecasting-with-recurrent-neural-networks-74674e289816

[24] He, Z., Zhou, J., Dai, H. N., & Wang, H. (2019, August). Gold Price Forecast Based on LSTM-CNN Model. 2019 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech). https://doi.org/10.1109/dasc/picom/cbdcom/cyberscitech.2019.00188

[25] YURTSEVER, M. (2021, December 6). Gold Price Forecasting Using LSTM, Bi-LSTM and GRU. *European Journal of Science and Technology*. https://doi.org/10.31590/ejosat.959405

[26] Arun, V. (2018, June 17). *From scratch—An LSTM model to predict commodity prices*. Medium. Retrieved September 27, 2022, from https://medium.com/@vinayarun/from-scratch-an-lstm-model-to-predict-commodity-prices-179e12445c5a