Deep Learning Based Models: Basic LSTM, Bi LSTM, Stacked LSTM, CNN LSTM and Conv LSTM to Forecast Agricultural Commodities Prices

R. Murugesan (rmurugan@nitt.edu)
National Institute of Technology Tiruchirappalli

Eva Mishra
National Institute of Technology Tiruchirappalli

Akash Hari Krishnan
New York University Tandon School of Engineering

Research Article

Keywords: Agricultural commodities price forecast, deep learning models, Basic LSTM, Bi LSTM, Stacked LSTM, CNN LSTM, and Convolutional LSTM

Posted Date: November 1st, 2021

DOI: https://doi.org/10.21203/rs.3.rs-740568/v1

Deep learning based models: Basic LSTM, Bi LSTM, Stacked LSTM, CNN LSTM and Conv LSTM to forecast Agricultural commodities prices

R. Murugesan[*, a], Eva Mishra[a] and Akash Hari Krishnan[b]

[a] Department of Humanities and Social Sciences, National Institute of Technology, Tiruchirappalli, Tamil Nadu, India, 620015

[b] Department of Computer Science and Engineering, New York University Tandon School of Engineering, USA, 10025

*Corresponding Author: rmurugan@nitt.edu

Abstract

The literature argues that an accurate price prediction of agricultural goods is a quintessence to assure a good functioning of the economy all over the world. Research reveals that studies with application of deep learning in the tasks of agricultural price forecast on short historical agricultural prices data are very scarce and insist on the use of different methods of deep learning to predict and to this reaction of filling the gap, this study employs five versions of LSTM deep learning techniques for the task of five agricultural commodities prices prediction on univariate time series dataset of Rice, Wheat, Gram, Banana, and Groundnut spanning January 2000 to July 2020. The study obtained good forecasting results for all five commodities employing all the five LSTM models. The study validated the results with lower values of error metrics, MAE, MAPE, MSE, and RMSE and two paired t-test with hypothesis and confidence level of 95% as a measure of robustness. The study predicted the one month ahead future price for all the five commodities and compared it with actual prices using said LSTM models and obtained promising results.

Keywords: Agricultural commodities price forecast, deep learning models, Basic LSTM, Bi LSTM, Stacked LSTM, CNN LSTM, and Convolutional LSTM

1. Introduction

Countries all over the world are in sight of frequent economic and political unrest due to global food price increase (Shiferaw, 2019). The price of agricultural products, a social signal between demand and supply, fluctuates enormously owing to many factors such as climate, policies, and lop-sidedness between production and information from the marketing side, and thus, has attracted a considerable amount of attention (Weng et al., 2019). Continuous rise in food prices caused by an expeditious increase in demand for food has left more than 800 million people around the world under direct threat due to chronic malnutrition, thus, has drawn big attention for food crop price prediction (Shao & Dai, 2018). Food price increases and its

fluctuations have important macro and microeconomic impacts such as adverse effect on the GDP of a country, inflation, household problems, raise in poverty, people's financial life, reduction in nutrition, curtail of consumption of essential services of education, and health care. Further, the price of agricultural goods affects the agricultural land market, government policies, and entire agro-based industries. The variation in prices of agricultural commodities keeps the farmers under emotional and financial stress as their hard work spent in years becomes waste (Chuluunsaikhan et al., 2020; Sabu & Kumar, 2020).

As food is essential to people's day-to-day life, the accuracy in price prediction as well as knowing the prices in advance are indispensable to properly guide agricultural production, make a correct balance between supply and demand, increase farmer's income, assist the farmers to plan their next crop, and to help the government, farmers, business people in agriculture and consumers to get a clear market awareness, devising business plans, fine-tuning individuals own finances, and reducing the risks and uncertainties to be handled (Zhang et al., 2020). Though forecasting of agricultural price is pertinent and considered as paramount for many economic actors (Ly, Traore, & Dia, 2021), as literature confirms research has not reaped the benefits of deep learning-based agricultural price forecasting and remains to be an unsolved problem on date (Sagheer & Kotb, 2019; Paroissien, 2020).

Literature exhibits that numerous approaches have been employed in the prediction of agricultural commodities prices - such as data-driven and heavy assumptions (stationarity, Gaussian, and linear correlation) dependent approaches like (Chimmula & Zhang, 2020) ARIMA (Jadhav, Chinnappa, & Gaddi, 2017), SARIMA (Shahid, Zameer, & Muneeb, 2020), BP Neural Network (Weng et al., 2019), Holt-Winter method (Faniband & Shaahid, 2021). Machine learning and Deep learning approaches like - ANN, SVR, ELM, RF, SVM, (Zhang et al., 2020), k-NN, HybridNNs-ARIMAX with Exogeneous Variables (Anggraeni et al., 2019), deep-learning techniques like TDNN and LSTM network (Manogna 2020; Sabu & Kumar, 2020). However, studies with the application of deep learning in the tasks of agricultural price forecast are very scarce. It is further argued in the literature (Sabu & Kumar, 2020) that predictive analytics with deep learning techniques makes the intensive analysis even on short historical agricultural price data, and it is expected to solve the problems of all stakeholders. It is felicitous to document that research warrants for the use of new methods from deep learning to forecast (Elsheikh, Yacout, & Ouali, 2019), and to this effect of filling this gap, this study employs five versions of LSTM deep learning techniques for the task of five agricultural commodities prices prediction. So, the primary aim of this research is to apply

five versions of LSTM methods on five agricultural commodities prices for the tasks of prediction. The study contributes to the literature as follows: As a novelty, for the first time, this study employs five variants of LSTM, 1. Vanilla LSTM, 2. Bi directional LSTM, 3. Stacked LSTM, 4. CNN LSTM, and 5. Conv LSTM on five agricultural commodities prices Rice, Wheat, Gram, Banana, and Groundnut. These five agricultural commodities are commonly used, especially Rice is one of the most important staple foods around the world (Maione & Barbosa 2019). With the applications of these five LSTM variants, the present scenario in agriculture is likely to improve and help the farmers in getting some basic knowledge about the best Minimum Support Price (MSP) for their crops. Further, it could act as a nerve centre for both peasants and buyers to delve into several choices and act accordingly (Rakhra et. al., 2021). All the models used have been evaluated by four error metrics viz, MAE, MSE, MAPE, and RMSE and also tested with t-test to enhance the reliability of the models. All the five deep learning models employed have shown better performance which sheds light into the black box of non-utilization of five LSTM variants in the agricultural prices forecast.

This study has been presented with seven sections, and they are 1.Introduction, 2. Related studies on agricultural commodity prices time series, 3. Methodology, 4. Forecasting Methodology, 5. Results and Discussion, 6. Statistical analysis for hypothesis testing, **7**. Conclusion.

2. Related studies on agricultural commodity prices time series

This section presents all the relevant methodologies belonging to three main domains, statistical, machine learning, and deep learning, applied in the prediction of agricultural prices.

Dairi et al. (2021) state that in this era, many advances have been seen in artificial intelligence (AI), especially in deep learning (DL), an important part of AI. DL extracts relevant characteristics of the data automatically. As the deep learning-driven methods do not depend on feature engineering, it benefits other ML methods. Nassar et al. (2020), while comparing the achievement of deep learning price prediction models with eight statistical as well as bench mark machine learning models, on the time series datasets of Vegetables, Fruits and Flowers, demonstrated that deep learning models, LSTM and CNN-LSTM are efficient in precise prediction of Fresh Produce prices for up to three weeks advance. Sabu and Kumar (2020) used time-series and machine learning models for predicting the monthly prices of arecanut in Indian Kerala state and found that LSTM neural network was good. Weng et al. (2019), while finding the suitability of ARIMA and Deep Learning models on different data sets, daily, weekly, and monthly, identified the deep learning method as the standard

agricultural goods prices forecast. In the context of development of effective models, authors Ribeiro, M. H. D. M, & dos Santos Coelho (2019) used RF, GBM, and XGB while adopting SVR, MLP and KNN as baseline models and ranked the models as 1. XGB, 2.GBM, 3. RF, 4.MLP, 5. SVR and 6. KNN and finally concluded that that the ensemble approach was found to be doing good in the investigation of price sequences data.

In another study by Chen et al. (2019), the noise of the cabbage data was reduced using Wavelet Analysis (WA). LSTM model then was applied on the fine-tuned normalized data which was found to be producing better results in achieving accuracy. While providing a concise summary of major deep learning techniques, Zhu et al. (2018) showed that DL methods such as CNN, RNN and GAN, are gaining momentum to help researchers in agriculture price forecast. Rasheed et al. (2021) analysed the wheat prices dataset with LSTM technique. Their study presented that LSTM was performing significantly when compared to other conventional machine learning and statistical time series models. The study also stated that deep learning is fairly a new direction in agriculture.

All the studies cited above and details of related studies found in Table 1 lead to the following few important inferences. 1. There are many studies with various models (statistical, ML, and DL-based) used for prediction tasks of many agricultural commodity prices. 2. Literature studies clearly indicate that there are very few researches with applications of deep learning models in the tasks of agricultural commodity price forecasting. 3. Further, as far as our knowledge is concerned, no study with applications of five LSTM variants, vanilla or Basic LSTM, Bi directional LSTM, Stacked LSTM, CNN LSTM, and Conv LSTM for agricultural price forecasting has been done before. This study attempts to fill this huge vacuum by applying five LSTM versions.

3. Methodology

3.1 LSTM Network for modeling time series

Literature has shown that the state-of-the-art DL methods could assure good results as the temporal dependencies and structures of the time series data are learned automatically (Dairi et al., 2021; Shin et al., 2021; Peng et al., 2020 and Zeroual et al., 2020). This section devoted to shortly describing the necessary concepts of deep learning models considered, namely, Basic LSTM, Bi directional LSTM, Stacked LSTM, CNN LSTM, and Convolutional LSTM.

3.1.1 Recurrent neural network (RNN)

RNN, aka vanilla RNN, and predecessor to LSTM (Deepa, Alli, & Gokila, 2021), due to its competence to recollect serial information from (Gautam, 2021) historical data, is one of

the most famous DL architecture (Gautam, 2021) and also one of the recurrent neural network techniques used for time series forecasts (TSF). As displayed in figure 1, RNN's network lag recursion makes the outcome of the network at time *t* linked with before time *t*. The equation defining the function of the single RNN cell is given below.

$$h_t = \tanh(WT\ [h_{t-1}, k_t] + 1) \quad (1)$$

Where WT, x, h_t, and h_{t-1} are weight matrix, bias matrix, and hidden state at current and previous time-steps, respectively. But when TSF requires long-term context memorization, RNN could not overcome vanishing and exploding gradients problems (Arun Kumar et al., 2021).

Table 1. Related studies

Name of the authors	Name of the commodities	Deep Learning Models used for prediction	Results
R L et al. (2021)	Cottonseed, Castor seed, Rape mustard seed, Guar seed, soybean seed	LSTM Base line models: ARIMA, TDNN	The LSTM model provided a better forecast**.**
Ouyang et al. (2019)	Cotton, Sugar, bean, bean II, soya bean oil, cardamom, strong Wheat, Corn, Coffee, cocoa, Frozen orange juice	LSTNet Base Line Models: CNN, RNN, ARIMA, VAR	The LSTNet performed better results over the r baseline methods on average.
Kurumatani K. (2020)	Cabbage, Tomato, Lettuce	LSTM (Recurrent neural network)	The LSTM performed the best result.
Jin et al. (2019)	Chinese cabbage, Radishes	LSTM	The optimum performance was obtained by the LSTM.
Prakash & Farzana, (2019)	Tomato	LSTM	The LSTM is one of the most effective models for dealing with nonlinear patterns in prediction.
Chen et al. (2021)	Chicken, Chili, Tomatoes	LSTM Baseline models: ARIMA, SVR, Prophet, XGBoost	Among the five baseline models, the LSTM was forecasted to produce the best results.

3.1.2 Long short-term Memory (LSTM)

LSTM network models are very useful for time-series data (Sezer, Gudelek, & Ozbayoglu, 2020). All its hyperparameters, viz, hidden layers, units in each layer, network weight initialization, momentum values, batch size, gradient clipping, gradient normalization, dropouts, number of epochs, decay rate, learning rate, activation functions, optimization and

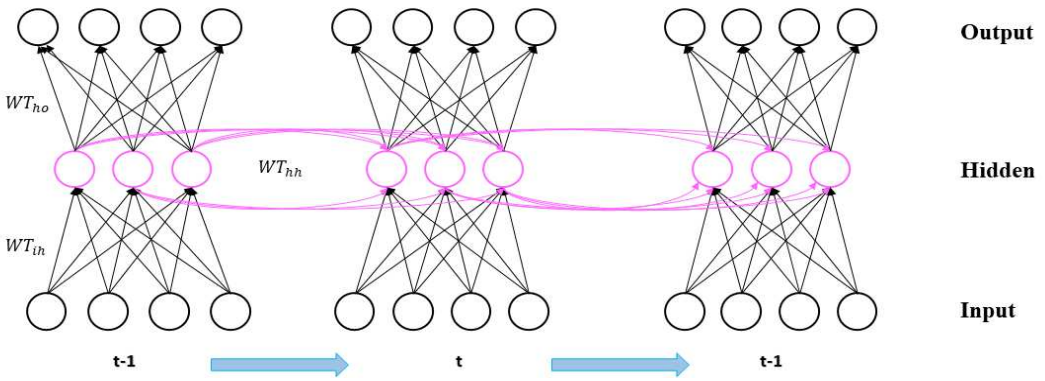


Fig 1. RNN time sequence process

sequence length for LSTM, are similar to those of RNN (Sezer, Gudelek & Ozbayoglu, 2020). LSTM, (Hochreiter & Schmidhuber, 1997; Arun Kumar et al., 2021) characterized by more memory cells compared to RNN truncates the gradients by enforcing constant error flows. It finally facilitates the traditional RNN algorithm to approximate the protracted information with significant delays and obtains better time series prediction (Arun Kumar et al., 2021**;** Chen et al., 2021; Deepa, Alli, & Gokila, 2021). Figure 2 exhibits the architecture of LSTM cell. As seen in figure 2, there are four main components in LSTM cell, identified as four gates, *FG*, *OG, IG* and *UG*. These four function in a unique way to acquire long-term remembrance, short-term remembrance, input serial at a given time period resulting in generating new long-term remembrance, new short-term remembrance, and output serial at a given time period. The different function served by every gate is explained as follows:

Step 1: The input gate IG_t which decides the input information be transferred to the memory cell, is mathematically given as:

$$\mathrm{IG}_t = \sigma\left(W_{\mathrm{IG}} * \left[h_{t-1}, k_t\right] + \mathbf{1}_{\mathrm{IG}}\right) \tag{2}$$

Step 2: Forget gate (FG_t), which controls the information to be neglected is mathematically expressed as:

$$\text{Forget Gate } \left(\mathrm{FG}_t\right) = \sigma\left(W_{\mathrm{FG}} * \left[h_{t-1}, k_t\right] + \mathbf{1}_{\mathrm{FG}}\right) \tag{3}$$

Step 3: Updated Gate ($\widetilde{FG}_t$) controls and updates the output information flowing out of the cell, is given by:

$$\text{Updated Gate}\left(\tilde{\text{UG}}_t\right) = \tanh\left(W_{\text{UG}} * \left[h_{t-1}, k_t\right] + \mathbf{1}_{\text{UG}}\right) \tag{4}$$

$$\text{UG}_t = \text{FG}_t * \text{UG}_{t-1} + \text{IG}_t * \tilde{\text{UG}}_{t-1}$$

(5)

Step 4: Output Gate (OG_t) which updates the previous hidden state is mathematically expressed as:

$$\text{Output Gate}\ (\text{OG}_t) = \sigma\left(W_{\text{OG}} * \left[h_{t-1}, k_t\right] + \mathbf{1}_{\text{OG}}\right) \tag{6}$$

$$h_t = \text{OG}_t * \tanh \tag{7}$$

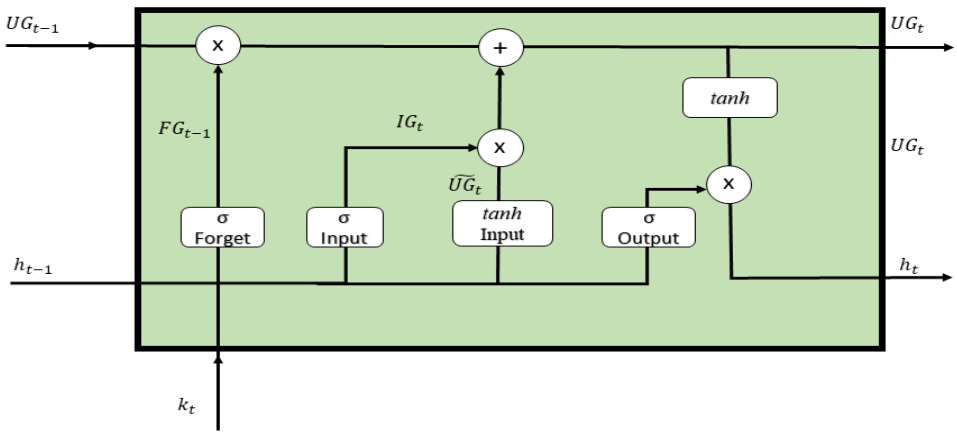


Fig 2. LSTM architecture cell

In equations 2-7,W_{IG},W_{FG}, W_{UG} ,W_{OG},l_{IG},l_{FG}, l_{UG}, l_{OG} are Weight and bias matrices. "*" is the element-wise multiplication of the vectors, and the sigmoid functions is mathematically given as found below.

$$f\ (x) = \frac{1}{(1+e^{-x})} \tag{8}$$

3.1.3 Conv LSTM

Convolutional Neural Network, another extensively applied deep learning architecture, has shown its state-of-the-art exploration in time series data too after its colossal accomplishments in the areas of image detection, speech, and facial recognition (Maity & Chatterjee, 2012; Vidal & Kristjanpoller, 2020). Its composition of many layers with distinctive architectures called convolutional layers, and subsampling (pooling) layers (Wang, Mu, & Liu, 2021; Zhang et al., 2021; Vidal & Kristjanpoller, 2020) extracts apt information by understanding the internal representation of the data by convolution operation and pooling operation (Livieris, Pintelas, & Pintelas, 2020; Ji et al., 2020). Its exclusively designed data pre-processing layers generate new feature values by applying convolution operation between the raw input data and convolution kernels, a technique originally used to extract features from image datasets (Livieris, Pintelas, & Pintelas, 2020). Conv LSTM (Shastri et al. 2020) solves

the shortcomings of standard fully connected long short-term memory- FCLSTM, through its 3D tensors characteristics of, 1. Input K_1, K_2, $K_3 \ldots K_t$, 2. Cell output, M_1, M_2, $M_3 \ldots M_t$, and 3. Hidden states h_1, h_2, $h_3 \ldots h_t$, (Figure 3). The input and past state in Conv LSTM fixes the future state of cell with its convolutional operator (*) in the transmission of state to state and input to state (Arora, Kumar, & Panigrahi, 2020). The key equations 9 to 13 express the Conv LSTM mathematically where "*" and "`" denote convolutional operator and Hadamark product (Shastri et. al 2020).

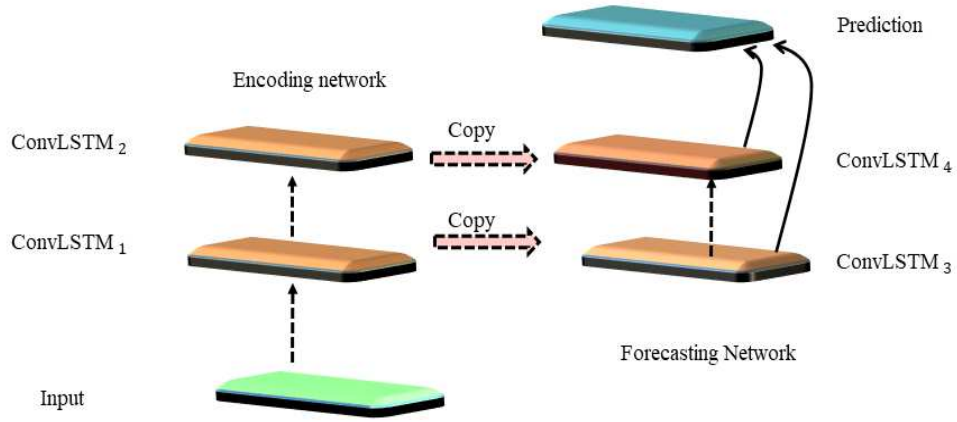


Fig 3. Architecture of convolutional LSTM

$$IN_t = \sigma\left(\mathbf{W}_{KIN} * IN_t + \mathbf{W}_{hIN} * h_{t-1} + \mathbf{W}_{MIN} \cdot M_{t-1} + x_{IN}\right) \quad (9)$$

$$FS_t = \sigma(\mathbf{W}_{KFS} * K_t + \mathbf{W}_{hFS} * h_{t-1} + \mathbf{W}_{MFS} \cdot M_t + x_{FS}) \quad (10)$$

$$M_t = FS_t \cdot M_{t-1} + IN_t \cdot \tanh\left(\mathbf{W}_{KM} * K_t + \mathbf{W}_{hM} * K_t + \mathbf{W}_{hM} * h_{t-1} + x_M\right) \quad (11)$$

$$N_t = \sigma\left(\mathbf{W}_{KN} * K_t + \mathbf{W}_{hN} * h_{t-1} + \mathbf{W}_{MN} \cdot M_t + x_N\right) \quad (12)$$

$$h_t = N_t \cdot \tanh\left(M_t\right) \quad (13)$$

3.1.4 Bidirectional LSTM (Bi LSTM)

Bi LSTM is a combination of bi-directional RNN and LSTM cell (Graves & Schmidhuber, 2005; Shastri et. al., 2020). Bi LSTM, a version of LSTM, is considered good for addressing time series data (Kulshrestha, Krishnaswamy**,** & Sharma, 2020). Contrary to LSTM, Figure 4, Bi-LSTM processes the sequences of data in both directions - first to last input and last to first and uses both backward and forward information through its architecture that comprises of forwarding and backward LSTM layers (Kulshrestha, Krishnaswamy, & Sharma, 2020; Mahto et al. 2021). The complete forecasted outputs are obtained by forwarding pass by running all the inputs for time $1 \leq k \geq K$. For time k = 1 to K, and for time k = K to 1, onward pass and reverse pass for onward states and reverse states are executed. In the same way, reverse pass for onward states for time k = K to 1 and for time k = 1 to K reverse states

are performed after finding the objective function derivative used in onward pass for time 1 ≤ k ≥K for reverse pass (Chen et. al., 2019, Kulshrestha, Krishnaswamy, & Sharma, 2020).

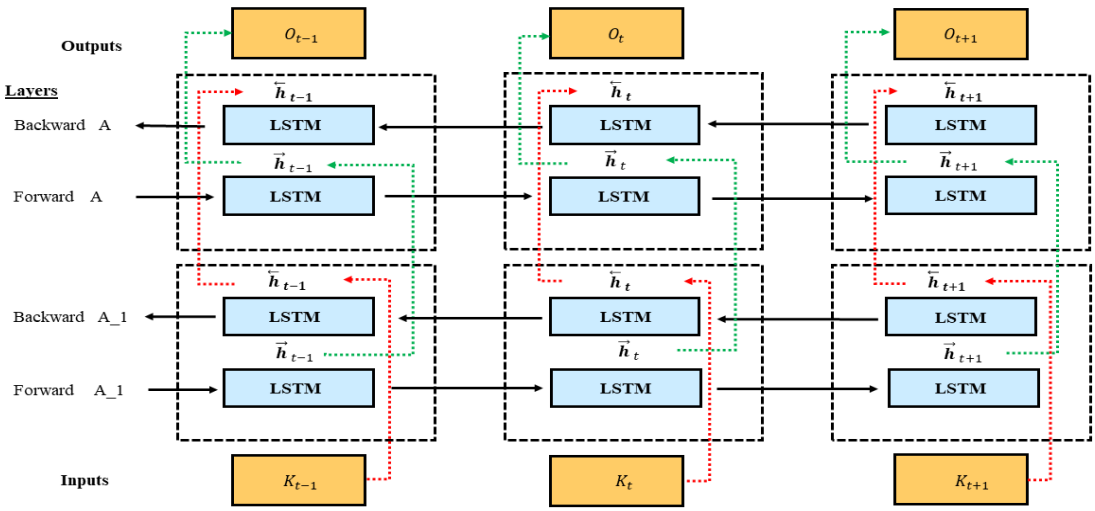


Fig 4. Bi directional architecture

As both the forward and backward layer brings the aggregate output of Bi LSTM, the mathematical equation for the Bi LSTM output is given in equation 20. The equations 14 to 20 mathematically express the layer A of Bi LSTM at time t.

$$O_t = \mathbf{W}_{hO}^{\text{ı}} \overset{\text{ı}}{h}_t + \mathbf{W}_{hO}^{\text{ɔ}} \overset{\text{ɔ}}{h}_t + x_O \tag{14}$$

$$\overset{\text{ɯ}}{FS}{}_t^A = \sigma\left(\mathbf{W}_{FSh}^{A\,\text{ɯ}} h_{t-1}^A + \mathbf{W}_{FSh}^{A\,\text{ɯ}} h_{t-1}^A + x_{FS}^{A\,\text{ɯ}}\right) \tag{15}$$

$$\overset{\text{ɯ}}{IN}{}_t^A = \sigma\left(\mathbf{W}_{INh}^{A\,\text{ɯ}} h_{t+1}^A + \mathbf{W}_{INK}^{A\,\text{ɯ}} h_t^{A+1} + x_{IN}^{A\,\text{ɯ}}\right) \tag{16}$$

$$\overset{\text{ɔ}}{\bar{M}}{}_t^A = \tanh\left(\mathbf{W}_{Mh}^{A\,\text{ɔ}} h_{t+1}^A + \mathbf{W}_{MK}^{A\,\text{ɔ}} h_t^{A-1} + x_M^{A\,\text{ɔ}}\right) \tag{17}$$

$$\overset{\text{ɔ}}{M}{}_t^A = \overset{\text{ɯ}}{FS}{}_t^A + \overset{\text{ɔ}}{M}{}_{t+1}^A + \overset{\text{ɯ}}{IN}{}_t^A \cdot \overset{\text{ɔ}}{\bar{M}}{}_t^A \tag{18}$$

$$\overset{\text{ɔ}}{N}{}_t^A = \sigma\left(W_{Nh}^{A\,\text{ɔ}} h_{t+1}^A + W_{Nh}^{A\,\text{ɔ}} h_t^{A+1} + x_N^{A\,\text{ɔ}}\right) \tag{19}$$

$$\overset{\text{ɔ}}{h}{}_t^A = \overset{\text{ɔ}}{N}{}_t^A \cdot \tanh\left(\overset{\text{ɔ}}{M}{}_t^A\right) \tag{20}$$

3.1.5 Stacked LSTM

Stacked LSTM or Deep LSTM also identified as a multilayer fully connected structure has numerous hidden layers with numerous memory cells. All the numerous LSTM layers are stacked together, resulting in more model complexity and augmented intensity of the model (Shastri et al., 2020; Devaraj et. al., 2021), In stacked LSTM, figure 5, the present layer obtains the value from the preceding layer, because preceding layers' input are learned by the next higher level layers for greater optimized results. Stacked LSTM administers output for every

time-step and not the single output for all time steps (Arora, Kumar, & Panigrahi, 2020). The mathematical formulation of model Ath LSTM layer for an unrolled stacked LSTM is found in equations.

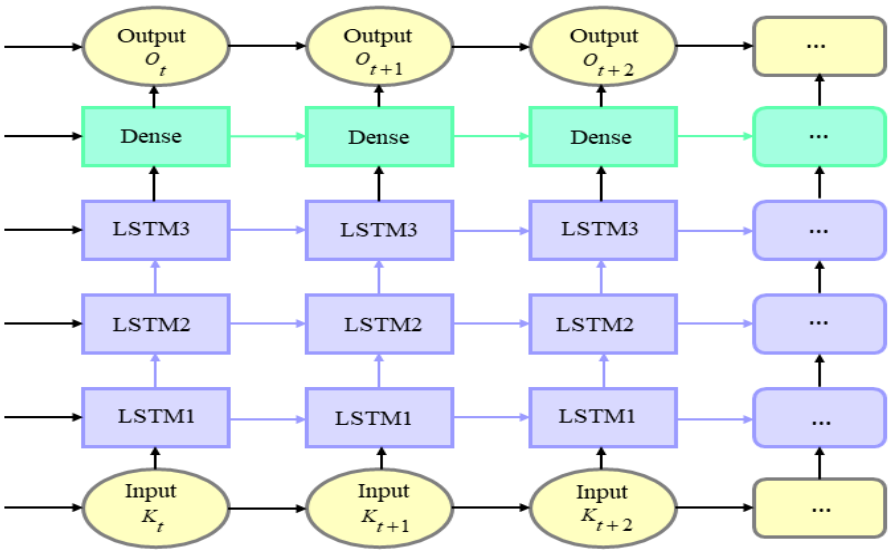


Fig 5. Stacked LSTM architecture

$$FS_t^A = \sigma\left(\mathbf{W}_{FSh}^A h_{t-1}^A + \mathbf{W}_{FSK}^A h_t^{A-1} + x_{FS}^A\right) \tag{21}$$

$$IN_t^A = \sigma\left(\mathbf{W}_{INh}^A h_{t-1}^A + \mathbf{W}_{INK}^A h_t^{A-1} + x_{IN}^A\right) \tag{22}$$

$$\bar{M}_t^A = \tanh\left(\mathbf{W}_{\bar{M}h}^A h_{t-1}^A + \mathbf{W}_{\bar{M}K}^A h_{t-1}^A + x_{\bar{M}}^A\right) \tag{23}$$

$$M_t^A = FS_t^A \cdot M_{t-1}^A + IN_t^A \cdot \bar{M}_t^A \tag{24}$$

$$N_t^A = \sigma\left(\mathbf{W}_{Nh}^A h_{t-1}^A + \mathbf{W}_{NK}^A h_{t-1}^A + x_N^A\right) \tag{25}$$

$$h_t^A = N_t^A \cdot \tanh\left(M_t^A\right) \tag{26}$$

3.1.6 CNN LSTM

The time series' big problem of not possessing the capability of providing inputs based on images and the feature matrices and elements like lines, vertices, angles, and shadows which are getting recognized by CNN, is overcome by a new framework called Gramian Angular Fields (GAF) and Markov Transition Fields (MTF). This new framework encodes sequences as different types of images through which the applications of computer vision techniques for classification have been made feasible (Vidal & Kristjanpoller, 2020). The overview of the CNN-LSTM forecasting model architecture is depicted in figure 6.

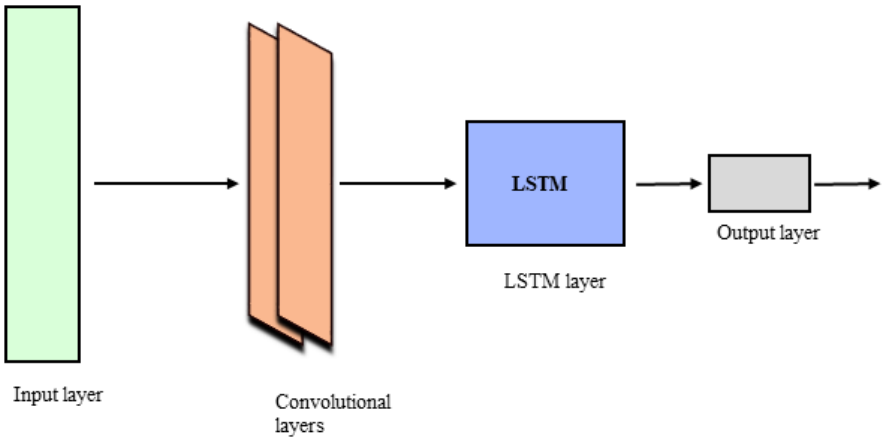


Figure 6. Overview of CNN-LSTM forecasting model architecture

4. Forecasting methodology

4.1 Model Description

Compared to other time-series data, agricultural commodities prices data are relatively small. The challenge here is to obtain a good forecast based on the investigation performance of these five deep learning models in these types of time series data. Parameters of the constructed five deep learning LSTM variants, Basic LSTM, Bi LSTM, Stacked LSTM, CNN LSTM, and Convolutional LSTM are furnished in Table 2.

4.2 Data preparation

While dealing with supervised classifiers, time series needs to be formatted into datasets consisting of independent and dependent variables. Hence each commodity in the original dataset is turned into the form of X (the independent variable) and y (the dependent variable):

$$X[i] = [\text{value in period } P_1 \quad \text{value in period } P_2 \quad \text{value in period } P_3]$$
$$y[i] = [\text{value in period } P_4]$$

4.3 Basic LSTM

LSTM is an improved version of RNN architecture created to model chronological sequences and their future dependencies more precisely than traditional RNNs. The basic version of LSTM in this implementation consists of one layer in input, hidden and output. The activation function used is $relu$, with $adam$ optimizer and loss function being mse. The serial data for each commodity is converted to the format mentioned earlier and fed into the LSTM for 200 epochs.

4.4 Stacked LSTM

Several hidden LSTM layers can be stacked on top of another to avail of the benefit is that the input values fed to the network go through more than a few LSTM layers and proliferate through time within one LSTM cell. Hence, parameters are well dispersed within several layers resulting in thorough processing of inputs in each time step. An LSTM layer needs a 3D input,

and LSTMs, by default, will produce a 2D output from the end of the sequence. The obstacle of producing 2D output from 3D is addressed by having the LSTM output a value for each time step in the input data by setting the $return\_sequences = True$ argument on the layer. This optional parameter is permitting us to have a three-dimensional output from the disguised LSTM layer as input to the next. In this implementation, stacked LSTM has all the same other parameters as Basic LSTM above.

Table 2. Model parameters

MODEL	PARAMETERS
Basic LSTM	50 output units, single dense hidden layer, 'relu' activation, 'adam' optimization, 'mse' loss
Stacked LSTM	50 output units, two dense hidden layer, 'relu' activation, 'adam' optimization, 'mse' loss
Bi directional LSTM	50 output units, single two-way dense hidden layer, 'relu' activation, 'adam' optimization, 'mse' loss
CNN LSTM	CNN: 64 filters, 2 kernels, pool_size = 0.5 LSTM: 50 output units, single dense hidden layer, 'relu' activation, 'adam' optimization, 'mse' loss
Convolutional LSTM	CNN: 64 filters, (1,2) kernels LSTM: single dense hidden layer, 'adam' optimization, 'mse' loss

4.5 Bi directional LSTM

On some serial prediction practices, it could be advantageous to permit the LSTM model to grasp the input serial in one plus one, binary directions and concatenate both interpretations. This model is called a Bi directional LSTM. A Bi directional LSTM allows the network to have both the forward and backward information about the sequences at each step. A Bi directional LSTM for single time series forecasts estimation is implemented by enclosing the first concealed layer in an enclosure layer called Bi directional. Similar to the previously defined LSTM models, this implementation of Bi directional LSTM has the same parameters.

4.6 CNN LSTM

A Convolutional LSTM uses a process called kernel convolution to break down a graphical image into its essential properties. By combining this with deep learning, CNNs can 'understand' image data, which leads to the possibility of making predictions about or diagnostics of it. It can effectively extricate and grasp characteristics from single-dimensional sequence data. A CNN combination of LSTM at the rear, where the CNN gets accustomed to understanding subsequence of all input that are subsequently jointly administered as input to an LSTM model, gives rise to a CNN LSTM. In order to work with CNN LSTMs, the data

needs to be reformatted into a suitable input for CNNs. The time series data is first split into samples of input and output with four times as input and one time as output, instead of the usual three inputs and one output done for the previous models. Each input sample is then divided into two sub-samples of two-period steps which will be interpreted by the CNN and furnish a time series to the LSTM model for further process. The LSTM at backend used is a basic LSTM model as defined before.

The total CNN model shall be enclosed for the purpose of reusing the same CNN for every sub-serial of the data where the Time Distributed wrapper will administer the whole model once per input. The number of filters is defined as the number of times the input sequence is read, two in this implementation. The kernel size is depending upon the availability number of 'read' operation of the input sequence, which is two again here. The layer of max-pooling follows the convolutional layer filters only the salient features and makes it down to fifty percent in size and finally flatten down to single one-dimensional vector. This vector is to be used as a single input time step by the LSTM at the backend.

4.7 Convolutional LSTM

Similar to the CNN LSTM discussed before, the convolutional LSTM removes the CNN in favour of a convolutional reading directly built into each LSTM unit. This model was specifically developed for working with two-dimensional spatial-temporal series, but it has been altered to work with univariate time-series sequences in this implementation. The serial data is modified from its original form to a format similar to the previously defined CNN LSTM model; in addition, the subsequence will need to be converted into two-dimensional vectors. This is achieved by considering them as vectors of shape$[2 \times 1]$, i.e., the second dimension is given size 1. The Conv LSTM model is built with a single layer with attributes being the number of filters, similar to CNN LSTM, and a two-dimensional kernel size compared to a single-dimensional kernel in CNN LSTM. The kernel size is defined in terms of [rows, columns]. As mentioned before, due to treating the one-dimensional series as two-dimensional, the number of rows is fixed to 1, and the number of columns will be the size of subsequence in the kernel. After this is done, the output of the Convolutional layer must be flattened and fed to the LSTM model to be used for prediction.

4.8. Performance evaluation metrics

R^2, MSE, MAPE, RMSE and MAE are the indexes used for the evaluation of all the NN-based predictive models for which the mathematical framework is found in equations 27 to 31. The conceptual model is presented in figure 7.

$$R^2 = \frac{\sum_{t=1}^{n}\left[\left(x_{t_i} - \bar{x}\right)\cdot\left(\hat{x}_t - \bar{x}\right)\right]^2}{\sqrt{\sum_{t=1}^{n}\left(x_{t_i} - \bar{x}\right)^2}\cdot\sqrt{\sum_{t=1}^{n}\left(\hat{x}_t - \bar{x}\right)^2}} \quad (27)$$

$$\text{MSE} = \frac{1}{N}\sum_{t=1}^{N}\left(\hat{x}_t - x\right)^2 \quad (28)$$

$$MAPE = \frac{100}{n}\sum_{t=1}^{n}\left|\frac{x_t - \hat{x}_t}{x_t}\right|\% \quad (29)$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(x_t - \hat{x}_t\right)^2} \quad (30)$$

$$\text{MAE} = \frac{\sum_{t=1}^{n}\left|x_t - \hat{x}_t\right|}{n} \quad (31)$$

Where x_t is the actual values and $\hat{x}_t$ is the predicted values of the models

4.9. System Specifications

All experiments were carried out in the Google Colaboratory basic version, with GPU hardware acceleration. The detailed specifications are given in Table 3.

4.10. Dataset

The data-set used in this research is obtained from RBI (Reserve Bank of India) website, Government of India. It has five independent time-series data of wholesale agricultural commodity prices from January 2000 to July 2020. The five agricultural commodities time series are Rice, Wheat, Gram, Banana, and Groundnut, which are the most important food crops for civilizations around the world, and the intrinsic complexity of the commodity market, has always been proven that the price forecast is an intractable task (Kamdem, Essomba & Berinyuy, 2020). Each time series dataset has been divided into training (80%), and testing (20%) datasets the 20 percent testing set, which ensures the evaluation of the forecasting models, has a considerable size of unseen ‘‘out-of-sample’’ data. (Livieris, Pintelas, & Pintelas, 2020).

Table 3. System Specification for Google Colab, and Modules, APIs, libraries used

Parameter	**Google Colab**	**Name of the modules, APIs, and Libraries used**
GPU	Nvidia KBO / T4	Keras == 2.4.3,
GPU Memory	12 GB / 16 GB	matplotlib == 3.2.2,
GPU Memory Clock	0.82GHz / 1.56GHz	numpy == 1.19.5,
Support Mixed Precision	No / Yes	pandas == 1.0.5,
GPU Release Year	2014 / 2018	protobuf == 3.15.6
No CPU Cores	2	scikit_learn == 0.24.2,
Available RAM	12 GB (upgradable to 26.75GB)	statsmodels == 0.11.1
Disk Space	358 GB	

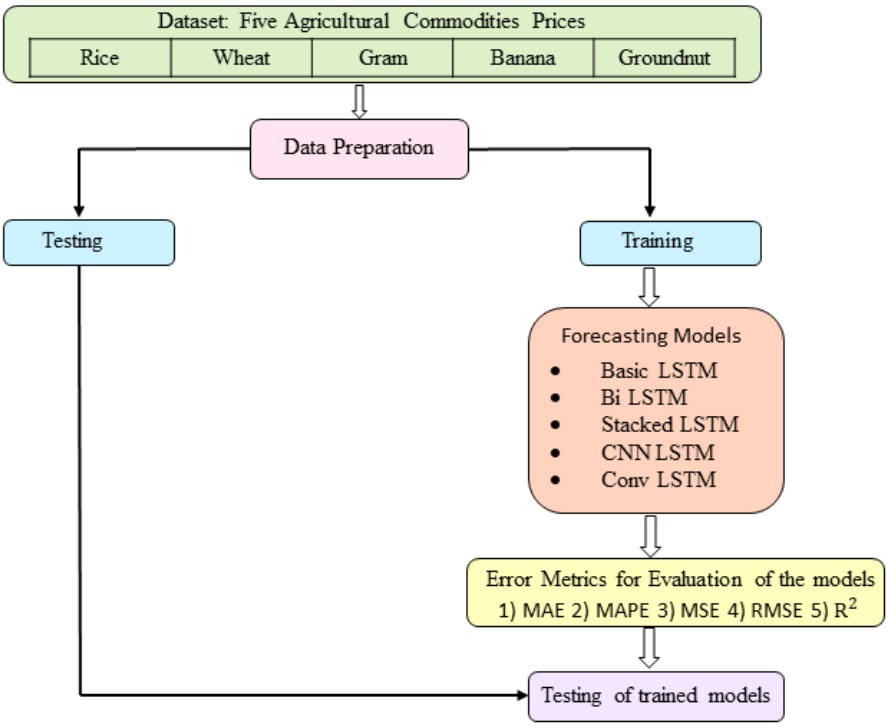


Fig 7. Architecture of forecasting model on agricultural commodities prices

5. Results and discussion

After initial analysis in connection with statistical properties and unit root test on time series data, all the five LSTM variants, Basic LSTM, Bi LSTM, Stacked LSTM, CNN LSTM and Convolutional LSTM have been applied to five (Rice, Wheat, Gram, Banana, and Groundnut) agricultural commodity prices time-series datasets from January 1st, 2000 to July, 1st 2020. Table 4 exhibits the statistical properties of agricultural commodity prices. The skewness data shows positive skewness, and also it is symmetrical as these are between -0.5 0.5 figures except for the products gram and onion. The result on kurtosis clearly presents that the excess kurtosis statistic is purely an indication of the Leptokurtic features of the data. Shapiro-Wilk (SW) test results (p-value less than 0.01) clearly show that the variables are not following a normal distribution and hence clearly pinpoint (Hansen, McDonald, & Nelson, 1999; Kulshrestha, Krishnaswamy, & Sharma, 2020) that non-parametric methods are deserving application for such a dataset. Table 5 reports the results of the standard stationarity tests ADF, and PP. Based on both ADF and PP, the time series data are stationary after I differencing. However, the five deep learning models, Bi LSTM, Basic LSTM, Stacked LSTM, CNN LSTM and Conv LSTM employed to handle the agricultural time series data are assumption-free and efficient in the extraction of apt information from time-dependent data.

But stationarity problem remains to be an important issue when statistical modeling is used as a benchmark model.

The loss vs epochs graphs for all the five LSTM versions of all the commodities plotted in figures 8(a-e) to 12(a-e), clearly exhibit a good learning rate, indicating the transformation of arrangement of data into the congruous format. This further confirms that each training sample consists of a series of data points. Thus, the series of data is fed into LSTM layers of all five variants, and the previous time step output is taken to the next subsequent input series.

The focus of the study is to predict the prices of five agricultural commodities using five state-of-the-art deep learning LSTM variants individually on each of all five agricultural commodities in the present of small-sized datasets. Figures 13(a-e) to 17(a-e) depicts the forecasting results from the five deep learning models applied on every agricultural commodity price time series. To enhance clarity and visibility, the forecasting results of five models based on the five agricultural commodity prices, have been displayed separately. So, figures 13(a-e) are an independent visual depiction of actual and forecasted values of five models for the agricultural commodity Rice. Similarly, figures 14(a-e), 15(a-e), 16(a-e) and figures 17(a-e) are displaying the actual and forecasted values of all the five models i.e. Basic LSTM, Bi directional LSTM, stacked LSTM, CNN LSTM and Convolutional LSTM for all the remaining four agricultural commodities to say, Wheat, Gram, Groundnut and Banana respectively. For all deep learning-based models and for all the commodities, the blue lines indicate actual values, and the red line indicates the forecasted values. From illustrations of actual and predicted values of considering commodities for all the five models shown in figures 13(a-e) to 17(a-e), it can be observed that all the five models show good forecasting performance. Both the actual values line and forecasted values line for all the models and for all the commodities are close to each other, which are visual evidence for the good forecasting performance. To assess quantitatively the performances of the five forecasting models, the forecast accuracy MAE (Mean absolute error), MAPE (Mean absolute percentage error), MSE (Mean square error), RMSE (Root mean square error), and R squared are computed and tabulated in Table 6. It can be easily seen that the forecasting performance of all the five state-of-the art deep learning techniques on all five agricultural commodities under study is showing promising results and considered as good since all the models are providing lower values of MAPE, MAE, RMSE. All the five models have the values ranging from 0.13 to 0.36 for MAE, 0.37999 to 0.65 for MSE, 0.000794 to 0.00235 for MAPE and 0.151657 to 0.487852 for RMSE respectively. Figures 18, 19, and 20 visually show the performance of all five LSTM variants with lower MAE, MAPE, and RMSE values.

Table 4. Statistical properties of agricultural commodity prices

Variable	Mean	SD	SE	95% Conf	Interval	Shapiro-Wilk stat	Shapiro-Wilk p value **	Skewness	Kurtosis
Rice	99.4371	35.5003	2.2588	94.9880	103.8862	0.86756	8.73519e-14	0.26029	1.52876
Wheat	99.8283	34.0883	2.1690	95.5561	104.1005	0.93793	3.48114e-11	0.25680	1.71620
Gram	107.8602	40.5852	2.5824	102.7738	112.9466	0.89342	3.40009e-12	1.31470	5.55831
Banana	93.4859	40.3166	2.5653	88.4332	98.5386	0.90272	1.47122e-11	0.21727	1.52965
Groundnut	91.3239	31.4668	2.0022	87.3803	95.2675	0.93859	1.20500e-08	0.27518	1.83390

** SW cut-off p value is 0.05.

Table 5. Unit root test results

Variable	ADF**								KPSS**							
	Level				1st Diff				Level				1st Diff			
	P value	CV 10%	CV 5%	CV 1%	P value	CV 10%	CV 5%	CV 1%	P value	CV 10%	CV 5%	CV 1%	P value	CV 10%	CV 5%	CV 1%
Rice	0.9876	-2.573	-2.874	-3.458	0.0	-2.573	-2.874	-3.458	0.0100	0.3470	0.4630	0.7390	0.1000	0.3470	0.4630	0.7390
Wheat	0.9967	-2.573	-2.874	-3.458	0.0	-2.573	-2.874	-3.458	0.0100	0.3470	0.4630	0.7390	0.1000	0.3470	0.4630	0.7390
Gram	0.2968	-2.573	-2.874	-3.458	0.0024	-2.573	-2.874	-3.458	0.0100	0.3470	0.4630	0.7390	0.1000	0.3470	0.4630	0.7390
Banana	0.6384	-2.573	-2.874	-3.458	0.0452	-2.573	-2.874	-3.458	0.0100	0.3470	0.4630	0.7390	0.1000	0.3470	0.4630	0.7390
Groundnut	0.808	-2.573	-2.874	-3.458	0.0005	-2.573	-2.874	-3.458	0.0100	0.3470	0.4630	0.7390	0.1000	0.3470	0.4630	0.7390

** The data is stationary, when p-value < 0.05 in ADF Test, and in KPSS Test, when p-value > 0.05

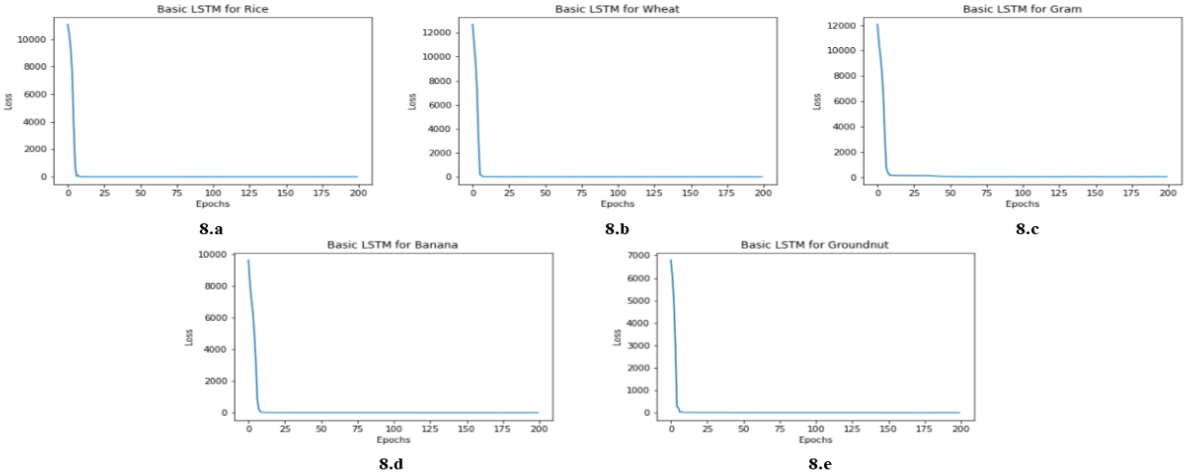


Fig 8a-e. Basic LSTM Loss

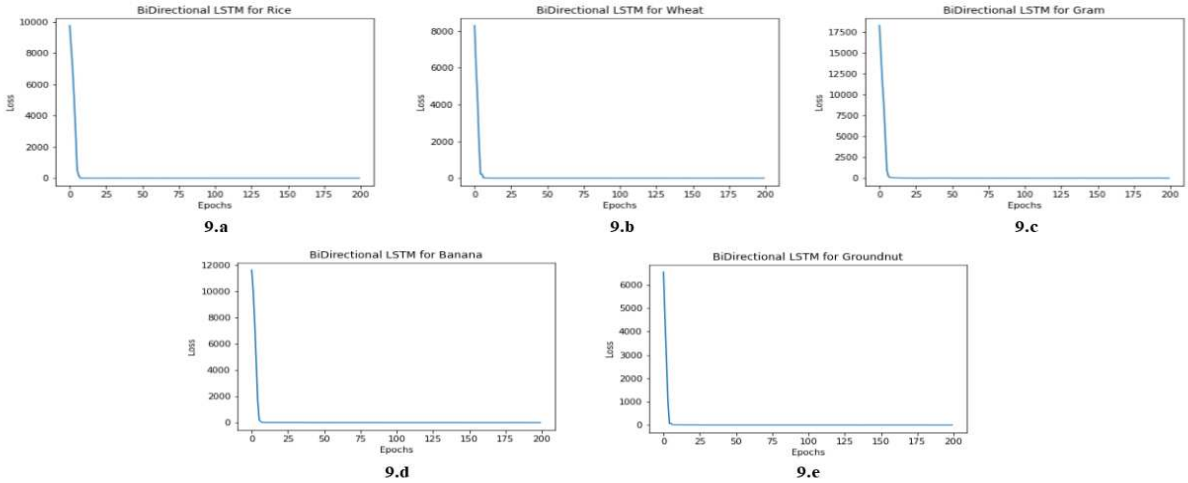


Fig 9a-e. Bi LSTM Loss

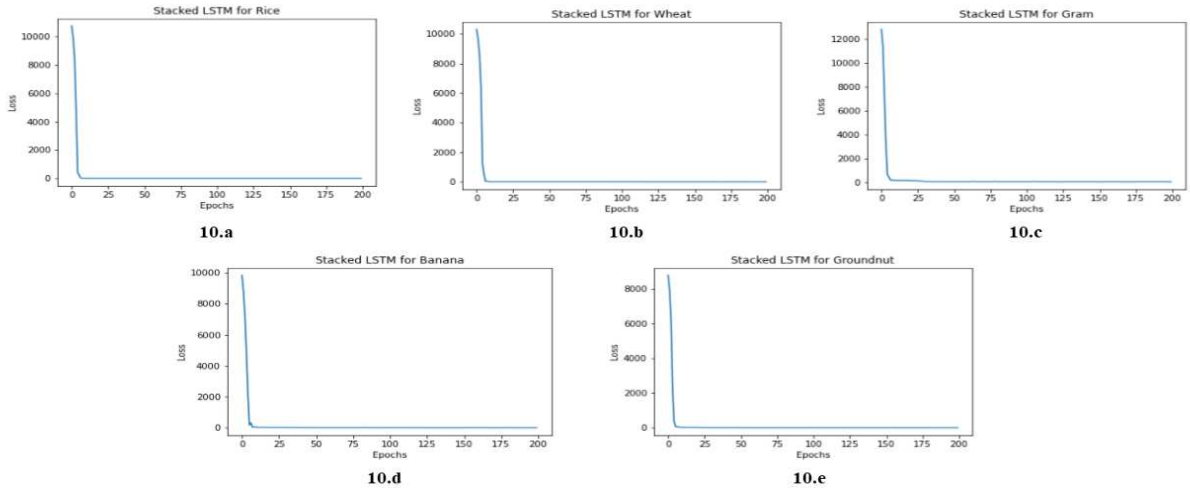


Fig 10a-e. Stacked LSTM Loss

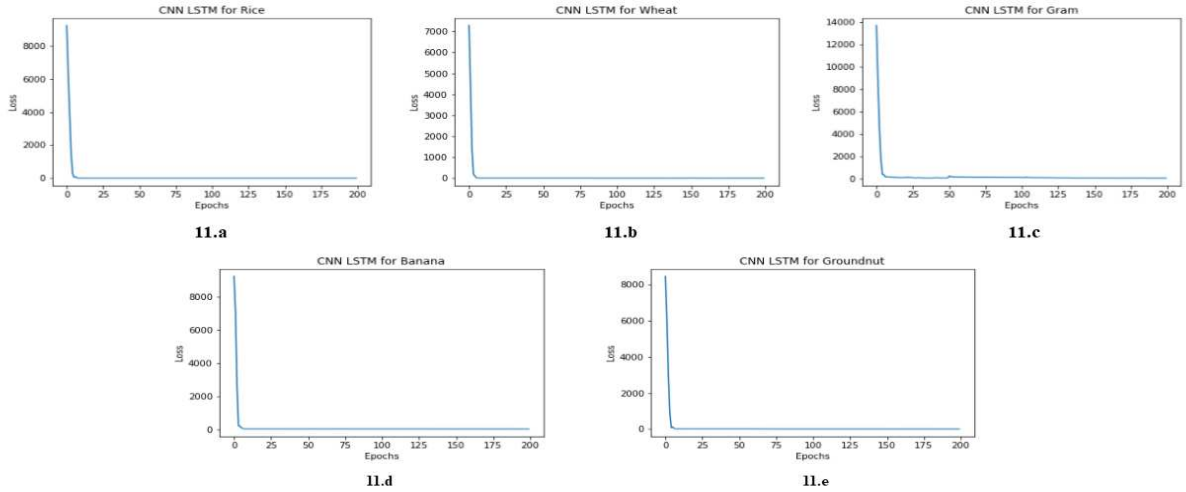


Fig 11a-e. CNN LSTM Loss

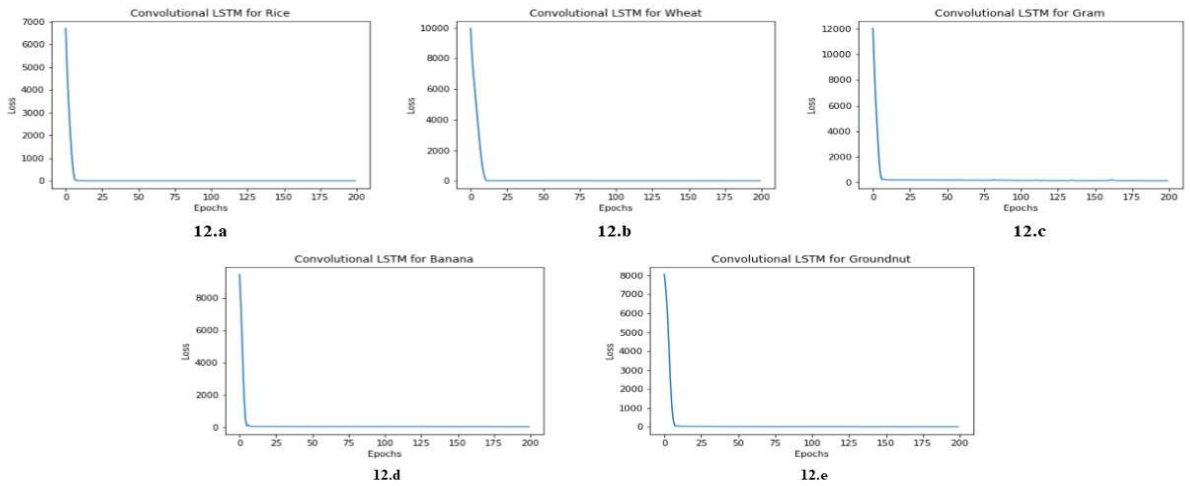


Fig 12a-e. Conv LSTM Loss

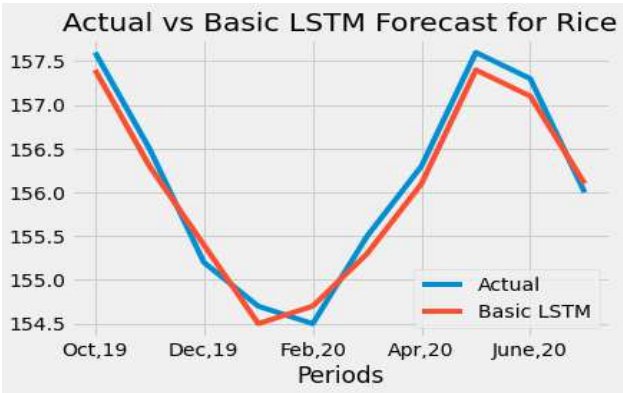


Fig 13a. Actual vs Basic LSTM Forecast for Rice

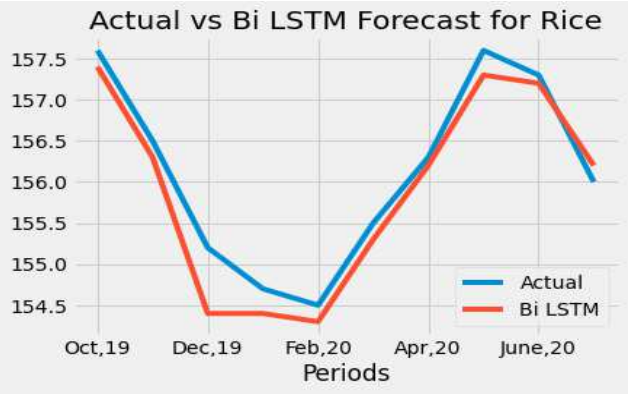


Fig 13b. Actual vs Bi LSTM Forecast for Rice

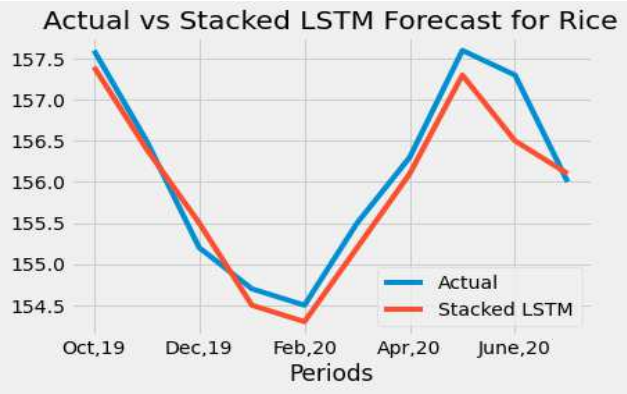


Fig 13c. Actual vs Stacked LSTM Forecast for Rice

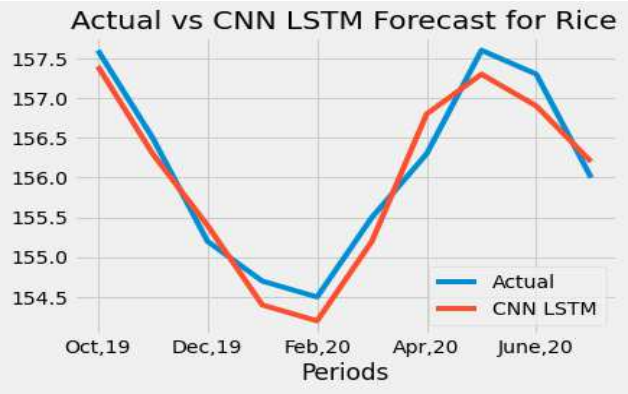


Fig 13d. Actual vs CNN LSTM Forecast for Rice

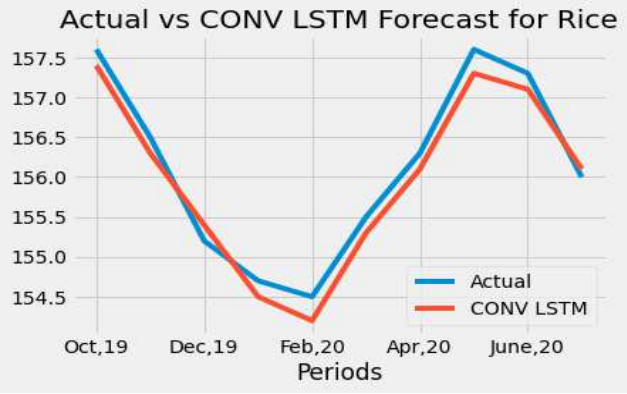


Fig 13e. Actual vs Conv LSTM Forecast for Rice

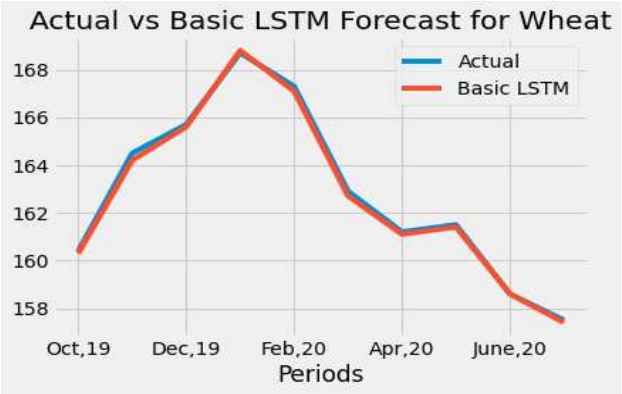


Fig 14a. Actual vs Basic LSTM Forecast for Wheat

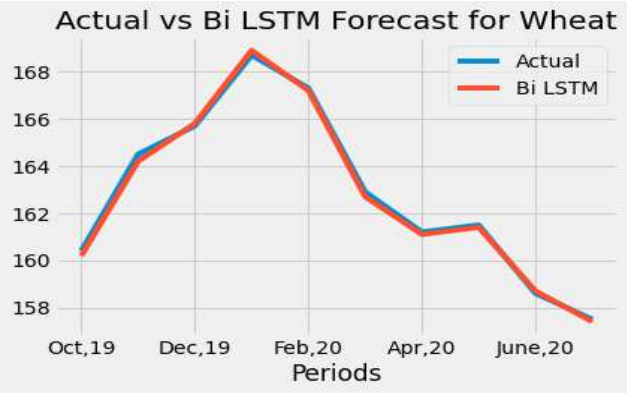


Fig 14b. Actual vs Bi LSTM Forecast for wheat

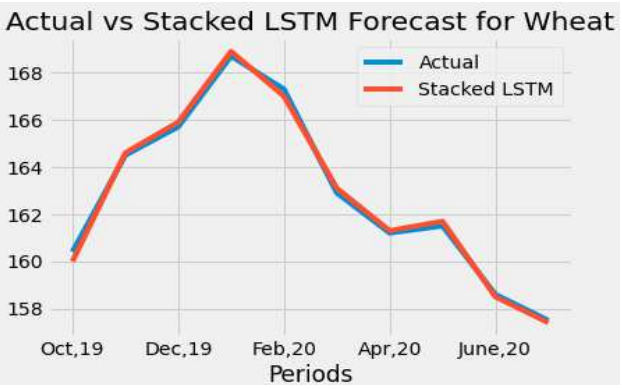


Fig 14c. Actual vs Stacked LSTM Forecast for Wheat

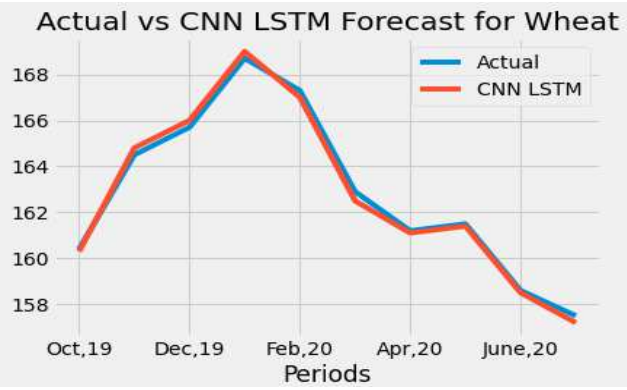


Fig 14d. Actual vs CNN LSTM Forecast for wheat

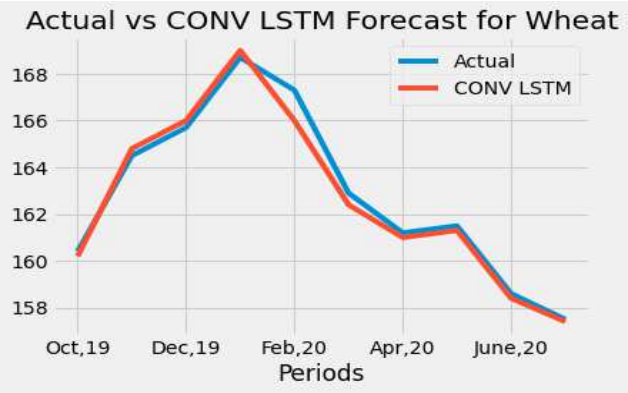


Fig 14e. Actual vs Conv LSTM Forecast for wheat

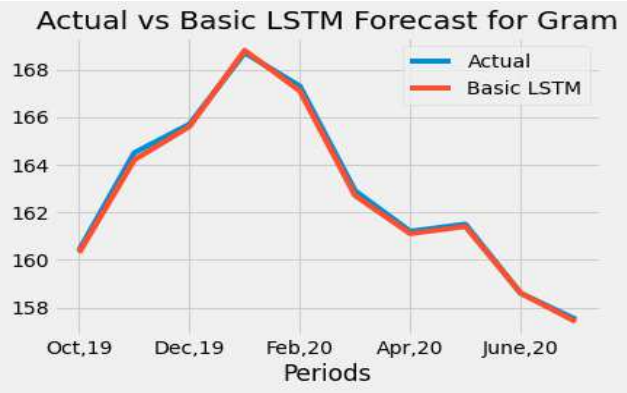


Fig 15a. Actual vs Basic LSTM Forecast for Gram

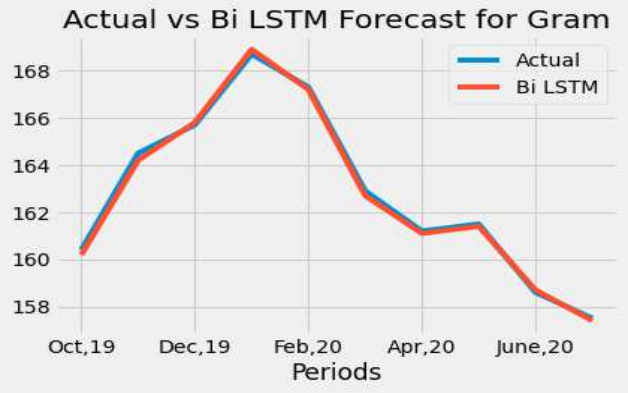


Fig 15b. Actual vs Bi LSTM Forecast for Gram

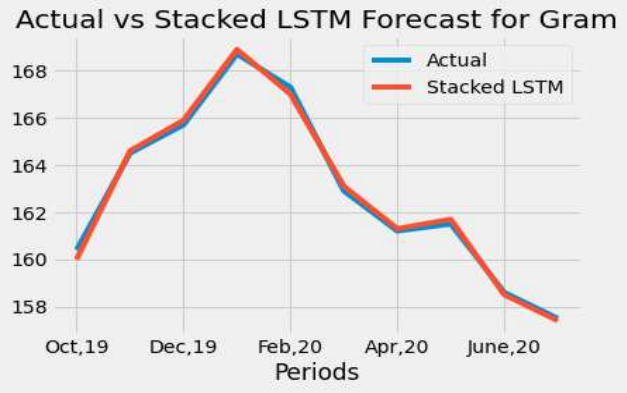


Fig 15c. Actual vs Stacked LSTM Forecast for Gram

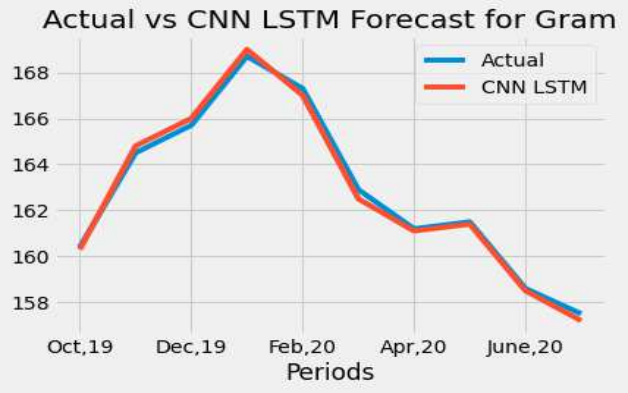


Fig 15d. Actual vs CNN LSTM Forecast for Gram

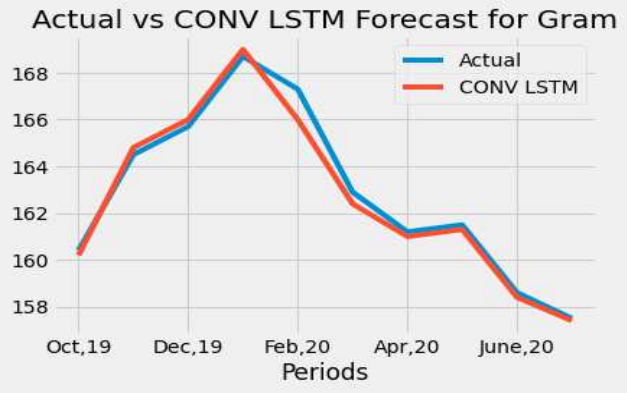


Fig 15e. Actual vs Conv LSTM Forecast for Gram

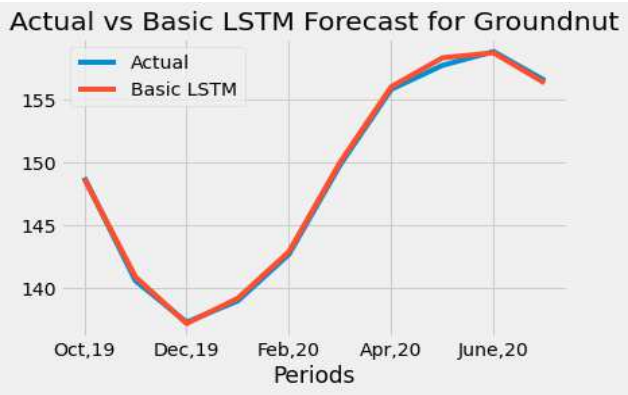


Fig 16a. Actual vs Basic LSTM Forecast for Groundnut

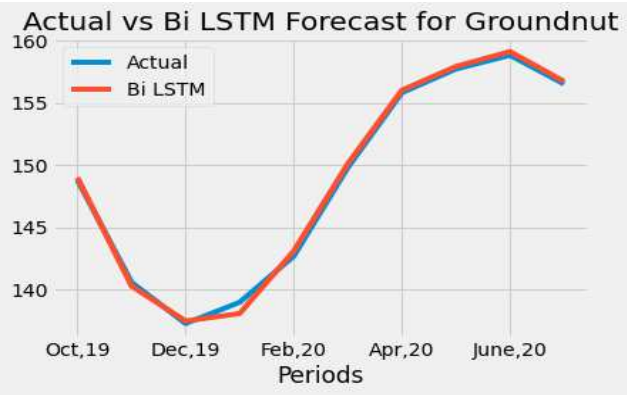
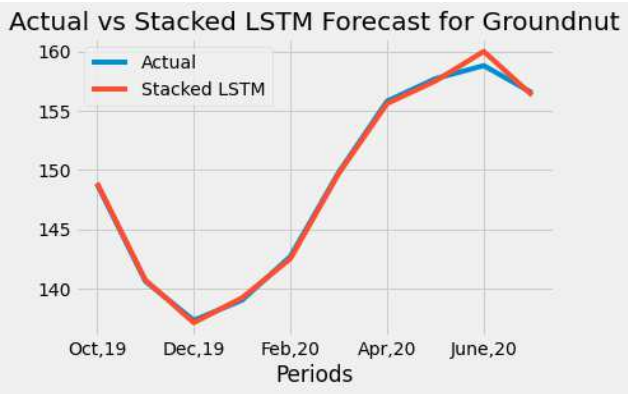
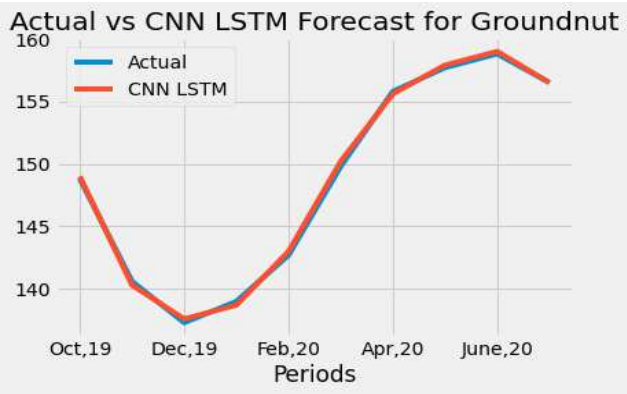


Fig 16b. Actual vs Bi LSTM Forecast for Groundnut



16c. Actual vs Stacked LSTM Forecast for Groundnut



16d. Actual vs CNN LSTM Forecast for Groundnut

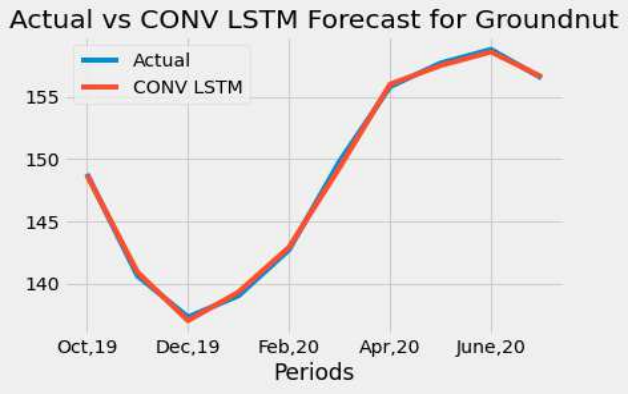


Fig 16.e. Actual vs Conv LSTM Forecast for Groundnut

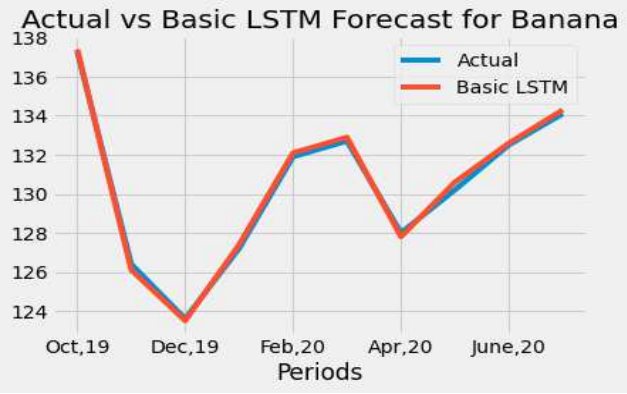


Fig 17a. Actual vs Basic LSTM Forecast for Banana

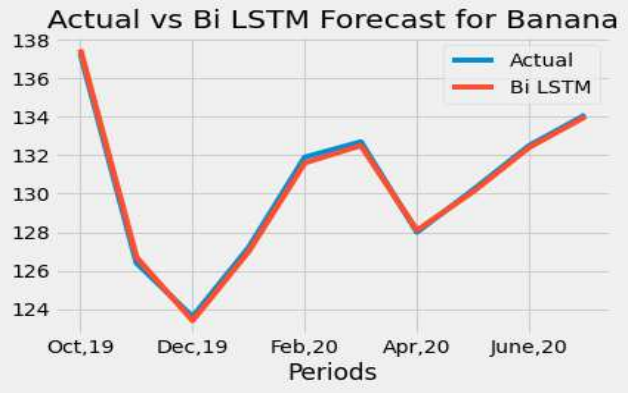


Fig 17b. Actual vs Bi LSTM Forecast for Banana

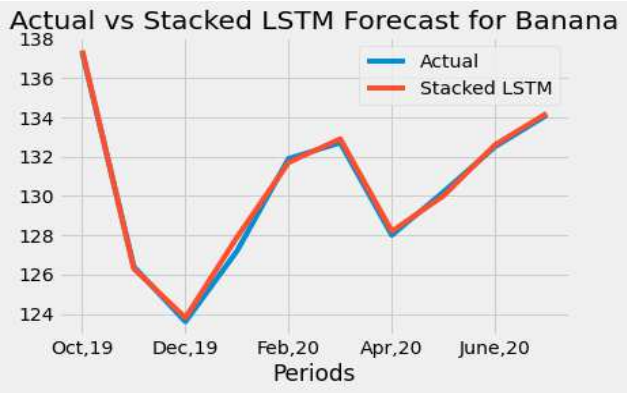


Fig 17c. Actual vs Stacked LSTM Forecast for Banana

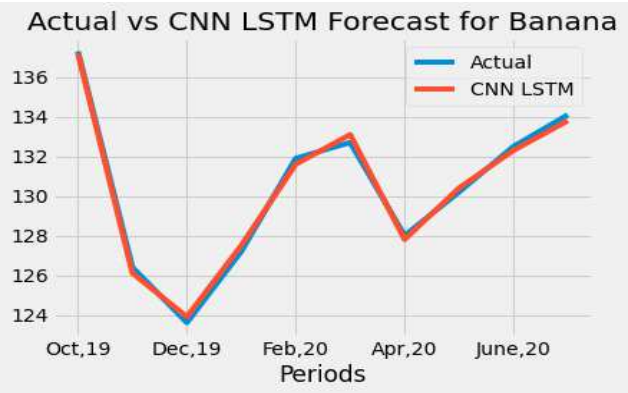


Fig 17d. Actual vs CNN LSTM Forecast for Banana

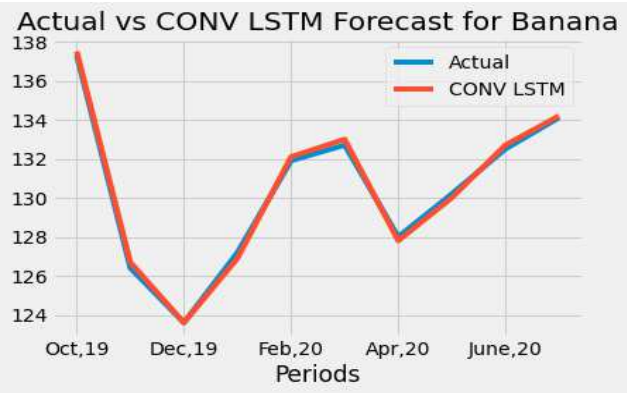


Fig 17e. Actual vs Conv LSTM Forecast for banana

Table 6. Evaluation metrics of all models for all commodities

Commodity	Model	MAE	MSE	MAPE	RMSE	R^2
Rice	Basic LSTM	0.18999	0.36999	0.00121	0.192353	0.96873
	Bi LSTM	0.259999	0.103999	0.001668	0.322490	0.91213
	Stacked LSTM	0.269999	0.109000	0.001726	0.330151	0.90790
	CNN LSTM	0.289999	0.092999	0.001857	0.304959	0.92142
	Conv LSTM	0.20999	0.469999	0.001345	0.216794	0.96029
Wheat	Basic LSTM	0.130000	0.230000	0.000794	0.151657	0.96811
	Bi LSTM	0.150000	0.270000	0.000919	0.163431	0.96778
	Stacked LSTM	0.190000	0.450000	0.001163	0.212132	0.96631
	CNN LSTM	0.230000	0.065000	0.001405	0.254950	0.95467
	Conv LSTM	0.360000	0.238000	0.002187	0.487852	0.94050
Gram	Basic LSTM	0.300000	0.152000	0.002071	0.389871	0.96848
	Bi LSTM	0.150000	0.270000	0.001038	0.164316	0.97617
	Stacked LSTM	0.230000	0.650000	0.001599	0.254950	0.97799
	CNN LSTM	0.290000	0.103000	0.001996	0.320936	0.96542
	Conv LSTM	0.340000	0.136000	0.002350	0.368781	0.95074
Banana	Basic LSTM	0.200000	0.480000	0.001538	0.219089	0.96683
	Bi LSTM	0.179999	0.037999	0.001385	0.194935	0.96492
	Stacked LSTM	0.210000	0.073000	0.001626	0.270185	0.96518
	CNN LSTM	0.260000	0.074000	0.002002	0.272029	0.96511
	Conv LSTM	0.199999	0.047999	0.001531	0.219080	0.96683
Groundnut	Basic LSTM	0.220000	0.068000	0.001470	0.260768	0.99890
	Bi LSTM	0.319999	0.143999	0.002193	0.379473	0.99219
	Stacked LSTM	0.269999	0.170990	0.001773	0.413521	0.99219
	CNN LSTM	0.239999	0.0679999	0.001640	0.260768	0.998905
	Conv LSTM	0.240000	0.070000	0.0016338	0.264575	0.99887

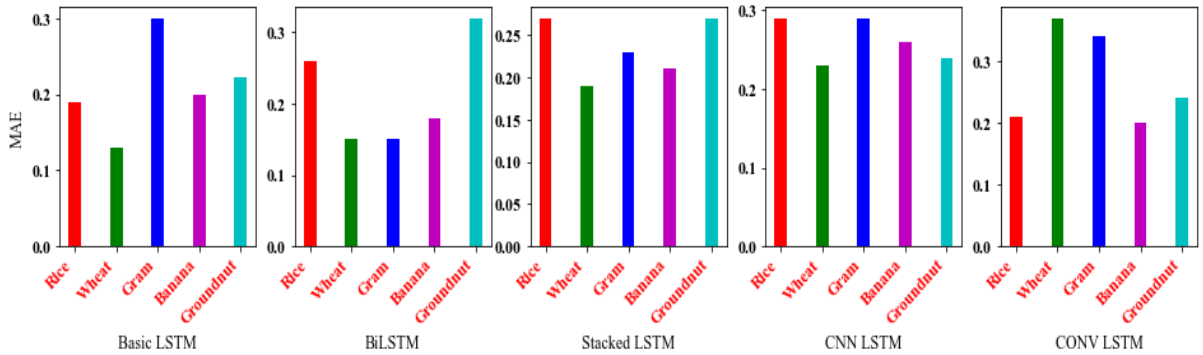


Fig 18. MAE comparison of all five LSTM variants

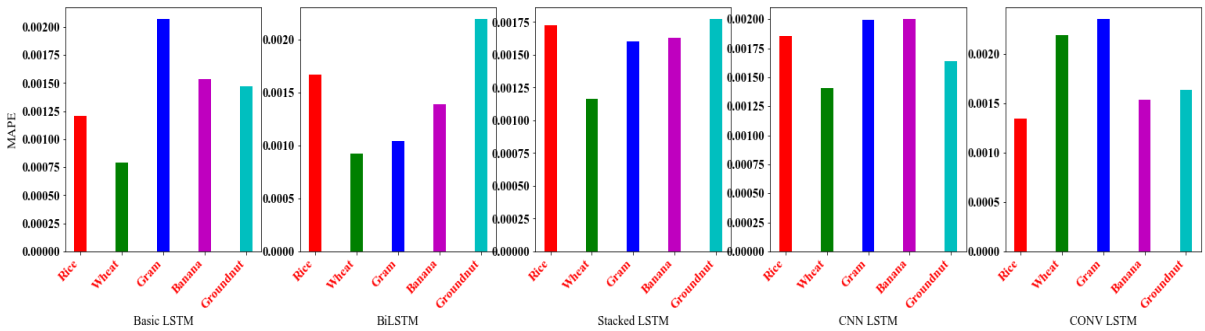


Fig 19. MAPE comparison of all five LSTM variants

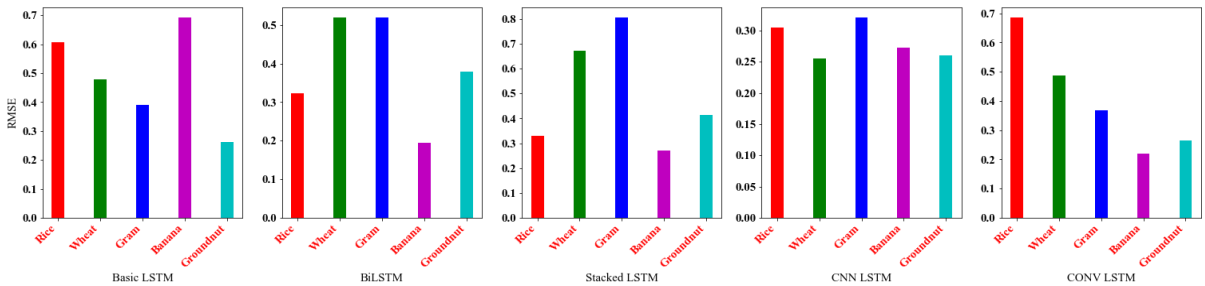


Fig 20. RMSE comparison of all five LSTM variants.

Table 7. Results of paired t-test with status of hypothesis

Pair	t-stat	p-value	Status
Wheat-Basic	1.761	0.112	H_O to be accepted
Wheat-Bi LSTM	1.413	0.191	H_O to be accepted
Wheat-Stacked LSTM	-0.142	0.891	H_O to be accepted
Wheat-CNN LSTM	0.600	0.563	H_O to be accepted
Wheat-Conv LSTM	1.191	0.264	H_O to be accepted
Rice-Basic	1.588	0.147	H_O to be accepted
Rice - Bi LSTM	1.799	0.121	H_O to be accepted
Rice -Stacked LSTM	1.111	0.164	H_O to be accepted
Rice -CNN LSTM	1.160	0.276	H_O to be accepted
Rice -Conv LSTM	1.775	0.018	H_O to be accepted
Gram-Basic	0.629	0.545	H_O to be accepted
Gram-Bi LSTM	0.188	0.239	H_O to be accepted
Gram-Stacked LSTM	0.118	0.909	H_O to be accepted
Gram-CNN LSTM	1.596	0.209	H_O to be accepted
Gram-Conv LSTM	1.882	0.115	H_O to be accepted
Banana-Basic	-1.77	0.269	H_O to be accepted
Banana-Bi LSTM	0.970	0.357	H_O to be accepted
Banana-Stacked LSTM	-1.194	0.200	H_O to be accepted
Banana-CNN LSTM	-1.327	0.310	H_O to be accepted
Banana-Conv LSTM	-1.327	0.275	H_O to be accepted
Groundnut-Basic	-1.555	0.154	H_O to be accepted
Groundnut-Bi LSTM	0.647	0.534	H_O to be accepted
Groundnut-Stacked LSTM	0.365	0.723	H_O to be accepted
Groundnut-CNN LSTM	0.967	0.359	H_O to be accepted
Groundnut-Conv LSTM	0.227	0.825	H_O to be accepted

Table 8. Predicted values and error % for all models and for all the commodities

Goods	Time step	Actual values	Basic LSTM		Bi LSTM		Stacked LSTM		CNN LSTM		Conv LSTM	
			PV*	Error%	PV	Error%	PV	Error%	PV	Error%	PV	Error%
Wheat	Aug 20	157.5	156.2	0.83	159.6	-1.33	159.8	-1.46	160.8	-2.10	154.3	2.03
Gram	Aug 20	142.3	137.1	3.65	136.4	4.15	135.4	4.85	134.5	5.48	148.6	-4.42
Banana	Aug 20	139.1	132.1	5.03	136.0	2.22	144.9	-4.17	145.1	-4.09	135.2	2.80
Groun-dnut	Aug 20	153.7	151.5	1.43	158.6	-3.19	158.1	-2.86	150.3	2.12	149.9	2.47
Rice	Aug 20	156.9	154.7	1.40	154.6	1.47	159.3	-1.53	156.2	0.45	155.4	0.96

*PV – predicted values

6. Statistical analysis for hypothesis testing

Paired samples t-test has been used to find out whether any significant difference exists between the forecasted results and the actual values on account of the application of five deep learning techniques on five agricultural commodities. To execute the paired samples t-test, a null hypothesis is formed based on 95% confidence interval. The description of the hypothesis test is found as: Null hypothesis, H_o: $\mu_1 = \mu_2$ and Alternative Hypothesis H_1: $\mu_1 \neq \mu_2$

Where μ_1 and μ_2 denote the mean values of actual values of the time series data and the mean values of forecasted of all the models used, the suggestion of t-test is the null hypothesis can be accepted at 0.05 level of significance. The result of the paired t-test has been shown in Table 7. As per the results of statistic testing, the agricultural price forecasting using all the five LSTM variants are acceptable. This further confirms that forecasting agricultural time's series prices of the small dataset with better accuracy using these five LSTM variants looks promising.

7. Prediction of future price of Rice, Wheat, Gram, Banana, and Groundnut

The predicted values of all the five commodities, Rice, Wheat, Gram, Banana, Groundnut, for August 2020 (one month ahead) and its comparison with reported actual values of these five commodities, are described in Table 8. As found in Table 8, the range of error percentage between predicted values and the actual values for all the models and for all the agricultural commodities is 0.83 to 5.48. The results show a maximum percentage error of 5.48 for CNN LSTM model for the commodity Gram and 4.85 and 4.17 for model stacked LSTM for the commodities Gram and Banana, respectively. Broadly speaking, all the results are very promising, and once again, it substantiates the result of getting better forecast and reliable agricultural commodities price forecast using deep learning-based state-of-the-art techniques employed by the study. This throws good information to all stakeholders. To enhance visual understanding, the particulars given in Table 8 is represented with figures. 21, 22, 23, 24, and 25 which display the actual and predicted values owing to applications of employed LSTM techniques. Each exhibits both the actual and predicted values and error percentage shown on the top of the respective bars for every commodity separately for better visual clarity.

8. Conclusion

Extant literature emphasizes that the role of accurate prediction of the prices of agricultural commodities is indispensable for various reasons, as mentioned in section 1. Research also finds the gap of not utilizing the state-of-the-art DL techniques in the fields of agriculture. This research attempts to fill this gap by employing deep learning-based five

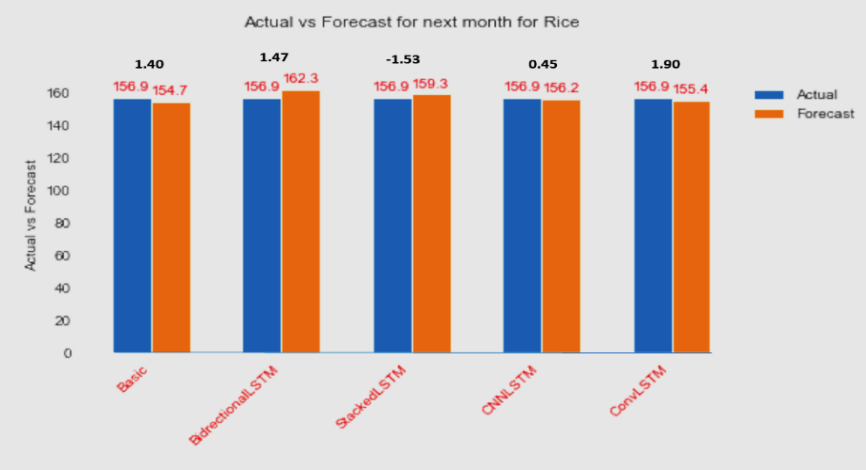


Fig 21. Actual vs. Forecast for next month for Rice

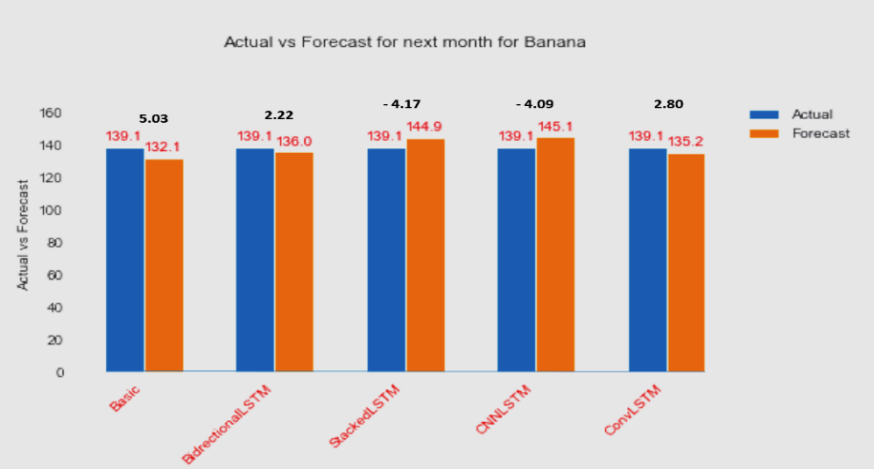


Fig 22. Actual vs. Forecast for next month for Banana

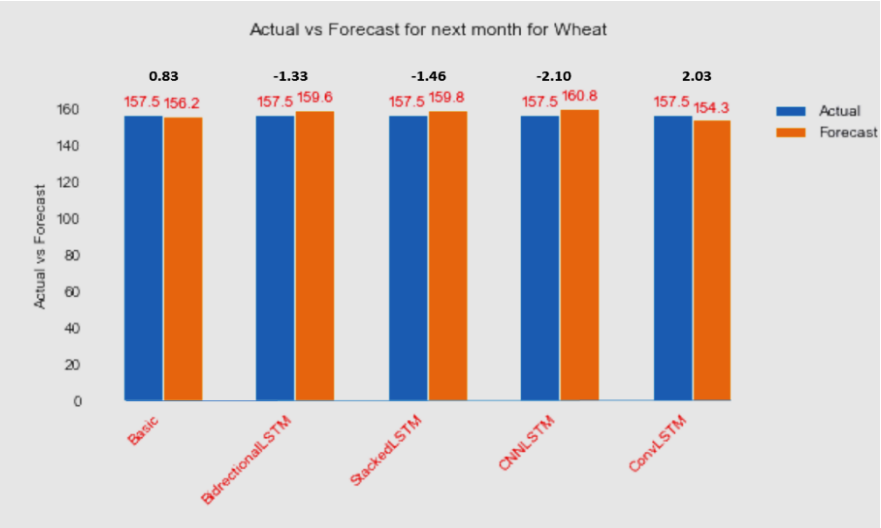


Fig 23. Actual vs. Forecast for next month for Wheat

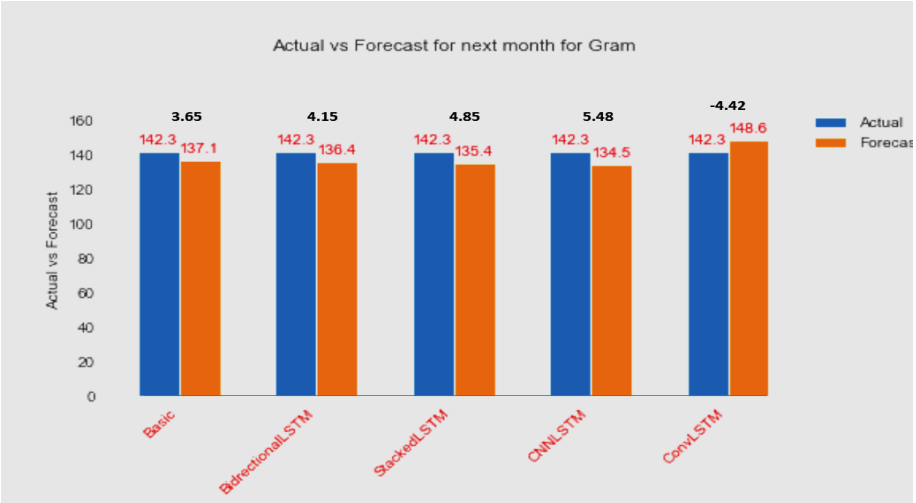


Fig 24. Actual vs. Forecast for next month for Gram

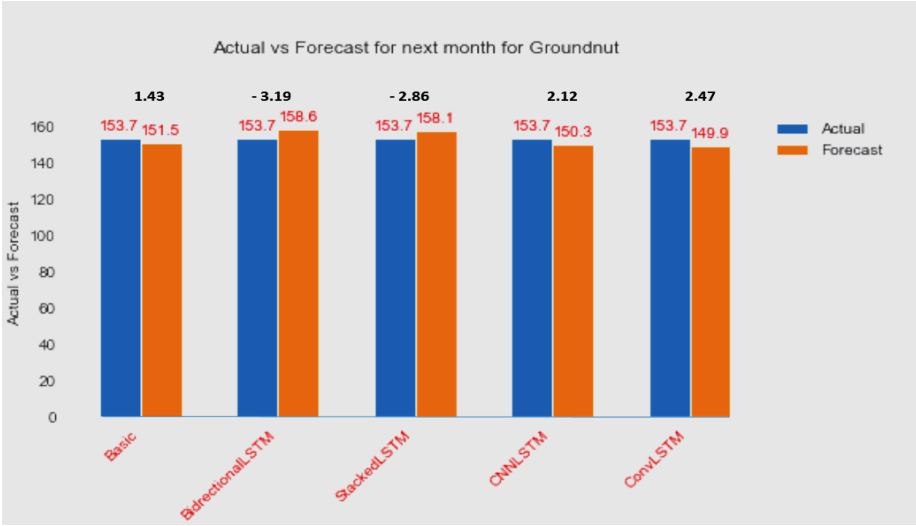


Fig 25. Actual vs. Forecast for next month for Groundnut

variants of LSTM techniques, called, 1. Basic LSTM, 2. Bidirectional LSTM, 3. Stacked LSTM, 4. CNN LSTM, and 5. Convolutional LSTM, on five agricultural commodities time series prices dataset spanning January 2000 to July 2020. After initial analysis related to statistical properties and unit root test of the datasets, each univariate commodity prices time series dataset has been split into training and testing sets for the purpose of feeding into LSTM models. The study obtained good and promising forecasting results using all the five LSTM techniques. The last ten periods result of actual vs. forecasted prices of all the commodities under consideration were close to each other which are validated by lower values of four error metrics, MAE, MAPE, MSE, and RMSE. All the five deep learning techniques have been tested statistically with a paired t-test, hypothesis, and confidence level of 95%. The t-test confirmed the performance of all the five models in the prediction tasks of agricultural commodities prices. In line with validation of actual vs. predicted values of agricultural

commodities price through lower values of four error metrics, the t-test result further confirmed the good agricultural price forecast employing these five LSTM techniques. The study also forecasted one month ahead future prices of all five agricultural commodities under consideration and compared the predicted prices with actual prices, and found that results are very promising. Hence this clearly observes that agricultural commodities prices can be forecasted more accurately using these five-deep learning-based LSTM models and provide valid and useful information to peasants, the public, the Government, and traders. In the future, adding more number of agricultural commodities around the world and comparing the results with popular machine learning techniques could provide more insights in the tasks of prediction of agricultural commodities prices.

Abbreviations

ADF - Augmented Dickey Fuller
AI - Artificial Intelligence
ANN - Artificial Neural Network
API - Application Programming Interface
ARIMA - Autoregressive Integrated Moving Average
ARIMAX - Autoregressive Integrated Moving Average with Explanatory Variable
BI LSTM - Bidirectional Long Short-Term Memory
BP Neural network - Back-propagation Neural Network
CNN- Convolutional Neural Networks
Conv LSTM - Convolutional Long Short-Term Memory
CPU - Central Processing Unit
DL - Deep Learning
ELM - Extreme Learning Machine
FCLSTM - Fully connected Long Short-Term Memory
FG - Forget Gate
GAF - Gramian Angular Fields
GAN - Generative Adversarial Networks
GB- Gigabyte
GBM - Gradient Boosting Machine
GDP - Gross Domestic Product
GHz – GigaHertZ
GPU - Graphic Processing Unit
Hybrid ANNs- Hybrid Artificial Neural Network
IG - Input Gate
KNN - k-Nearest Neighbours
LSTM - Long Short-Term Memory
MAE - Mean Absolute Error
MAPE - Mean Absolute Percentage Error
ML - Machine Learning
MLP - Mmultilayer Perceptron Neural Network
MSE - Mean Squared Error
MSP - Minimum Support Price
MTF- Markov Transition Fields
NN - Neural Network
OG - Output Gate
PP - Philips Pherron
R^2 - R squared
RAM - Random-access Memory
RBI - Reserve Bank of India
RF - Random Forest
RMSE - Root Mean Square Error
RNN - Recurrent Neural Networks
SARIMA - Seasonal Autoregressive Integrated Moving Average
SVM - Support Vector Machine
SVR - Support Vector Regressor
TDNN - Time-Delay Neural Network
TSF- Time Series Forecast
UG - Update Gate
XGB - Extreme Gradient Boosting Machine

Ethical Approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Funding details

Not applicable

Conflict of interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Informed Consent

Not applicable

Authorship Contribution

Dr. R. Murugesan: Conceptualization, Methodology, Software, Validation, Formal analysis, Writing - original draft, software for visualization, complete review. *Eva Mishra:* Literature review, tabulation, writing, data curation, and editing. *Akash Hari Krishnan:* Software, Analysis of forecasting methodology, software implementation.

References

Anggraeni, W., Mahananto, F., Sari, A. Q., Zaini, Z., & Andri, K. B. (2019). Forecasting the Price of Indonesia's Rice Using Hybrid Artificial Neural Network and Autoregressive Integrated Moving Average (Hybrid NNs-ARIMAX) with Exogenous Variables. *Procedia Computer Science*, *161*, 677-686.

Arora, P., Kumar, H., & Panigrahi, B. K. (2020). Prediction and analysis of COVID-19 positive cases using deep learning models: A descriptive case study of India. *Chaos, Solitons & Fractals*, *139*, 110017.

Chen, W., Xu, H., Jia, L., & Gao, Y. (2021). Machine learning model for Bitcoin exchange rate prediction using economic and technology determinants. *International Journal of Forecasting*, *37*(1), 28-43.

Chen, Z., Goh, H. S., Sin, K. L., Lim, K., Chung, N. K. H., & Liew, X. Y. (2021). Automated Agriculture Commodity Price Prediction System with Machine Learning Techniques. *arXiv preprint arXiv:2106.12747*.

Chimmula, V. K. R., & Zhang, L. (2020). Time series forecasting of COVID-19 transmission in Canada using LSTM networks. *Chaos, Solitons & Fractals*, *135*, 109864.

Chuluunsaikhan, T., Ryu, G., Yoo, K. H., Rah, H., & Nasridinov, A. (2020). Incorporating Deep Learning and News Topic Modeling for Forecasting Pork Prices: The Case of South Korea. *Agriculture*, *10*(11), 513.

Dairi, A., Harrou, F., Zeroual, A., Hittawe, M. M., & Sun, Y. (2021). Comparative study of machine learning methods for COVID-19 transmission forecasting. *Journal of Biomedical Informatics*, *118*, 103791.

Deepa, S., Alli, A., & Gokila, S. (2021). Machine learning regression model for material synthesis prices prediction in agriculture. *Materials Today: Proceedings*.

Elsheikh, A., Yacout, S., & Ouali, M. S. (2019). Bidirectional handshaking LSTM for remaining useful life prediction. *Neurocomputing*, *323*, 148-156.

Faniband, Y. P., & Shaahid, S. M. (2021). Univariate Time Series Prediction of Wind speed with a case study of Yanbu, Saudi Arabia. *International Journal*, *10*(1).

Gautam, Y. (2021). Transfer Learning for COVID-19 cases and deaths forecast using LSTM network. *ISA transactions*.

Graves, A., & Schmidhuber, J. (2005). Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural networks*, *18*(5-6), 602-610.

Hansen, J. V., McDonald, J. B., & Nelson, R. D. (1999). Time series prediction with Genetic-Algorithm designed neural networks: An empirical comparison with modern statistical models. *Computational Intelligence*, *15*(3), 171-184.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*(8), 1735-1780.

Jadhav, V., CHINNAPPA, R. B., & Gaddi, G. M. (2017). Application of ARIMA model for forecasting agricultural prices. *Journal of Agricultural Science and Technology*, 19(4), 981-992

Ji, L., Zou, Y., He, K., & Zhu, B. (2019). Carbon futures price forecasting based with ARIMA-CNN-LSTM model. *Procedia Computer Science*, *162*, 33-38.

Jin, D., Yin, H., Gu, Y., & Yoo, S. J. (2019, November). Forecasting of Vegetable Prices using STL-LSTM Method. In *2019 6th International Conference on Systems and Informatics (ICSAI)* (pp. 866-871). IEEE.

Kamdem, J. S., Essomba, R. B., & Berinyuy, J. N. (2020). Deep learning models for forecasting and analyzing the implications of COVID-19 spread on some commodities markets volatilities. *Chaos, Solitons & Fractals*, *140*, 110215.

Kulshrestha, A., Krishnaswamy, V., & Sharma, M. (2020). Bayesian BILSTM approach for tourism demand forecasting. *Annals of tourism research*, *83*, 102925.

Kurumatani, K. (2020). Time series forecasting of agricultural product prices based on recurrent neural networks and its evaluation method. *SN Applied Sciences*, *2*(8), 1-17.

Livieris, I. E., Pintelas, E., & Pintelas, P. (2020). A CNN–LSTM model for gold price time-series forecasting. *Neural computing and applications*, *32*(23), 17351-17360.

Ly, R., Traore, F., & Dia, K. (2021). *Forecasting commodity prices using long-short-term memory neural networks* (Vol. 2000). Intl Food Policy Res Inst.

Mahto, A. K., Alam, M. A., Biswas, R., Ahmad, J., & Alam, S. I. (2021). Short-Term Forecasting of Agriculture Commodities in Context of Indian Market for Sustainable Agriculture by Using the Artificial Neural Network. *Journal of Food Quality*, *2021*.

Maione, C., & Barbosa, R. M. (2019). Recent applications of multivariate data analysis methods in the authentication of rice and the most analyzed parameters: A review. *Critical reviews in food science and nutrition*, *59*(12), 1868-1879.

Maity, B., & Chatterjee, B. (2012). Forecasting GDP growth rates of India: An empirical study. *International Journal of Economics and Management Sciences*, *1*(9), 52-58.

Manogna, R. L. (2020). Innovation and firm growth in agricultural inputs industry: empirical evidence from India. *Journal of Agribusiness in Developing and Emerging Economies*.

Nassar, L., Okwuchi, I. E., Saad, M., Karray, F., & Ponnambalam, K. (2020, July). Deep learning based approach for fresh produce market price prediction. In *2020 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-7). IEEE.

Ouyang, H., Wei, X., & Wu, Q. (2019). Agricultural commodity futures prices prediction via long-and short-term time series network. *Journal of Applied Economics*, *22*(1), 468-483.

Paroissien, E. (2020). Forecasting bulk prices of Bordeaux wines using leading indicators. *International Journal of Forecasting*, *36*(2), 292-309.

Peng, L., Zhu, Q., Lv, S. X., & Wang, L. (2020). Effective long short-term memory with fruit fly optimization algorithm for time series forecasting. Soft Computing, 1-21.

Prakash, N., & Farzana, G. S. (2019, December). Short Term Price Forecasting of Horticultural Crops Using Long Short Term Memory Neural Network. In *International Conference*

On Computational And Bio Engineering (pp. 111-118). Springer, Cham.
Rakhra, M., Soniya, P., Tanwar, D., Singh, P., Bordoloi, D., Agarwal, P., & Verma, N. (2021). Crop Price Prediction Using Random Forest and Decision Tree Regression:-A Review. *Materials Today: Proceedings*.
Rasheed, A., Younis, M. S., Ahmad, F., Qadir, J., & Kashif, M. (2021). District Wise Price Forecasting of Wheat in Pakistan using Deep Learning. *preprint2103.04781*
Ribeiro, M. H. D. M., & dos Santos Coelho, L. (2020). Ensemble approach based on bagging, boosting and stacking for short-term prediction in agribusiness time series. *Applied Soft Computing*, *86*, 105837.
RL, M., & Mishra, A. K. (2021). Forecasting spot prices of agricultural commodities in India: Application of deep-learning models. Intelligent Systems in Accounting, Finance and Management, 28(1), 72-83.
Sabu, K. M., & Kumar, T. M. (2020). Predictive analytics in Agriculture: Forecasting prices of Arecanuts in Kerala. *Procedia Computer Science*, *171*, 699-708.
Sagheer, A., & Kotb, M. (2019). Time series forecasting of petroleum production using deep LSTM recurrent networks. *Neurocomputing*, *323*, 203-213.
Sezer, O. B., Gudelek, M. U., & Ozbayoglu, A. M. (2020). Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. *Applied Soft Computing*, *90*, 106181.
Shahid, F., Zameer, A., & Muneeb, M. (2020). Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM. *Chaos, Solitons & Fractals*, *140*, 110212.
Shao, Y. E., & Dai, J. T. (2018). Integrated feature selection of ARIMA with computational intelligence approaches for food crop price prediction. *Complexity*, 2018.
Shastri, S., Singh, K., Kumar, S., Kour, P., & Mansotra, V. (2020). Time series forecasting of Covid-19 using deep learning models: India-USA comparative case study. *Chaos, Solitons & Fractals*, *140*, 110227.
Shiferaw, Y. A. (2019). Time-varying correlation between agricultural commodity and energy price dynamics with Bayesian multivariate DCC-GARCH models. *Physica A: Statistical Mechanics and its Applications*, *526*, 120807.
Shin, D., Ha, E., Kim, T., & Kim, C. (2021). Short-term photovoltaic power generation predicting by input/output structure of weather forecast using deep learning. Soft Computing, 25(1), 771-783.
Vidal, A., & Kristjanpoller, W. (2020). Gold volatility prediction using a CNN-LSTM approach. *Expert Systems with Applications*, *157*, 113481.
Wang, S., Mu, L., & Liu, D. (2021). A hybrid approach for El Niño prediction based on Empirical Mode Decomposition and convolutional LSTM Encoder-Decoder. *Computers & Geosciences*, *149*, 104695.
Weng, Y., Wang, X., Hua, J., Wang, H., Kang, M., & Wang, F. Y. (2019). Forecasting horticultural products price using ARIMA model and neural network based on a large-scale data set collected by web crawler. *IEEE Transactions on Computational Social Systems*, *6*(3), 547-553.
Zeroual, A., Harrou, F., Dairi, A., & Sun, Y. (2020). Deep learning methods for forecasting COVID-19 time-Series data: A Comparative study. *Chaos, Solitons & Fractals*, *140*, 110121.
Zhang, D., Chen, S., Liwen, L., & Xia, Q. (2020). Forecasting agricultural commodity prices using model selection framework with time series features and forecast horizons. *IEEE Access*, *8*, 28197-28209.
Zhang, L., Liu, P., Zhao, L., Wang, G., Zhang, W., & Liu, J. (2021). Air quality predictions with a semi-supervised bidirectional LSTM neural network. *Atmospheric Pollution Research*, *12*(1), 328-339.

Zhu, N., Liu, X., Liu, Z., Hu, K., Wang, Y., Tan, J., & Guo, Y. (2018). Deep learning for smart agriculture: Concepts, tools, applications, and opportunities. *International Journal of Agricultural and Biological Engineering*, *11*(4), 32-44.