

An Enhanced Deep Learning Approach in Forecasting Banana Harvest Yields

Mariannie A Rebortera¹

Graduate Programs
Technological Institute of the Philippines
Quezon City, Philippines

Arnel C Fajardo²

Dean, School of Graduate Studies
Manuel L. Quezon University
Diliman, Quezon City Philippines

Abstract—This technical quest aspired to build deep multifaceted system proficient in forecasting banana harvest yields essential for extensive planning for a sustainable production in the agriculture sector. Recently, deep-learning (DL) approach has been used as a new alternative model in forecasting. In this paper, the enhanced DL approach incorporates multiple long short term memory (LSTM) layers employed with multiple neurons in each layer, fully trained and built a state for forecasting. The enhanced model used the banana harvest yield data from agrarian reform beneficiary (ARB) cooperative of Dapco in Davao del Norte, Philippines. The model parameters such as *epoch*, *batch size* and *neurons* underwent tuning to identify its optimal values to be used in the experiments. Additionally, the root-mean-squared error (RMSE) is used to evaluate the performance of the model. Using the same set of training and testing data, experiment exhibits that the enhanced model achieved the optimal result of 34.805 in terms of RMSE. This means that the enhanced model outperforms the single and multiple LSTM layer with 43.5 percent and 44.95 percent reduction in error rates, respectively. Since there is no proof that LSTM recurrent neural network has been used with the same agricultural problem domain, therefore, there is no standard available with regards to the level of error reduction in the forecast. Moreover, investigating the performance of the model using diverse datasets specifically with multiple input features (multivariate) is suggested for exploration. Furthermore, extending and embedding this approach to a web-based along with a handy application is the future plan for the benefit of the medium scale banana growers of the region for efficient and effective decision making and advance planning.

Keywords—Yield forecasting; Deep Learning; Long short-term memory; Banana harvest yield forecasting

I. INTRODUCTION

Deep learning (DL) is a method that has been enticing attention in recent years of machine learning and its continuous growth gains more popular among researchers in diverse disciplines [1] where advancement and progression are fast and incremental. Frequently, development takes place in a well-resourced area (e.g. medical, security) and this budding application of DL is dispersed to the agricultural sector [2]. In agriculture, production is an essential phenomenon in natural aspect where progressions fuse in multifaceted ways and production patterns are specifically influence by market factors and relentlessly affected by extreme events (i.e. floods, droughts) and revealed to deteriorate and decreases its yield. Thus, management technologies and timely interventions

should be put in place. Otherwise, it would remain unmitigated or even intensified and could suffer shortfalls and will continue to exert pressure on agricultural produce.

Yields from crops play a noteworthy role in the economic progression. Among the major fruit crops, banana (*Musa sp.*) is one of the important tropical fruit crops and part of the rising economy of many developing countries like Philippines and the world's most important goods following rice, wheat and maize. Hence, yield assessment of banana production is essential for policy decisions regarding procurement, distribution, buffer stocking, import-export, price-fixation and marketing [3]. In view of that, more accurate forecasts of the harvest yields and crop production provides an aid to an effective and efficient decision making using timely information. It is a significant phase for an emerging economy so that adequate planning is undertaken for sustainable growth [4] and for the overall development of the country. However, studies have shown that agricultural problems like forecasting yields remain difficult due to the lack of the necessary infrastructures and there is no proof of optimal model to handle time series (TS) data to be used in forecasting such as the banana harvest yields dataset.

Previously, different conventional models such as autoregressive integrated moving average (ARIMA) [5], [6], [7], [8] are used in forecasting more specifically using TS data. However, a major drawback in its used in forecasting is its incapability to recognized nonlinearities [9], [10], [4]. These classical techniques have been replaced by DL algorithms [11]. DL approaches are capable of identifying non-linearity and complexity of data [12], [8] in TS forecasting. Hence, this advance approach is referring to as a future promising tool [13] in forecasting yields in the field of agriculture.

The more recent structures of DL are Deep Neural Networks, Convolutional Neural Networks, Recurrent Neural Networks (RNN), and Q-learning [11]. Among these DL architecture, RNN model presents elevated performance in prediction, as it can capture the time features and the architectures demonstrate dynamic temporal behavior [14]. However, the training is difficult and the major weakness of RNN is carried out during the requirement of learning long-range time dependencies [15]. This limitation is addressed through the development of LSTM algorithm [16], [17], [18]. LSTM is developed to seize the “vanishing gradient problem” encountered in RNN. It is also capable of learning long term dependences [19], [10]. It is the state-of-the-art technique for

sequence learning [20], [21], [16] and TS prediction such as in financial market [22], hydrology [19], petroleum production [23], energy [24], [25], [26], [27], neurocomputing [28], [24], [18], [17], expert systems [16], internet of things [10]. However, the weakness of this technique has perceived as it did not perform satisfactorily in dealing with TS forecasting. Its shallow architecture makes it incapable to exemplify the complex characteristics of TS data more specifically in handling extremely nonlinear and long interval TS datasets [15] such as in banana harvest yields data. Furthermore, this limitation compels LSTM to be unclear if it is the best design to work out real problem especially in using the harvest yield dataset and the optimization issues due to the size of the data and the model tuning strategy applied. Also, LSTM has not been used in forecasting harvest yields such as in bananas since its inception.

In this paper, the enhanced deep learning-based approach is used in forecasting harvest yields of banana production. The performance of the model is then evaluated in terms of accuracy measures. The result of this study will be a great contribution to the consistent management for the improvement of harvest yields and to the overall production. It would also provide a new technique to assist the agrarian reform beneficiary (ARB) cooperative of Dapco in its individual farming scheme, decision-making process and advance planning.

The following parts of the paper are structured as follows: Section II looks into the idea of Deep Learning, time series forecasting, application, and challenges. Section III features the fundamentals of LSTM and the enhanced model. Section IV highlights the experiments as well as the outcomes and finally, suggestions and conclusions are offered in Section V.

II. RELATED LITERATURE

A. Deep Learning Approach

DL establishes a current, modern technique for data analysis, with likely results and significant capability [29]. As DL has been effectively applied in several domains, it has recently entered also in agriculture. Moreover, the performance of the model is generally high and its potentials can also be applied in a wide variety of agricultural related problems not only involving images but also in forecasting TS data. The efficiency of its testing time is relatively faster than any other outmoded methods and the likelihood to develop simulated datasets to train model [13] is another advantage to solve real-world problems.

DL algorithm consists of various components like RNN [11] and is the most popular method in forecasting task based on the solution intricacy, the desired accuracy in prediction and features of data [15]. More approaches are adopting RNN particularly LSTM exploiting the time dimension to perform higher performance prediction [13] and more sophisticated architecture in dealing with large datasets which could improve its performance [30], [31].

B. Time Series Forecasting Applications, Challenges and Methods

Investigating time series and dynamic demonstration is an interesting exploration. Analysis of TS data aims to research the observations trail and construct a model to depict the structure of data and forecast future values. Hence, it is vital to conceptualize an applicable model aiming at improving accuracy of the forecast. There are some different domains which already tested the capacity and adopted the used of LSTM in TS forecasting problems such as predicting emergency event occurrence [16] which solves classification and regression problem and exhibits better performance which is proven effective over conventional methods. In forecasting petroleum production, [15] which case study involves the production of two separate oil depot at a particular time period, proves the capability and eligibility of LSTM to be applied in the nonlinear forecasting problems and outperforms outmoded and traditional mathematical forecasting models. In predicting traffic flow [24], where it involves short term traffic flow at a time interval of 1 to 5 minutes, LSTM and Attention Mechanism shows excellent performance in dealing with 5- or 1-minute-long historical data. However, the performance of the model declines swiftly as the length of the sequence increases because of the collection of errors: the longer the sequence in LSTM, the greater the error. Thus, it considers time and space features combination. Otherwise, in predicting the remaining useful life of proton exchange membrane fuel cell [18], LSTM model quickly and accurately predicts the remaining service life and suitable for online residual life prediction but the robustness and generalizations performance need to be further strengthened and improved. It also outperforms the outmoded techniques in predicting hourly day-ahead solar irradiance [25] though error is encountered when using weather forecast but it shows less overfitting and better generalization. Thus, further evaluation and assessment is needed.

Moreover, it also shows simplicity and effectivity than ARIMA and back propagation neural network in the tourism flow prediction [10]. It suggests further that more hidden states are to be tested and superiority over the classic feed forward neural network and the double-LSTM models in predicting water table depth [19]. A dropout strategy is being implemented to avoid overfitting though it has a resilience to learn TS data but has insufficiency in its fitting ability. In financial market predictions, it is effective in extracting meaningful information from a noisy financial TS [22] compared to random forest, standard deep nets and logistic regression and it turns out to be an advancement of the domain with respect to prediction accuracy and daily returns after transaction cost. It also outperforms conventional techniques in short-term load forecasts [26] and exhibits consistency for the ambiguity with snowball in hours to forecast. LSTM proves appropriateness in TS modelling and forecasting with reduction in percentage errors and shows simplicity towards managing the information instead of working over complex equations.

III. THE ENHANCED APPROACH

This segment depicts the enhanced deep learning approach used in forecasting banana harvest yields from data preparation to evaluating the forecast. Also, important steps are discussed and described in this section.

A. Data

The banana harvest yield data set features the number of stems cut, the number of boxes produce and box-stem ratio. The author chose the “number of stems cut” as the harvest yields time series data to be used during the experiment. In this study, the term “stems” is referring to the whole bunch of the banana fruit cut from the plant. The harvest data came from the ARB cooperative of Dapco in Davao del Norte, Philippines. The cooperative is one of the key players from the small-medium scale banana growers in the region and has contributed to the overall production on the exportation of bananas to other countries like Japan. The dataset contains thirty-five thousand series of observations approximately from year 2014 to 2018 where each year is composed of thirteen (13) periods and it usually starts from second half of the first month of the year to the first half of the following year.

B. Data Preprocessing

Series of data transformations are done before fitting the model to the dataset and making a forecast. It includes converting the TS data into supervised learning to make it stationary. A lag differencing is used to strip off the increasing trend in the data. Transformation of data into a supervised learning problem and scaling to values to meet the hyperbolic tangent activation function of the model is also done. All these transforms will be inverted back on forecast to revert data into its original scale before evaluation and in determining the error score and splitting data into training and testing sets respectively. Aiming to obtain the best outcomes, iterative optimization is used which means attaining the outcomes several times and select the utmost optimal iteration that has less errors. Each important parameter, such as epoch, batch size and neurons, is given a varied value and the experiment for each parameter are run in several repetitions as desired. The best value for each parameter is identified through the summary of performance report using the RMSE scores from each population of results.

C. The Multi-Dynamic Long Short Term Memory (mdLSTM) Model

It is indispensable to briefly explain the fundamentals of LSTM memory block as it is the precedent of the enhanced model prior to its introduction.

1) Fundamentals of the Long Short Term Memory (LSTM) Model.

LSTM was first proposed in 1997 (Sepp Hochreiter, Jürgen Schmidhuber), driven by an analysis of error flow in prevailing RNNs (Hochreiter et al., 2001) [16]. It shows suitability for processing and forecasting using TS data. The LSTM block, shown in Fig. 1, depicts a cell state (C_t) which resembles a conveyor belt. It thoroughly takes turn in the chain and controlled by constitution of gates; an elective inlet means of information. Gates are comprised of a sigmoid neural net layer

and a *pointwise multiplication* and *addition* operation. These gates and the memory cell allow an LSTM unit to respectively *forget*, *memorize* and *expose* [14] the memory content.

It has an input at time step t denoted by (x_t), and the hidden state from the previous time step (S_{t-1}) that is introduced to LSTM block, and then the hidden state (S_t) is computed through the forget gate (f_t), input gate (i_t) and output gate (o_t) where the *input and forget gates* are responsible of how much new content should be remembered (*memorized*) and how much old content should be disregarded (*forgotten*). Gates are computed using the following set of formulas arrange in steps:

Step 1: Determine what content is going to be disregarded from the *cell state* which will be decided by the *forget gate* (f_t):

$$f_t = \text{sigmoid}(W_{xf}x_t + W_{sh}h_{t-1} + b_f) \quad (1)$$

Step 2: Determine which new content is going to be kept in the *cell state* which will be decided in two phases: First, the *input gate* (i_t) layer decides which values to be updated. Second, a *tanh* layer that forms a vector of new candidate values (\hat{C}_t). These two phases can be illustrated as follows:

$$i_t = \text{sigmoid}(W_{xi}x_t + W_{si}h_{t-1} + b_i) \quad (2)$$

$$\hat{C}_t = \tanh(W_{ci}x_t + W_{ci}h_{t-1} + b_c) \quad (3)$$

Step 3: Update the previous *cell state* (C_{t-1}) into the new *cell state* (C_t), which can be conveyed as:

$$C_t = f_t * C_{t-1} + i_t * \hat{C}_t \quad (4)$$

Step 4: Lastly, determine the desired output to be produced. The output will be a streamed form and will be based on the cell state. The output gate (o_t) in this step will decide what part of the cell state is going to be produced as output and then goes through the *tanh* layer, impelling values to be between -1 and 1, and multiplying it to the *output gate* as illustrated in the equation below:

$$o_t = \text{sigmoid}(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \quad (5)$$

$$S_t = o_t * \tanh(C_t) \quad (6)$$

The LSTM is represented with the two sets of parameters from the preceding six equations. These are: W_{xf} , W_{sf} , W_{xi} , W_{si} , W_{ci} , W_{xo} , W_{ho} which are referred to as weights and b_f , b_i , b_c and b_o are biases, respectively.

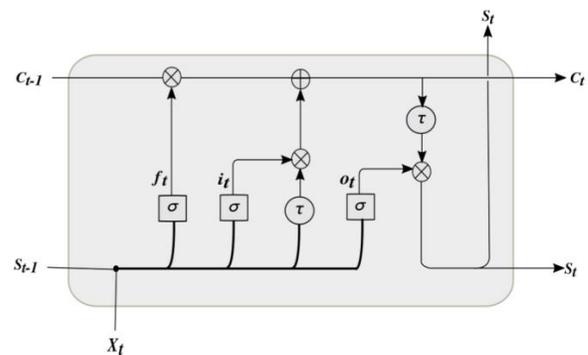


Fig. 1. The LSTM Memory Block, where \oplus is the Pointwise Addition, \otimes is the Pointwise Multiplication, τ Hyperbolic Tangent Activation Function and σ is the Sigmoid Activation.

2) The multi-dynamic Long Short-Term Memory (mdLSTM)

An effective means to a better overall performance of the neural network is to augment its deepness [11]. The development of expound LSTM recurrent neural network is encouraged by the remarkable learning capabilities of profound recurrent network design to be utilized in TS forecasting applications. The enhanced model, shown in Fig. 2, has several LSTM layers, heaped one after the other joined to blend the advantage of a sole LSTM layer in an expound recurrent network manner feed with multiple value of neurons. The aim of the enhanced model is to construct the characteristics in a hierarchical design where the lower layer separates the input data disparity factors and these demonstrations are merged at the upper layer. Such deep structure will simplify well owing to a trimmed representation than a shallow design in case of large or complex datasets [32], [33] such as in the banana harvest yields.

D. Training and Forecasting

To execute the model, Keras library along with Theano and Tensorflow backend are properly installed and configured, splitting of dataset into 80% training and 20% testing is done respectively to capture the thirteen (13) periods of the year 2018 as test dataset. The number of epochs, batch size and neurons are assigned a minimal value from the summary result of the performance done in the tuning step. The “mean squared error” loss function with “ADAM” optimization algorithm are used in compiling the model. RMSE [34], [26], [22], [8] is used to evaluate the performance of the model to forecast. It calculates the variance between actual value and the predicted value and used to evaluate different models for a certain data and not amid datasets. The following formula is used for computing the RMSE:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i^{obs} - y_i^{pred})^2} \tag{7}$$

This measure is calculated by associating the target values for the time series and its consequent time series predictions, where n is the total number of observations, y_i^{obs} is the actual value; whereas, y_i^{pred} is the predicted value.

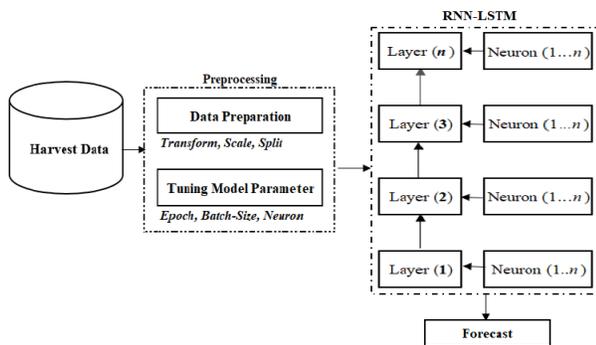


Fig. 2. The Flowchart of the Enhanced Model.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The enhanced deep learning model was implemented using the banana harvest yields dataset. Given the nature of the dataset, long interval time and the missing (zero) observations are seen and considered. In view of the fact that missing values were observed, simply removing records containing zero value is the simplest strategy ever devised during the experiment with the goal of obtaining the optimal outcome of the enhanced model.

For consistency and fair assessment, splitting of data to 80% training and 20% for testing is done. Moreover, the model uses the optimal outcome from the lowest mean RMSE of the model parameter tuning for the values to be assigned for *epochs*, *batch size* and *neurons*. Noticeably, the number of epoch is very minimal because there is no proof that suggests the exact number of cycles (epochs) to train a model. Experiment wise, it is evident that setting the value of epoch and neurons minimally produces a sensible model outcome in terms of reduction of errors and forecasting accuracy most likely if using large datasets. The size of the diverse dataset is considerably important in setting the number of epochs because it shows different behavior to optimally train the network.

During the experiment, the model uses the optimal value of the parameters such as the number of epochs, batch size and neurons with respect to the result obtains from the model parameter tuning. Also, it has been noticed that executing the model more than once with the same parameter values does not guarantee better performance and at some instances, worsen the result. There are three sets of experiment done in forecasting using banana harvest yield dataset: *first*, using single LSTM layer, *second*, using multiple LSTM layers assigned with same value of neurons in each layer and *third*, using the enhanced deep learning model where multiple layers feed with multiple value of neurons. The latter used the precedent numbers of the optimal value of neuron obtained from the model parameter tuning. The results of experiments are exhibited in Table I, Table II and Table III, respectively. For uniformity, all experiments were done with and without a dropout rate. The second and third experiment used two up to four LSTM layers only.

TABLE. I. RESULTS OBTAIN USING SINGLE LSTM LAYER

LSTM Layer	Parameters			RMSE	
	Epoch	Batch Size	Neurons	No Dropout	With Dropout
1	4	1	5	61.602	68.331

TABLE. II. RESULTS OBTAIN USING MULTIPLE LSTM LAYERS

LSTM Layers	Parameters			RMSE	
	Epoch	Batch Size	Neurons	No Dropout	With Dropout
2	4	1	5, 5	64.025	72.502
3	4	1	5, 5, 5	63.748	63.362
4	4	1	5, 5, 5, 5	63.225	68.702

TABLE. III. RESULTS OBTAIN USING THE ENHANCED APPROACH

LSTM Layers	Parameters			RMSE	
	Epoch	Batch Size	Neurons	No Dropout	With Dropout
2	4	1	1,2,3,4,5	34.805	46.867
2	4	1	1,2,3,4	36.670	37.020
2	4	1	1,2,3	44.606	47.688
2	4	1	1,2	35.030	36.909
3	4	1	1,2,3	37.947	36.670
3	4	1	1,2,3,4	47.498	49.837
3	4	1	1,2,3,4,5	36.371	45.562
4	4	1	1,2,3,4,5	50.104	50.085
4	4	1	1,2,3,4	39.290	36.771

Results showed that all the experiments convey better performances in the absence of dropout rate. The optimal result of the second experiment is composed of four LSTM layers using the same value of neurons with 63.225 in terms of RMSE. The third experiment performs best with 34.805 (in terms of RMSE) in the composition of two LSTM layers with the different minimal value of neurons. The enhanced deep learning approach outperforms the single LSTM layer and multiple LSTM layers with 43.5% and 44.95% reduction of error rates, respectively. Therefore, the deep structure of LSTM is proficient in forecasting highly nonlinear and long interval time data such as in banana harvest yields dataset. The result of the performance of the enhanced approach is considered significant since there is no standard (reduction of errors) available and no proof that it has been used in forecasting harvest yields of any agricultural crop such as bananas. Furthermore, the graphical representation of the best results of the performance obtained from each experiment is shown in the following figures: Fig. 3 represents the graphical plot of the single LSTM layer; Fig. 4 represents the multiple LSTM layers, while Fig. 5 is the graphical representation of the enhanced approach.

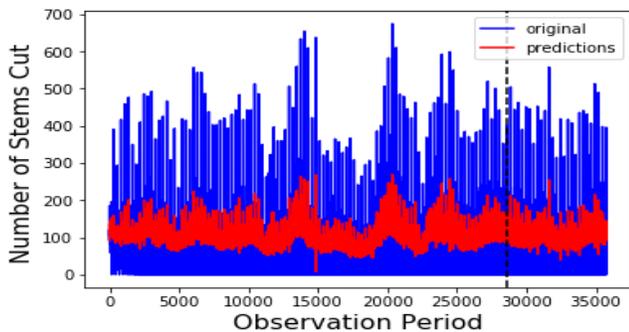


Fig. 3. Graphical Representation of the Optimal Result using the Single LSTM Layer.

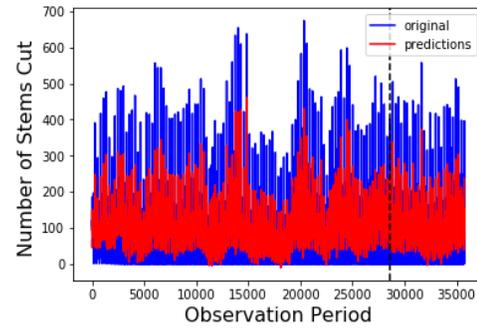


Fig. 4. Graphical Representation of the Optimal Result using Multiple LSTM Layer.

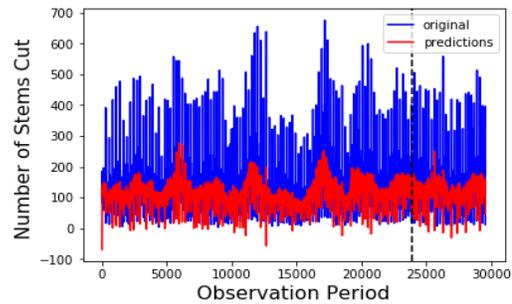


Fig. 5. Graphical Representation of the Optimal Result of the Enhance Approach.

V. CONCLUSION

This technical quest aspired to build deep multifaceted system proficient in forecasting banana harvest yields. In this paper, the enhanced model was trained from raw observations. The model achieved better performance with 43.5% and 44.95% reduction of error rates compared to single LSTM and multiple LSTM layers, respectively. It has been found out that applying dropout strategy and employing larger number of neurons does not guarantee optimal results on the datasets described in Section III-A. This outcome is a contribution to the unceasing growth of research in the agriculture domain since it has not been used for forecasting yields of crop such as bananas. It is also significant to the extensive planning for sustainability of ARB cooperatives subsidizing the over-all production. For future endeavors, it is suggested to investigate the performance of the model in diverse datasets specifically in handling multiple input features (multivariate) datasets. Furthermore, with the aim of providing help and support to the ARB cooperatives-IFS strategy and officers to the efficient and extensive planning for production sustainability, embedding the model to a web-based along with a handy application for recording and monitoring the interventions and factors influencing yields are well-thought-out for future work.

ACKNOWLEDGMENT

The author of this paper would like to thank the Agrarian Reform Beneficiaries (ARB) cooperative of Dapco in Davao del Norte, Philippines for the support and for providing the relevant data for this study.

REFERENCES

- [1] “Kentarō Kuwata and Ryosuke Shibasaki The University of Tokyo IIS , The University of Tokyo , 4-6-1 Komaba , Meguro-ku , Tokyo 153-8505 , JAPAN,” pp. 4–6, 2015.
- [2] A. Koirala, K. B. Walsh, Z. Wang, and C. McCarthy, “Deep learning – Method overview and review of use for fruit detection and yield estimation,” *Comput. Electron. Agric.*, vol. 162, no. January, pp. 219–234, 2019.
- [3] S. Nagini, “Agriculture Yield Prediction Using Predictive Analytic Techniques,” pp. 783–788, 2016.
- [4] S. Rathod and K. N. Singh, “Hybrid Time Series Models for Forecasting Banana Production in Karnataka Hybrid Time Series Models for Forecasting Banana Production in Karnataka State , India,” no. January 2018, 2017.
- [5] B. Garg, S. Aggarwal, and J. Sokhal, “model R,” *Comput. Electr. Eng.*, vol. 0, pp. 1–21, 2017.
- [6] D. Elavarasan, D. Raj, V. Sharma, and A. Y. Zomaya, “Forecasting yield by integrating agrarian factors and machine learning models : A survey,” *Comput. Electron. Agric.*, vol. 155, no. August, pp. 257–282, 2018.
- [7] P. Surya and I. L. Aroquiara, “Crop Yield Prediction in Agriculture using Data Mining Predictive Analytic Techniques,” vol. 5, no. 4, pp. 783–787, 2018.
- [8] S. Siami-namini and N. Tavakoli, “A Comparison of ARIMA and LSTM in Forecasting Time Series,” 2018 17th IEEE Int. Conf. Mach. Learn. Appl., pp. 1394–1401, 2018.
- [9] S. Rathod and G. C. Mishra, “Statistical Models for Forecasting Mango and Banana Yield of,” vol. 20, pp. 803–816, 2018.
- [10] Y. Li and H. Cao, “ScienceDirect ScienceDirect Prediction for Tourism Flow based on LSTM Neural Network Prediction for Tourism Flow based on LSTM Neural Network,” *Procedia Comput. Sci.*, vol. 129, pp. 277–283, 2018.
- [11] Y. Lecun, Y. Bengio, and G. Hinton, “Deep learning,” 2015.
- [12] R. Zuo, Y. Xiong, J. Wang, and E. J. M, “PT NU SC,” *Earth-Science Rev.*, 2019.
- [13] A. Kamilaris and F. X. Prenafeta-boldú, “Deep learning in agriculture : A survey,” vol. 147, no. July 2017, pp. 70–90, 2018.
- [14] J. Chung, “Gated Feedback Recurrent Neural Networks,” vol. 37, 2015.
- [15] A. Sagheer and M. Kotb, “Neurocomputing Time series forecasting of petroleum production using deep LSTM recurrent networks,” *Neurocomputing*, vol. 323, pp. 203–213, 2019.
- [16] B. Cortez, B. Carrera, Y. Kim, and J. Jung, “PT US CR,” *Expert Syst. Appl.*, 2017.
- [17] Y. Rizk and M. Awad, “Neurocomputing On extreme learning machines in sequential and time series prediction : A non-iterative and approximate training algorithm for recurrent neural networks,” *Neurocomputing*, vol. 325, pp. 1–19, 2019.
- [18] A. Elsheikh, S. Yacout, and M. Ouali, “PT US CR,” *Neurocomputing*, 2018.
- [19] J. Zhang, Y. Zhu, X. Zhang, M. Ye, and J. Yang, “Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas,” *J. Hydrol.*, vol. 561, no. April, pp. 918–929, 2018.
- [20] K. Greff, R. K. Srivastava, J. Koutn, and B. R. Steunebrink, “LSTM : A Search Space Odyssey,” pp. 1–12.
- [21] K. Yao, T. Cohn, K. Vylomova, and C. Dyer, “Depth-Gated Recurrent Neural Networks,” pp. 1–5.
- [22] T. Fischer and C. Krauss, “Deep learning with long short-term memory networks for financial market predictions,” *Eur. J. Oper. Res.*, vol. 270, no. 2, pp. 654–669, 2018.
- [23] A. Sagheer and M. Kotb, “AC US CR,” *Neurocomputing*, 2018.
- [24] B. Yang, S. Sun, J. Li, X. Lin, and Y. Tian, “US CR,” *Neurocomputing*, 2018.
- [25] X. Qing and Y. Niu, “Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM,” *Energy*, vol. 148, pp. 461–468, 2018.
- [26] S. Muzaffar, A. Afshari, and S. Muzaffar, “ScienceDirect ScienceDirect ScienceDirect Short-Term Load Forecasts Using LSTM Networks on District and Cooling Short-Term Load Forecasts LSTM feasibility of using the heat demand-outdoor Masdar for,” *Energy Procedia*, vol. 158, pp. 2922–2927, 2019.
- [27] Y. Qin et al., “Hybrid forecasting model based on long short term memory network and deep learning neural network for wind signal,” *Appl. Energy*, vol. 236, no. October 2018, pp. 262–272, 2019.
- [28] J. Liu, Q. Li, W. Chen, Y. Yan, Y. Qiu, and T. Cao, “ScienceDirect Remaining useful life prediction of PEMFC based on long short-term memory recurrent neural networks,” pp. 1–11, 2018.
- [29] A. Singh, B. Ganapathysubramanian, A. K. Singh, and S. Sarkar, “Machine Learning for High-Throughput Stress Phenotyping in Plants,” *Trends Plant Sci.*, vol. 21, no. 2, pp. 110–124, 2016.
- [30] R. Sensing, S. I. Sciences, and C. Vision, “Multi-Temporal Land Cover Classification with Long Short-Term Memory Neural Networks,” vol. XLII, no. June, pp. 6–9, 2017.
- [31] D. Ienco, R. Gaetano, C. Dupaquier, and P. Maurel, “Land Cover Classification via Multitemporal Spatial Data by Deep Recurrent Neural Networks,” pp. 1–5, 2017.
- [32] M. Långkvist, L. Karlsson, and A. Loutfi, “A review of unsupervised feature learning and deep learning for time-series modeling q,” vol. 42, pp. 11–24, 2014.
- [33] M. Hermans and B. Schrauwen, “Training and Analyzing Deep Recurrent Neural Networks,” pp. 1–9.
- [34] W. Xu, Z. Zhang, D. Gong, and X. Guan, “Neural Network Model for the Risk Prediction in Cold Chain Logistics,” vol. 9, no. 8, pp. 111–124, 2014.