



Crop Yield Forecasting Through an Attention-Driven LSTM Deep Learning Mode

Homa Rizvi¹, Shamim Ahmad², Dr. Yusuf Perwej³, Farheen Siddiqui⁴, Dr. Nikhat Akhtar⁵

¹ Assistant Professor, Department of Computer Science & Engineering, Shri Ramswaroop Memorial University, Deva Road, Lucknow

² Assistant Professor, Department of Computer Science & Information System, Shri Ramswaroop Memorial University, Deva Road, Lucknow

³ Professor, Department of Computer Science & Engineering, Shri Ramswaroop Memorial University, Deva Road, Lucknow

⁴ Assistant Professor, Department of Computer Science & Engineering, Shri Ramswaroop Memorial University, Deva Road, Lucknow ⁵ Professor, Department of Computer Science & Engineering, Goel Institute of Technology & Management, Lucknow

Abstract: Agriculture is a vital profession globally, reliant on climatic conditions and precipitation. The goal of this article is to use climate, soil, and temperature data to estimate crop output early. This study proposes a classification-based approach for agricultural production prediction using Long Short-Term Memory (LSTM) with an Attention Mechanism. The Economics and Statistics department of the Government of Karnataka collects the manual data. This technique used data from the Department of Economics and Statistics on three crops: jowar, rice, and ragi. To fill in the missing and null values in the dataset, the linear interpolation approach is used. The feature selection procedure is useful for the Correlation based Feature Selection Algorithm (CBFA) and the Variance Inflation Factor Algorithm (VIF) because it helps them choose and remove groups of features that are linked. We use Accuracy, R2, Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE) to see how well the model works. The proposed LSTM model gives results with assessment metrics including accuracy, R2, MAE, MSE, and RMSE values of roughly 99.10%, 0.44, 0.132, and 0.233, respectively.

Keywords: Agriculture Data, Feature Selection, Long Short-Term Memory, Crop Selection, Pre-Process, Crop Yield Prediction.

1. Introduction

The Indian economy's most significant industry is agriculture. Many nations throughout the world are still having trouble managing their food supply chains because the population is growing so quickly. Agriculture has become a part of growing important food crops in our generation [1]. Jowar, paddy, and ragi are the most important food crops and the second most important in terms of output. The jowar is the main part of the Indian economy, and it is grown in both the rainy and after-rainy seasons [2]. Farmers can make better judgments about when to plant crops based on environmental considerations that will give them the best yield. The characteristics that determine crop yield output include climate, soil, temperature, biological, geographical, and other variables [3]. At every level, from local to global, farmers' decision-makers find it incredibly hard to anticipate how much of a crop will grow. Many people, such as agronomists, dealers, farmers, and policy officials, find it useful to be able to predict crop yields [4]. It is hard to reach the goal when there is a lot of crop production in a little field area. Farmers are looking at the best way to get a good yield based on the agricultural data they gather and coming up with ways to estimate crop yields that will help them learn more about rural life and farming [5]. The crop yield prediction model aids farmers in determining the optimal timing and kinds of crops to plant, considering environmental conditions to enhance production. Compared to traditional farming methods, precision agriculture is a new way of doing things that saves farmers time and money [6]. Machine Learning (ML) can learn on its own from past experiences by constantly training and giving better prediction and classification results [7]. In the past 10 years, the contemporary data-based modelling technique has been used in many areas of agriculture to make more accurate forecasts, improve efficiency, and add useful features [8]. To tell the difference between healthy and

unhealthy crops, ML image classifiers are used. The predictive model is created by utilizing several characteristics, such as model parameters that were found using historical data during the training step [9]. The ML algorithms, including Naïve Bayes (NB), Decision Tree (DT), and Random Forest (RF), exhibit both parametric and non-parametric characteristics, significantly influencing agricultural production prediction [10]. In certain cases, an Artificial Neural Network (ANN) is utilized to find and sort out crop production prediction difficulties by looking at things like CO₂ fixation, solar radiation [11], and water content. This research proposes Long Short-Term Memory (LSTM) for the early prediction of agricultural productivity [12].

2. Background

Sometimes, traditional agricultural methods aren't particularly precise or effective, which wastes time and money. For instance, if you apply insecticides and fertilizers uniformly throughout a wide field, you can end up utilizing too much or too little of these inputs, which would hurt the environment, cost more money, and give you worse outcomes [13]. Farmers have a hard time making judgments because they can't easily get reliable and up-to-date information on the weather, the health of the soil, and what people want to buy [14]. Also, not having enough money and infrastructure, particularly in rural regions, makes all of these difficulties worse. This makes it impossible for farmers to adopt the best techniques and newest technology [15]. Khaki et al. [16] used deep learning (DL) methods to predict corn and soybean yields in the Corn Belt of the United States. These methods included Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Random Forest (RF), Deep Feedforward Neural Networks (DFNN), and the Least Absolute Shrinkage and Selection Operator (LASSO). The ensemble model, which utilized data from 2016 to 2018, had a root mean square error (RMSE) of 9% for corn yields and 8% for soybean yields. Bi et al. [17] used Genetic Algorithms (GA), Neural Networks (NN), and GA-enhanced Deep Learning (DL) for crop prediction, achieving a reduction in RMSE of around 10%. Shahhosseini et al. [18] developed a hybrid model for forecasting maize production that combines crop modeling and machine learning techniques, resulting in a reduction in RMSE by 7 to 20%. Woittiez et al. [19] explored production gaps in oil palm cultivation, emphasizing the need of understanding contributing factors to improve crop management strategies and refine yield forecasts.

The findings from these researches elucidate the many machine learning algorithms used to predict agricultural productivity. This will support ongoing research that seeks to improve the precision of predictions via the use of ensemble learning techniques. Oikonomidis et al. [20] proposed a deep learning model to assess the effectiveness of machine learning algorithms based on certain criteria. Their research focused on the XGBoost algorithm and several hybrid models that included a CNN with other techniques, including DNN, RNN, and LSTM. They used these models using [21] a publicly available soybean dataset of 25,345 samples and 395 characteristics related to meteorological and edaphic conditions. Their findings suggest that forthcoming advancements may amalgamate XGBoost with deep learning methodologies, such as LSTMs or RNNs, particularly for tasks necessitating sequential data, like forecasting agricultural output. The [22] proved how effectively random forest models function by swiftly and correctly analyzing huge datasets of agricultural output. This is particularly

crucial for projecting how much food will be grown, which requires a lot of data [23]. Random forests are an example of a data mining technique that can identify hidden patterns and trends in huge datasets [24]. Data mining uncovers insights that empower organizations to make informed decisions about impending agricultural trends and conditions [25]. Hasan et al. (2023) introduced the K-nearest Neighbour Random Forest Ridge Regression (KRR) model in this work. The idea is to make accurate guesses on how much food will be produced, focusing on major crops like rice, wheat, and potatoes. This model has done better than traditional machine learning approaches. It also lets you utilize a recommender system to assist you find out which crops are optimal for improved planning and production in agriculture [26]. Boppudi (2024) introduces the Deep Ensemble Classifier Integrated Bird Swarm Butterfly Optimization Algorithm (DEC-IBSBOA) model for predicting agricultural production. This model uses the IBS-BOA approach for advanced data pre-processing, feature extraction, and finding the best features. The DEC-IBSBOA model is quite accurate, with a low MAE of around 1.0, which is better than any other method [27].

Other research projects have employed Time Series Forecasting methods, including ARIMA models, to find seasonal and temporal patterns in yield data [28]. These models use historical patterns and how they change over time to estimate what future yields will be. But these models have certain problems. For one, they rely on stable assumptions, and for another, they don't consider many factors at the same time [29]. However, the effectiveness of these models may fluctuate considerably across various crops and situations [30]. Research using UAV-based multispectral data and several machine learning algorithms for yield prediction indicates that Random Forest is the optimal model for forecasting maize yields, while Gaussian Process regression [31] is the superior model for predicting wheat and soybean yields. Support Vector Machines (SVM) have shown exceptional effectiveness in predicting broad bean yields, whereas Convolutional Neural Networks (CNN) have displayed amazing accuracy in forecasting rice yields. These discrepancies suggest that broader framework techniques might be used to address yield prediction challenges across various crops and environmental conditions. Ensemble learning (EL) [35] enhances prediction accuracy by amalgamating many foundational models using techniques like as bagging, boosting, and stacking to use their respective strengths. In several situations [36], these techniques have consistently surpassed individual models for generalization performance.

But there are a number of intricate reasons why it's impossible to anticipate agricultural output. The quality of the soil, pests, genotypes [37], the weather, the time of year, and other factors all have an effect on crop yields. Second, the methods and steps used to anticipate yield change with time and are not always linear [38]. In agricultural systems, a large portion often eludes depiction using basic stepwise calculations, particularly in scenarios where datasets are complex, incomplete, or unclear. Some of the most frequent methods that computers produce predictions include via decision trees, linear regression, and ensemble learning. Linear regression is a simple and popular machine learning method that uses a linear model to predict the relationship between crop yield and other factors that affect it. Deep learning has been widely used in agriculture [43] due to its efficacy in managing spatiotemporal data dependencies and extracting pertinent features without requiring human feature engineering [44]. Deep learning use multi-layer neural networks to extract abstract features from large datasets. These datasets might be

organized, semi-structured, or unstructured. This strategy focuses on figuring out how functional characteristics and interaction elements are related, which is important for making accurate forecasts about crop yield [45].

3. Things influencing Crop Yield Prediction

Several factors, such the weather, the kind of soil, the type of crop, and how the farm is run, affect how well you can estimate crop yields. The weather has a big effect on how much crops grow. Rain, temperature, and humidity [46] are all important for plant development. But one of the hardest things is that we can't make long-term weather predictions. Migdall says that bad weather might make it hard to harvest if the fields are too wet for machines. To fix this, she says we should make long-term projections better by using seasonal forecasts with more than one scenario [47]. The kind and fertility of the soil are also very important since they directly affect the health and development of plants. The kind of crop also affects yield forecast since various types of crops are more or less resistant to pests and diseases. Migdall said that although [48] individual farmers know what they have planted, getting information on the sorts of crops and how they grow on a broader scale, such across whole countries or continents, is very hard [49]. Also, how you manage your farm might affect how well your crops do. These include crop rotation, irrigation systems, fertilization regimens, and pest control, all of which have a big effect on how much food is produced.

3.1 Kaggle Datasets

Farming is a very important part of the world economy. The ever-increasing number of people on Earth means that we need to know a lot about global agricultural production in order to deal with food security problems and lessen the consequences of climate change. Estimating how much crops will yield is an important agricultural problem. The main things that affect agricultural production are the weather (like rain and temperature), the use of pesticides, and the availability of accurate historical crop yield data [50]. This data is very important for making smart decisions about agricultural risk management and predicting the future, as shown in figure 1.

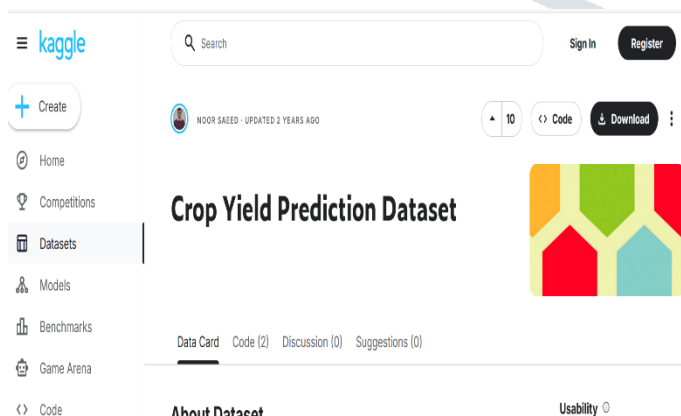


Figure 2. Kaggle Datasets for Crop Yield Prediction

4. The Suggested Model

The suggested technique involves gathering data from the yield output of three primary crops: jowar, rice, and ragi. Next, compare the three crops to see which one has the best yield projection. The main parts of this framework are dataset, pre-processing, feature selection, and LSTM [51]. Using these approaches, the attention mechanism makes it

easy to anticipate improved crop yields based on categorization. The flowchart for the suggested crop production forecast is shown in Figure 2. The most major crops grown in Karnataka are jowar, paddy, and ragi. The dataset comes from Kaggle datasets.

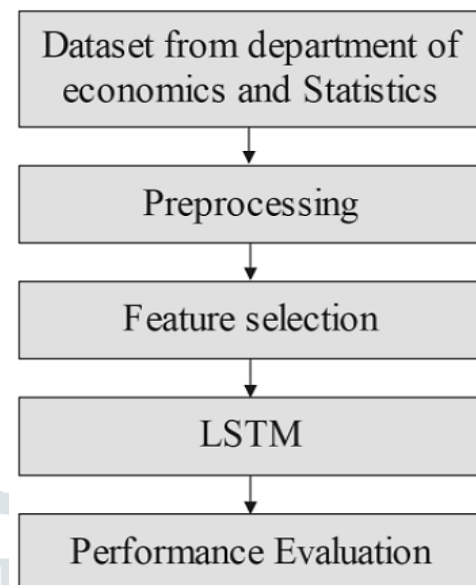


Figure 2. The LSTM Based Model for Crop Yield Prediction

After gathering the datasets, the next step is to clean up the data. Deep Learning [19] does not handle noisy data like outliers and mistakes. Before classifying the data, it is pre-processed since certain districts in Karnataka have missing or null values. This means that the undesirable data is removed and the right range of data is kept in the production row [20]. The mean values may take the place of the values in that row, and the dataset has string values that should be changed to numbers so that the data can be separated into training and testing sets. We employ the Linear Interpolation [21, 22] approach to fill in the gaps and missing data. This approach uses math to find the new data points by drawing a straight line across the current data points in the same order as the previous value. If you have time-series data and some of the values are missing, interpolation is a common way to fill in the gaps. The linear interpolation equation (1) does the following.

$$f(X) = f(x_0) + \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x - x_0)$$

In this case, x is the independent variable, x_0 and x_1 are the known values for the independent variable, and $f(x)$ is the dependent variable value for the independent variable value x . After the data has been pre-processed, the feature selection takes place. When datasets include a lot of properties, high-level feature selection is very important for getting the most accurate predicting results. The main reason for using feature selection is because the ML method makes training quicker, makes the model less complicated, and makes it easier to understand the process. This may make the model more accurate by choosing the proper subset and keeping it from overfitting. There are three ways to choose features in attribute selection: wrapper, filter, and embedding. These say that the wrapper and filter techniques are first utilized to choose the best qualities. The wrapper technique usually works better than the filter method, although the model costs a lot to run. The embedded method has both the filter and wrapper method, and it utilizes its own way of choosing the property.

4.1 Long Short-Term Memory (LSTM)

During the training, there were 11 hidden layers and 50 neurons in each layer. Long Short-Term Memory (LSTM) is the best model for predicting time series and figuring out which crop production prediction is the best. The LSTM is used to figure out which crop will provide the best yield by comparing it to the yield predictions for the other three crops. When training a model using Deep Learning (DL) algorithms, there are more hyperparameters to think about. These include the number of neurons, hidden layers, and learning rate. Setting parameters by hand isn't always possible. Hyperparameter optimization is the process of picking the right set of hyperparameters to make the model work better. In LSTM, the time steps, the number of input and hidden layers, and the number of hidden neurons may all be thought of as hyperparameter optimization variables. You should be cautious while tweaking the hyperparameters since it might change how well the model works. The layers in the deep LSTM model will cause overfitting and delay convergence. The neurons in the hidden layer function in the same way as the LSTM layer. To predict time-series data, you need data from the past, but a regular neural network will just look at the data it gets now. RNN and LSTM can both remember past data, however LSTM models can remember it for longer than RNN models. LSTM has two main ideas that it uses to learn about time-based aspects in data. The memory idea, which brought up the cell state, and the cell concept, which can successfully train the completely linked layers, are the first two things. In LSTM, there are distinct memory cells in the hidden layer for reading, writing, and deleting data. These cells are controlled by three gates: the input gate, the output gate, and the forget gate. These three gates decide what data should be retained in memory. The cell state moves data from one layer to the next. The first step is the forget gate, which lets just the data that is needed past the cell state. The sigmoid layer is the first stage of the input gate. It controls the output value. The Tanh layer is the second step. It creates the vectors of new feature values. Both of them are kept in the cell state. The output gate shows the updated information about the cell. The LSTM analyzes past data and present unknown patterns by controlling them at a basic level to find patterns. This lets it make predictions about the future sooner. Figure 3 explains how LSTM works.

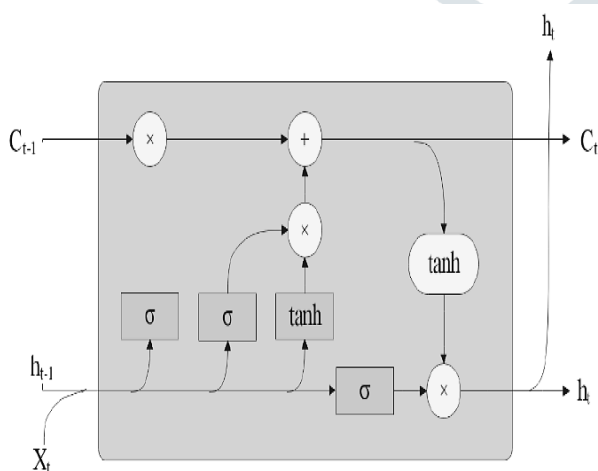


Figure 3. The LSTM Model Functionality

h_{t-1} – previous memory output

C_t – current memory output.

LSTM cell is described in

$$cg_t = \text{Tanh}(wt_{cg} \times [hd_{cg-1}, x_{cg}] + bs_{cg})$$

where, (cg_t) – current memory

(wt_{cg}) – weight matrix

(bs_{cg}) - bias

The input gate controls the current memory input data update to the value of the memory cell and it is calculated in

$$ig_t = \sigma(wt_{ig} \times [hd_{ig-1}, x_{ig}] + bs_{ig})$$

The input gate controls the previous memory data update to the value of the memory cell and it is calculated in

$$fg_t = w_{tf}wt_{fg} \times [hd_{fg-1}, x_{fg}] + bs_{fg}$$

$$cu_t = fit \times lc_{t-1} + cg_t$$

Where, cu_t – current memory cell

lc_{t-1} – last LSTM cell value.

You may also use LSTM in both the stacked and bidirectional forms. In stacked mode, LSTM first works on the input and the next LSTM, and then it works on the outputs of the temporal characteristics that the previous models generated. Stacked LSTM is used in advanced temporal learning features. The bidirectional LSTM trains the extra model, whereas the unidirectional LSTM does not. One LSTM can read the input data from the start of the series to the finish ($t_0 \rightarrow t_n$), while the other can read it from the end to the start ($t_n \rightarrow t_0$). Then, these two models are joined together to make the temporal feature output. Bidirectional learned the model characteristics from both ends of the input sequences. The LSTM model also has the Attention Mechanism to make the predictions more accurate. The LSTM serves as a mechanism for attentiveness. The decoder checks all of the encoder's states again at each step. The last time step hidden state is utilized to encode the input progression during training. The Recurrent Neural Network (RNN) functions as long-term memory, thus facilitating the attention process and enhancing the accuracy of output regression updates. So, the concealed state has all the information about the input sequence. This technique of categorization gives you great tools to make reliable predictions about crop yields.

5. Performance Evaluation

We made sure that our machine learning models would function well in real-life agricultural circumstances by carefully examining how well they operated throughout testing and outcomes evaluation. In this research, the suggested approach is reproduced using LSTM in accordance with the system specifications. We used many metrics to evaluate the experimental outcome, including Accuracy, Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and R-Squared (R2). This part demonstrates how the LSTM model did in terms of the total rate that could be achieved, both in numbers and words. Table 1 and figure 4 indicate how accurate the suggested technique is compared to other methods. The suggested LSTM with an attention mechanism is compared to the existing approaches, which include Convolutional Neural Network (CNN), Deep Neural Network (DNN), Recurrent Neural Network (RNN), and Generative Adversarial Network (GAN). The accuracy results reveal that CNN got 84.88%, DNN got 87.49%, RNN got 90.66%, and GAN got 92.77%. The suggested LSTM with Attention mechanism got a superior accuracy score of 99.10% than the other approaches that are already out there.

Table 1: The Experimental Results of the Proposed Model with Existing Model

Model	Performance Summary for 80% - 20%				
	Accuracy	R2	MAE	MSE	RMSE
DNN	87.49	0.48	0.137	0.060	0.243
CNN	84.88	0.50	0.140	0.063	0.249
RNN	90.66	0.47	0.136	0.059	0.241
LSTM	99.10	0.44	0.132	0.055	0.233
GAN	92.77	0.45	0.133	0.056	0.235

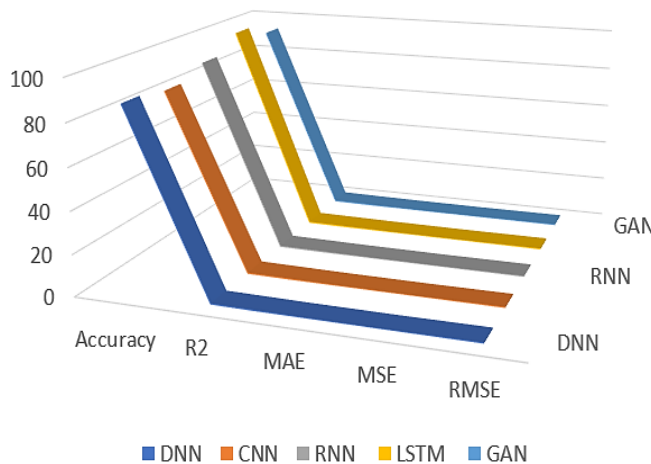


Figure 4. The Different Graphical Representations of Crop Yields from Experimental Results of Proposed Models Compared to Existing Models

The results from the IoT-based [58] soil monitoring system show that it has a lot of promise for improving precision farming methods. The technology provides you vital information that might help crops develop better and utilize resources more efficiently by helping you keep an eye on critical soil parameters all the time [59]. This sensitivity is necessary for making informed changes to farming methods that would boost overall productivity [60]. It's also important to think about how the system may be utilized on different sorts of farms and how it can be altered to match their demands. The model has worked well in the regions that were looked at [61], but it has to be evaluated in other places with varied kinds of soil and weather [62]. To ensure the system functions well across various agricultural contexts.

6. Conclusion

Forecasting how much food crops will produce is a crucial part of growing food in the globe. Farmers benefit greatly from being able to accurately anticipate agricultural yields, which also raises the value of food and the economy. Changes in the weather, such as rain, temperature, and soil, may impact crop yield output. This suggested technique gathers the yield data for three important crops. Then, compare the three crop production predictions to see which one is superior. The suggested model employs linear interpolation to fill in the gaps in the Kaggle datasets. The Long Short-Term Memory (LSTM) with attention mechanism is used to classify better agricultural production predictions for three crops: jowar, paddy, and ragi. The model's performance is measured using Accuracy, R2, MAE, MSE, RMSE, and MAPE. The suggested technique yields superior outcomes, as shown by performance metrics such as accuracy, R2, MAE, and RMSE values of around 99.10%, 0.44, 0.132, and 0.233, respectively, which are relatively more favourable than those of current methods. In subsequent endeavours, the suggested

methodology will be expanded to include more soil factors for the enhancement of predictive outcomes.

References

- [1] D. Paudel, H. Boogaard, A. de Wit, M. van der Velde, M. Claverie, L. Nisini, S. Janssen, S. Osinga, and I. N. Athanasiadis, "Machine learning for regional crop yield forecasting in Europe," *Field Crops Res.*, vol. 276, p. 108377, Feb. 2022
- [2] D. Paudel, H. Boogaard, A. de Wit, S. Janssen, S. Osinga, C. Pylaniadis, and I. N. Athanasiadis, "Machine learning for large-scale crop yield forecasting," *Agric. Syst.*, vol. 187, p. 103016, Feb. 2021
- [3] Ben Hassen, T.; El Bilali, H. Impacts of the Russia-Ukraine War on Global Food Security: Towards More Sustainable and Resilient Food Systems? *Foods* 2022, 11, 2301.
- [4] WHO. World Hunger Is Still Not Going Down after Three Years and Obesity Is Still Growing—UN Report. Available online: <https://www.who.int/news/item/15-07-2019-world-hunger-is-still-not-going-down-after-three-years-and-obesity-is-stillgrowing-un-report>
- [5] Y. Perwej, Firoj Parwej, "A Neuroplasticity (Brain Plasticity) Approach to Use in Artificial Neural Network", *International Journal of Scientific & Engineering Research (IJSER)*, France , ISSN 2229 – 5518, Volume 3, Issue 6, Pages 1- 9, 2012, DOI: 10.13140/2.1.1693.2808
- [6] Venkata K. S. Maddala, Dr. Shantanu Shahi, Yusuf Perwej, H G Govardhana Reddy, "Machine Learning based IoT application to Improve the Quality and precision in Agricultural System", *European Chemical Bulletin (ECB)*, ISSN: 2063-5346, SCOPUS, Hungary, Volume 12, Special Issue 6, Pages 1711 – 1722, May 2023, DOI: 10.31838/ecb/2023.12.si6.157
- [7] KDV Prasad, Yusuf Perwej, E. Nageswara Rao, Himanshu Bhaidas Patel, "IoT Devices for Agricultural to Improve Food and Farming Technology", *Journal of Survey in Fisheries Sciences (JSFS)*, ISSN: 2368-7487, SCOPUS, Vol. 10, No. 1S (2023): Special Issue 1, Pages 4054-4069, Canada, March 2023
- [8] M. Fathi, R. Shah-Hosseini, and A. Moghimi, "3D-ResNet-BiLSTM Model: A Deep Learning Model for County-Level Soybean Yield Prediction with Time-Series Sentinel-1, Sentinel-2 Imagery, and Daymet Data," *Remote Sens.*, vol. 15, no. 23, Art. no. 23, Jan. 2023, doi: 10.3390/rs15235551.
- [9] Champaneri, M.; Chachpara, D.; Chandvidkar, C.; Rathod, M. Crop Yield Prediction Using Machine Learning. *Int. J. Sci. Res.* 2020, 9, 2.
- [10] Sarvesh Kumar, Dr. Shobhit Sinha, Y. Perwej, Ankit Shukla, Dr. Nikhat Akhtar, "Integrated Ensemble Learning Techniques for Precision Crop Yield Prediction", *International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT)*, ISSN: 2456-3307, Volume 11, Issue 3, Pages 1072-1083, June 2025, DOI: 10.32628/CSEIT25113391
- [11] A. Kaneko, T. Kennedy, L. Mei, C. Sintek, M. Burke, S. Ermon, and D. Lobell, "Deep learning for crop yield prediction in Africa," in *International Conference on Machine Learning AI for Social Good Workshop*, Long Beach, us, 2019
- [12] Farheen Siddiqui, Homa Rizvi, Y. Perwej, Shamim Ahmad, Dr. Nikhat Akhtar, "Leveraging AI for Social Impact in Environmental Sustainability", *International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET)*, Print ISSN: 2395-1990, Online ISSN: 2394-4099, Volume 12, No. 4, Pages 253-266, August 2025, DOI: 10.32628/IJSRSET2512506

- [13] Farheen Siddiqui, Shilpi Shukla, Y. Perwej, Sweta Singh, N. Akhtar, "IoT-Enabled Decision Support System for Crop Recommendation Using Machine Learning Algorithms", *Journal of Emerging Technologies and Innovative Research (JETIR)*, ISSN-2349-5162, Volume 12, Issue 11, Pages 803 - 809, November 2025, DOI: 10.6084/m9.jetir.JETIR2511200
- [14] Satterthwaite, D.; Mcgranahan, G.; Tacoli, C. Urbanization and its implications for food and farming. *Philos. Trans. R. Soc. Ser. B* 2010, 365, 2809–2820
- [15] Lin, T.; Zhong, R.; Wang, Y.; Xu, J.; Jiang, H.; Xu, J.; Ying, Y.; Rodriguez, L.; Ting, K.C.; Li, H. DeepCropNet: A deep spatial-temporal learning framework for county-level corn yield estimation. *Environ. Res. Lett.* 2020, 15, 034016
- [16] T. K. Fegade and B. V. Pawar, "Crop prediction using artificial neural network and support vector machine," *Data Management, Analytics and Innovation*, Springer, Berlin, Germany, pp. 311-324, 2020.
- [17] Sweta Singh, Shilpi Shukla, Yusuf Perwej, Farheen Siddiqui, Nikhat Akhtar, "Optimizing Crop Yield Forecasts Through Deep Neural Network Architectures Using Omdena Dataset", *International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET)*, Print ISSN: 2395-1990, Online ISSN: 2394-4099, Volume 12, No. 5, Pages 377-389, October 2025, DOI: 10.32628/IJSRSET2513819
- [18] Alberto Gonzalez-Sanchez, Juan Frausto-Solis, Waldo Ojeda-Bustamante, "Attribute Selection Impact on Linear and Nonlinear Regression Models for Crop Yield Prediction", *The Scientific World Journal*, vol. 2014, Article ID 509429, 10 pages, 2014.
- [19] Woittiez, L. S., Van Wijk, M. T., Slingerland, M., Van Noordwijk, M., & Giller, K. E. (2017). Yield gaps in oil palm: A quantitative review of contributing factors. *European Journal of Agronomy*, 83, 57-77.
- [20] Kajal, Neha Singh, Nikhat Akhtar, Ms. Sana Rabbani, Y. Perwej, Susheel Kumar, "Using Emerging Deep Convolutional Neural Networks (DCNN) Learning Techniques for Detecting Phony News", *International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT)*, ISSN: 2456-3307, Volume 10, Issue 1, Pages 122-137, 2024, DOI: 10.32628/CSEIT2410113
- [21] N.Akhtar, Kumar Bibhuti B. Singh, Devendra Agarwal, Y. Perwej, "Improving Quality of Life with Emerging AI and IoT Based Healthcare Monitoring Systems", *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, ISSN: 2456-3307, Volume 11, Issue 1, Pages 96-107, January 2025, DOI: 10.32628/CSEIT2514551
- [22] Y. Perwej, "The Bidirectional Long-Short-Term Memory Neural Network based Word Retrieval for Arabic Documents", *Transactions on Machine Learning and Artificial Intelligence (TMLAI)*, which is published by Society for Science and Education, United Kingdom (UK), ISSN 2054-7390, Volume 3, Issue 1, Pages 16 - 27, 2015, DOI: 10.14738/tmlai.31.863
- [23] Y. Perwej, "Recurrent Neural Network Method in Arabic Words Recognition System", *International Journal of Computer Science and Telecommunications (IJCST)*, which is published by Sysbase Solution (Ltd), UK, London, (<http://www.ijcst.org>), ISSN 2047-3338, Volume 3, Issue 11, Pages 43-48, 2012
- [24] Oikonomidis A, CatalCand Kassahun A 2022 Hybrid deep learning-based models for crop yield prediction *Appl. Artif. Intell.* 36 2031822
- [25] Nagarjuna Tandra, Nikhat Akhtar, K Padmanaban, L. Gaganathan, "A finite-element dual-level contextual informed neural network with swarm space hopping algorithm based optimal feature selection and detection for EEG-based epileptic seizure detection", *Swarm and Evolutionary Computation*, Elsevier, SCIE, Volume 97, Pages 1- 19, August 2025, DOI: 10.1016/j.swevo.2025.102072
- [26] Hasan M, MarjanMA, UddinMP, AfjalMI, Kardy S, MaS and NamY 2023 Ensemble machine learning-based recommendation system for effective prediction of suitable agricultural crop *Frontiers in Plant Science* 14 1234555
- [27] Sarvesh Kumar, Dr. Shobhit Sinha, Dr. Yusuf Perwej, Ankit Shukla, Dr. Nikhat Akhtar, "Integrated Ensemble Learning Techniques for Precision Crop Yield Prediction", *International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT)*, ISSN: 2456-3307, Volume 11, Issue 3, Pages 1072-1083, June 2025, DOI: 10.32628/CSEIT25113391
- [28] Kumar Bibhuti B. Singh, N. Akhtar, Devendra Agarwal, Susheel Kumar, Y. Perwej, "An Evaluation of OpenCV's Investigation into Hand Gesture Recognition Methods", *International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET)*, Print ISSN: 2395-1990, Online ISSN: 2394-4099, Volume 12, Issue 1, Pages 01-14, January 2025, DOI: 10.32628/IJSRSET25121150
- [29]. Sadenova, M., Beisekenov, N., Varbanov, P.S. and Pan, T., 2023. Application of machine learning and neural networks to predict the yield of cereals, legumes, oilseeds and forage crops in Kazakhstan. *Agriculture*, 13(6), p.1195.
- [30] Chlingaryan, A.; Sukkarieh, S.; Whelan, B. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Comput. Electron. Agric.* 2018, 151, 61–69.
- [31] Marques Ramos, A.P.; Prado Osco, L.; Elis Garcia Furuya, D.; Nunes Goncalves, W.; Cordeiro Santana, D.; Pereira Ribeiro Teodoro, L.; Antonio da Silva Junior, C.; Fernando Capristo-Silva, G.; Li, J.; Henrique Rojo Baio, F.; et al. A random forest ranking approach to predict yield in maize with uav-based vegetation spectral indices. *Comput. Electron. Agric.* 2020, 178, 105791
- [32] Shilpi Shukla, Sweta Singh, Yusuf Perwej, Farheen Siddiqui, Nikhat Akhtar, "AI-Powered Crop Recommendation for Smart Farming, Current Barriers, and Future Perspectives", *International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT)*, Volume 11, Issue 5, Pages 308-323, October 2025, DOI: 10.32628/CSEIT251117134
- [33] Ji, Y.; Chen, Z.; Cheng, Q.; Liu, R.; Li, M.; Yan, X.; Li, G.; Wang, D.; Fu, L.; Ma, Y.; et al. Estimation of plant height and yield based on UAV imagery in faba bean (*Vicia faba* L.). *Plant Methods* 2022, 18, 26
- [34] Sunny Kumar, Apoorva Dwivedi, Yusuf Perwej, Moazzam Haidari, Siddharth Singh, Dr. Nagarajan Gurusamy, "A Smart IoT-Image Processing System for Real-Time Skin Cancer Detection", *Journal of Neonatal Surgery (JNS)*, SCOPUS, ISSN: 2226-0439 (Online), Volume 14, Issue S14, Pages 823-831, April 2025, DOI: 10.52783/jns.v14.4330
- [35] Yang, Q.; Shi, L.; Han, J.; Zha, Y.; Zhu, P. Deep convolutional neural networks for rice grain yield estimation at the ripening stage using UAV-based remotely sensed images. *Field Crops Res.* 2019, 235, 142–153.
- [36] Farheen Siddiqui, Sana Rabbani, Dr. Yusuf Perwej, Hina Rabbani, Dr. Nikhat Akhtar, "Leveraging Cloud Computing, IoT and Big Data for Intelligent Infrastructure Management in Smart Cities", *Journal of Emerging Technologies and Innovative Research (JETIR)*, ISSN-2349-5162, Volume 12,

Issue 8, Pages 301 - 310, August 2025, DOI: 10.6084/m9.jetir.JETIR2508335

[37] Pantazi, X.E.; Moshou, D.; Alexandridis, T.; Whetton, R.L.; Mouazen, A.M. Wheat yield prediction using machine learning and advanced sensing techniques. *Comput. Electron. Agric.* 2016, 121, 57–65.

[38] Whetton, R.; Zhao, Y.; Shaddad, S.; Mouazen, A.M. Nonlinear parametric modelling to study how soil properties affect crop yields and NDVI. *Comput. Electron. Agricult.* 2017, 138, 127–136.

[39] Y. Perwej, Firoj Parwej, Nikhat Akhtar, “An Intelligent Cardiac Ailment Prediction Using Efficient ROCK Algorithm and K- Means & C4.5 Algorithm”, *European Journal of Engineering Research and Science (EJERS)*, Bruxelles, Belgium, ISSN: 2506-8016 (Online), Vol. 3, No. 12, Pages 126 – 134, 2018, DOI: 10.24018/ejers.2018.3.12.989

[40] N. Akhtar, Hemlata Pant, Apoorva Dwivedi, Vivek Jain, Y. Perwej, “A Breast Cancer Diagnosis Framework Based on Machine Learning”, *International Journal of Scientific Research in Science, Engineering and Technology*, Print ISSN: 2395-1990, Online ISSN: 2394-4099, Volume 10, Issue 3, Pages 118-132, 2023, DOI: 10.32628/IJSRSET2310375

[41] Anjali Yadav, Shruti Dwivedi, Anubhav Dwivedi, Ujjwal Thakur, Dr. Nikhat Akhtar, “Intelligent Disease Diagnosis: A Multi-Disease Prediction Approach Using Machine Learning”, *International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET)*, Volume 12, No. 3, Pages 98 -109, May 2025, DOI: 10.32628/IJSRSET251235

[42] Amanullah Ansari, Shrejal Singh, Dr. Nikhat Akhtar, “AI-Driven Crop Disease Detection and Management in Smart Agriculture”, *International Journal of Scientific Research in Science and Technology (IJSRST)*, SSN: 2395-6011, Volume 12, Issue 3, Pages 309-319, May 2025, DOI: 10.32628/IJSRST2512341

[43] Apoorva Dwivedi, K. Manivannan, Sunny Kumar, Neha Anand, Y. Perwej, Rakhi Kamra, “A Real-Time Environmental Pollution Monitoring Framework Using IoT and Remote Sensing Technologies”, *International Journal of Environmental Sciences (IJES)*, SCOPUS, ISSN: 2229-7359, Volume 11, Number 7s, Pages 1064 - 1075, June 2025

[44] Tian, H.; Wang, P.; Tansey, K.; Han, D.; Zhang, J.; Zhang, S.; Li, H. A deep learning framework under attention mechanism for wheat yield estimation using remotely sensed indices in the Guanzhong Plain, PR China. *Int. J. Appl. Earth Observ. Geoinform.* 2021, 102, 102375.

[45] Elavarasan, Raj Vincent, P.M. Fuzzy deep learning-based crop yield prediction model for sustainable agronomical frameworks. *Neural Comput. Appl.* 2021, 33, 13205–13224.

[46] Y. Perwej, Nikhat Akhtar, Devendra Agarwal, “The emerging technologies of Artificial Intelligence of Things (AIoT) current scenario, challenges, and opportunities”, Book Title “Convergence of Artificial Intelligence and Internet of Things for Industrial Automation”, SCOPUS, ISBN: 978-1-032-42844-4, CRC Press, Taylor & Francis Group, 2024

Link: <https://www.taylorfrancis.com/chapters/edit/10.1201/9781003509240-1/emerging-technologiesartificial-intelligence-things-aiot-current-scenario-challenges-opportunities-yusuf-perwej-nikhatakhtar-devendra-agarwal?context=ubx&refId=537f1a8f-6a94-4439-b337-3ad3d1ce8845>, DOI: 10.1201/9781003509240-1

[47] N. Akhtar, Nazia Tabassum, Asif Perwej, Y. Perwej, “Data Analytics and Visualization Using Tableau Utilitarian for COVID-19 (Coronavirus)”, *Global Journal of Engineering and Technology Advances (GJETA)*, ISSN :

2582-5003, Volume 3, Issue 2, Pages 28-50, 2020, DOI: 10.30574/gjeta.2020.3.2.0029

[45] Mahmoud AbouGhaly, Y. Perwej, Mumdouh Mirghani Mohamed Hassan, Nikhat Akhtar, “Smart Sensors and Intelligent Systems: Applications in Engineering Monitoring” ,

International Journal of Intelligent Systems and Applications in Engineering, SCOPUS, ISSN: 2147- 6799, Volume 12, Issue 22s, Pages 720–727, July 2024

[48] Anmol Chauhan, Ms. Sana Rabbani, Devendra Agarwal, Nikhat Akhtar, Yusuf Perwej, “Diffusion Dynamics Applied with Novel Methodologies”, *International Journal of Innovative Research in Computer Science and Technology (IJIRCST)*, ISSN (Online): 2347-5552, Volume-12, Issue-4, Pages 52 - 58, July 2024, DOI: 10.55524/ijircst.2024.12.4.9

[49] KDV Prasad, Yusuf Perwej, E. Nageswara Rao, Himanshu Bhaidas Patel, “IoT Devices for Agricultural to Improve Food and Farming Technology”, *Journal of Survey in Fisheries Sciences (JSFS)*, ISSN: 2368-7487, SCOPUS, Volume 10, No. 1S (2023): Special Issue 1, Pages 4054-4069, Canada, 2023

[50] López-Aguilar, K.; Benavides-Mendoza, A.; González-Morales, S.; Juárez-Maldonado, A.; Chinas-Sánchez, P.; Morelos-Moreno, A. Artificial Neural Network Modeling of Greenhouse Tomato Yield and Aerial Dry Matter. *Agriculture* 2020, 10, 97.

[51] Dunderski, D.; Jac'ımovic', G.; Crnobarac, J.; Viskovic', J.; Latkovic', D. Using Digital Image Analysis to Estimate Corn Ear Traits in Agrotechnical Field Trials: The Case with Harvest Residues and Fertilization Regimes. *Agriculture* 2023, 13, 732.

[52] Fernandez-Gallego, J.A.; Buchailot, M.L.; Gracia-Romero, A.; Vatter, T.; Diaz, O.V.; Aparicio Gutiérrez, N.; Nieto-Taladriz, M.T.; Kerfal, S.; Serret, M.D.; Araus, J.L.; et al. Cereal Crop Ear Counting in Field Conditions Using Zenithal RGB Images. *J. Vis. Exp.* 2019, 144, e58695.

[53] Neha Anand, Arpita Vishwakarma, Y. Perwej, Neeta Bhusal Sharma, Atifa Parveen, “A Hybrid Deep Learning Ensemble Approach for Enhanced Data Mining Efficiency”, *Journal of Emerging Technologies and Innovative Research (JETIR)*, ISSN-2349-5162, Volume 12, Issue 8, Pages 268 - 276, August 2025, DOI:10.6084/m9.jetir.JETIR2508238

[54] Sarvesh Kumar, Y. Perwej, Farheen Siddiqui, Ankit Shukla, Dr. Nikhat Akhtar, “A Data-Driven Framework for Fake News Detection Via Web Scraping and Machine Learning Approach”, *International Journal of Innovative Science and Research Technology (IJSRT)*, ISSN- 2456-2165, Volume 10, Issue 6, Pages 1391 - 1404, June 2025, DOI: 10.38124/ijisrt/25jun1003

[55] Yang,W.; Nigon, T.; Hao, Z.; Paiao, G.D.; Fernández, F.G.; Mulla, D.; Yang, C. Estimation of corn yield based on hyperspectral imagery and convolutional neural network. *Comp. Electr. Agric.* 2021, 184, 106092

[56] Ma, Y.; Zhang, Z. A Bayesian Domain Adversarial Neural Network for Corn Yield Prediction. *IEEE Geosci. Remote Sens. Lett.* 2022, 19

[57] Nikhat Akhtar, Devendra Agarwal, “An Efficient Mining for Recommendation System for Academics”, *International Journal of Recent Technology and Engineering(IJRTE)*, ISSN 2277-3878 (online), SCOPUS, Volume-8, Issue-5, Pages 1619-1626, January 2020, DOI: 10.35940/ijrte.E5924.018520

[58] Vaishali Singh, Soumya Verma, Ayush Srivastava, Abhishek Dubey, Dr. Nikhat Akhtar, “Eco- Sensing System for Water Pollution and Microplastic Detection”, *International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT)*,

ISSN: 2456-3307, Volume 11, Issue 3, Pages 679-690, May 2025, DOI: 10.32628/CSEIT25113333

[59] Nikhat Akhtar, Dr. Hemlata Pant, Apoorva Dwivedi, Vivek Jain, Yusuf Perwej, "A Breast Cancer Diagnosis Framework Based on Machine Learning", International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Print ISSN: 2395-1990, Online ISSN: 2394-4099, Volume 10, Issue 3, Pages 118-132, May-June-2023, DOI: 10.32628/IJSRSET2310375

[60] Neha Kulshrestha, Nikhat Akhtar, Yusuf Perwej, "Deep Learning Models for Object Recognition and Quality Surveillance", Accepted International Conference on Emerging Trends in IoT and Computing Technologies (ICEICT-2022), ISBN 978-10324-852-49, SCOPUS, Routledge, Taylor & Francis, CRC Press, Chapter 75, Pages 508-518, Goel Institute of Technology & Management, Lucknow, May 2022, Link - <https://www.routledge.com/Emerging-Trends-in-IoT-and-Computing-Technologies-Proceedings-ofInternational/Tripathi-Verma/p/book/9781032485249#> DOI: 10.1201/9781003350057-75

[61] Rezk, N.G.; Hemdan, E.E.D.; Attia, A.F.; El-Sayed, A.; El-Rashidy, M.A. An efficient IoT based smart farming system using machine learning algorithms. Multimed. Tools Appl. 2021, 80, 773–797

[62] Elbasi, E.; Zaki, C.; Topcu, A.E.; Abdelbaki, W.; Zreikat, A.I.; Cina, E.; Shdefat, A.; Saker, L. Crop prediction model using machine learning algorithms. Appl. Sci. 2023, 13, 9288

