ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

OPTIMIZING CROP RECOMMENDATION SYSTEMS USING ADVANCED DEEP LEARNING TECHNIQUES

PAPPALA MOHAN RAO[1], KUNJAM NAGESWAR RAO[2]

[1]Research Scholar, Department of CS & SE, Andhra University College of Engineering, Andhra University, India.
[2]Professor, Department Of CS&SE, Andhra University College of Engineering, Andhra University, India.
[1]mohanpappala@gmail.com, [2]kunjamnag@gmail.com

ABSTRACT:

The agricultural sector plays a significant role in economic development, especially for developing nations, but farmers experience hurdles such as climate change interventions, soil management, and crop yield choices. To this end, this study turns to Deep Learning (DL) to solve these problems and develop a holistic approach that employs a range of DL approaches for crop selection based on soil condition, climate, and past farming records. Predicted models and methods: These are Long Short-Term Memory (LSTM), Bidirectional LSTM (BiLSTM), and Transformers; precision agriculture through crop prediction; data-driven decision-making on resource allocation to boost farm revenue; and soil management through the prediction of the soil pH and recommended crops. This plan of action is designed to help farmers make the most beneficial decision to enhance the farming results.

Keywords: *Agricultural, Deep Learning (DL), Long Short-Term Memory (LSTM), Bidirectional LSTM (BiLSTM), Transformers, crop recommendation.*

1. INTRODUCTION:

Agriculture is the mainstay of India's economy, as it sustains the majority of people. However, agricultural output is still difficult to predict because some of the variables, like climatic change, soil fertility, and the nature of crops that can be cultivated, may not be easy to monitor. Data mining provides a way to dangle patterns and extract useful information from voluminous data sets that can help farmers in choosing crops and cultivating methods. This has resulted in it being ranked as the second largest agricultural country in the world, having 6 million square kilometres under cultivation. In sharp contrast, only 45% of the current water utilisation is used to grow crops, with most farmers still depending on rain-fed farming. Challenges to agriculture include the lack of focus, which reduces the number of active agricultural workers, and the growing farmers' suicide cases. In such a scenario, they had to reach out to more yield-enhancing technologies and newer methods. This specialized agriculture involves correct spacing, application, and use of soil, fertilizers, and other resources, which has also become popular among farmers, producing high-demand crops that will help them in tiding over these difficulties. The following research will develop an exemplary crop forecasting model with integrated machine learning algorithms to aid precision agriculture. This is intended to identify the crop type that would be expected to give maximum returns, based on performance indicators such as NPK, rainfall, moisture content, and temperature on the ground. These are relevant factors to crop development and in identifying crop varieties suitable to be nurtured. Some of the variables used in collecting the data include NPK[1] for the soil, rainfall, temperature, and pH level. With high-level machine learning algorithms, the results of this research will provide better and more accurate crop prediction. Other attributes include depth of topsoil, its acidity or alkalinity, erodibility, permeability, texture, capability to drain, water retention, and color of the topsoil considered in choosing the most appropriate

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

crop. It also applies support vector machine SVM, random forest[4], and logistic regression in classification to improve the accuracy. The research makes use of support vector machines SVM, random forests[4], and logistic regression as machine learning techniques in improving the accuracy in classification. This study also uses models like LSTM[5], BiLSTM, and Transformer networks. Proposed here is an advanced machine learning system to build a robust, accurate, and reliable system of recommendation that would help farmers in their functioning and thereby increase efficiency and sustainability in Indian agriculture.

The study has many independent variables like Nitrogen, Phosphorous, Potassium in soil, Rainfall intensity, Temperature and PH level. But other important parameters like Topsoil depth, Acidity/Alkalinity, Erodibility, Permeability, Texture, Drainage capacity, Water holding capacity and Topsoil colour also matter.

The proposed crop recommendation model is a recommendation system to help farmers to decide which crop to grow. With the help of machine learning algorithms and environmental datasets the model is developed to optimize the agricultural process and make it more sustainable in India. These findings should be practical for farmers and help them to choose the crop as per current and projected climate situation and ensure more efficiency in agricultural sector. So this study highlights the need to incorporate machine-learning solutions in agricultural domain to resolve these emerging issues among Indian farmers. So the crop forecasting model developed in the study should be viewed as a step towards improving the existing methods and practices in agriculture, better utilization of resources and food and economic security in Indian context.

2. LITERATURE REVIEW:

[7] Sivanandam uses Selection Model and K-nearest Neighbor algorithms to enhance crop prediction for farmers in developing nations. Such tools in a nation like India, where the domination of the majority is agriculture, help to predict the output so that people can chalk out a strategy for selecting crops. The data mining categorizes and analyzes the information, and KNN gives real-time forecasts to optimize the yield and minimize the risks of loss in the event of crop failure phenomena.

[6] CH.Rakesh looks into some of the machine learning techniques, namely Naive Bayes, Decision Trees, and Random Forest, in the realm of precision agriculture in India. All of these techniques use soil data to suggest crops and their practices in order to increase productivity. It works effectively in the classification of crops and prediction of yield, making it useful for web-based data analysis.

[8] Daneshwari Modi designed an SVM-based crop recommendation system to increase productivity and reduce food shortages in India. The soil parameters are considered by the system, after which it predicts the suitable crop with 97% accuracy. It uses ensemble techniques, random forests, and Bayesian networks. Future enhancements could incorporate the internet of things for real-time soil monitoring.

[9] Ramchandra, in his study, allows SVM, Random Forest, and Naive Bayes machine learning algorithms for the prediction of crop type of India, considering its soil and weather conditions; the last one gave the best result. The system is developed to improve crop selection and reducing agricultural risk. The plans are to have an application in a national and international setup.

[10] Jaichandram R elaborates on the application of a Machine Learning Algorithm: Light GBM in making crop recommendations for enhanced results by investigating environmental information and soil nutrition. The system is robust and accurate, capturing temperature, humidity, rainfall, pH levels, and soil content. It also foresees crop diseases and prescriptions for fertilizers. This algorithm leads to increased productivity and profitability in agriculture by

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

having the backing of neural networks for tailored guidance.

As articulated by [11] P. Parameswari, it drives economic growth and competitiveness. Sustainable agriculture features such as crop rotation, harvesting, and soil management are achievable with the help of PART, a machine-learning algorithm that gives an accuracy of 99.4%. All these technologies highly developed precision agriculture by the means of real-time data available with crops, soil, and weather. Researchers focused on developing the effective crop prediction models that help farmers to take better decisions on what to plant.

[12] The system by Raj Kumar utilizes Convolutional Neural Networks in the detection of plant diseases and in the suggestion of crops to be grown based on soil quality. In utilizing the Plant Village Dataset, an accuracy of 98.2% in disease classification is achieved. It outperforms the other classifiers with the SVC. The system further makes suggestions for improving protection and productivity by suggesting crops that can be grown in the available soil.

[13] Sujata Chakravarty's crop recommendation framework uses machine learning techniques to perform crop recommendations with a high percentage, in this case, up to 99.54% using Naive Bayes. It has preprocessing, classification, and performance evaluation implemented with data. Future work would cover deep neural networks with a cloud-based analysis.

[14] Prema T. Akkasaligar's system combines soil pH with season and temperature to recommend crops using neural networks with inputs of Random Forest. The system is expected to improve practice in agriculture and the selection of crops in India.

[15] Tapas Kumae Mishra presents a system for crop prediction using soil quality, N-P-K values, humidity, and rainfall determination to recommend crops to the users. The Crops Prediction System for Indian regions and crop yield problems related to soil impurity and the weather is done through K-Neighbors and Random Forest Classifiers. The system enhances an increase in productivity and seeks to attract new and younger farmers through enhancements of the future with larger datasets.

[16] This study focuses on enhancing crop recommendations using machine learning and cloud-based solutions. By integrating datasets from diverse sources, including Indian agricultural records and local district data, and applying advanced classifiers like SGDC, it achieves high accuracy in predictions. The approach personalizes recommendations, accounting for climate, soil, and water availability, aiding sustainable farming.

Recent advancements in machine learning (ML) provide valuable tools for analyzing agricultural data, enabling better crop selection, resource management, and cultivation strategies. In 2024, ML is widely acknowledged for its role in mitigating climate risks, boosting yields, and promoting sustainable farming practices.

While existing models focus on traditional machine learning techniques like KNN, Random Forest, and SVM for crop prediction, they lack the ability to process temporal agricultural data effectively. This paper addresses these gaps by introducing advanced deep learning models, including LSTM, Bi-LSTM, and Transformers, for more accurate and dynamic crop recommendations. We trained our model using these three deep learning techniques and introduced the TRANSFORMER ENCODING MODEL for better crop recommendations. Following is the design flow of the proposed system as shown in Fig(ii).

3. PROPOSED SYSTEM:

Figure (a) shows the end-to-end flow of the complex yet scalable crop recommendation system that ingests multiple data inputs and uses deep learning models. The process starts at the Data Collection Layer of the system where it fetches a wide range of soil data including soil nutrient status that is NPK contents and pH of the soil. It also fetches essential weather data like temperature, rainfall, humidity and other weather conditions for the crops. With all these data sources the system is guaranteed to have a comprehensive and rich data set to handle the complexity of agricultural ecosystems.

Then comes the preprocessing layer which is a critical layer for data quality and usability. In this layer the raw data is preprocessed and cleaned to remove errors like missing values and outliers. This also includes data formatting which means putting the data into a structure that the machine learning models can understand. This is important because poorly prepared data will always lead to bad performance; hence the importance of the preprocessing phase is to improve the performance of the next phases.

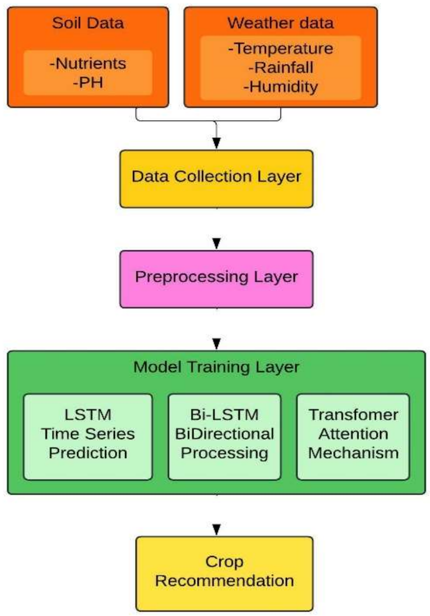
Next, the data goes through the Model Training Layer, where the system has a number of sophisticated deep learning models, all of which constitute the system, and each of which is designed to process a particular type of data. The first model employed is called Long Short-Term Memory (LSTM), which takes advantage of the temporal analysis as applied to weather and soil data. The LSTM has high capability in handling the sequence in the data, which is essential to understand how previous context impacts future context.

Subsequently, the system incorporates BiLSTM, an improved variation of LSTM, which analyses the data in a forward direction as well as in a backward direction. This two-way approach enables the model to capture all the aspects of the data, making the interpretation of the interconnections of various variables easier. BiLSTM is especially useful in processing the sequential data where some features might be hidden in the reverse direction while being evident in the forward direction from the time series analysis.

The third model integrated into the system is the Transformer Attention Mechanism, which has the pass-through capability for various data input forms and excels in modelling the dependency of features across different distances. The Transformer model excels at detecting complex coupling between different pieces of information, which is particularly useful for the data that contains multiple interconnected variables. An advantage of this model is that it incorporates the attention mechanism in focusing on some parts of the input data, which makes crop recommendations more precise.

The crop recommendations that are made available to the users are a composite of three deep-learning models with a high level of accuracy and reliability. This integration process ensures that all the advantages of the two models will be incorporated, hence enhancing the formulation of the recommendation system. What the company provides the farmer with is a map of the farmer's plot with details of the type of soil and the climatic conditions. Opinions and options provided permit farmers to make unprecedented decisions as to which plants should be given priority in their farms, therefore increasing their revenue and, at the same time, honouring the natural world.

Therefore, with the assistance of such a complex approach, the specified crop recommendation system is aimed at significant advancement in the sphere of precise agriculture. It can enhance the efficiency of work to be done on the farm. Moreover, it is ecological because farmers cultivate them in such a way that they are friendly not only to the local climate but to the tract of land as well. The system provides valuable recommendations in accurate, efficient, and relevant ways that contribute to assisting farmers in getting higher yields and seriously attempting food security and sustainable farming.



Fig(a): Design architecture of Crop Recommendation System.

4. ADVANCED DEEP LEARNING MODELS FOR CROP RECOMMENDATION:

In this paper, advanced deep learning models—Long Short-Term Memory, Bidirectional Long Short-Term Memory, and Transformer Encoding Models—are going to be applied for crop yield prediction. These models have been trained with time-series data regarding the advanced and complex factors governing the growth of crops and give output to farmers about their future crop yields with a very high degree of accuracy and reliability.

4.1 LSTM DEEP LEARNING MODEL:

The Long Short-Term Memory (LSTM) networks are an extended class of recurrent neural networks (RNNs that are created especially for the interpretation of sequential data – the kind of data in which the order of items matters significantly. Original RNNs, which represent a specific class of neural networks, have a number of drawbacks, one of which is the failure to pre-train, owing to such problems as vanishing and exploding gradients in terms of the learning process with regard to the long sequences of data. To this end, LSTMs overcome these challenges by incorporating specific components meant for holding and processing information for more extended time periods. They are particularly ideal for any usage that requires managing temporal data, which covers language modelling, time series, and speech, to mention but a few. However, the basic structure of LSTM networks has several essential parts, such as a memory cell (Cell State), Input gate, Forget gate and Output gate. These components help in regulating the flow of information through the network, letting it store important information it has gained more than useless information it may have received. That is, long-term memory enables LSTMs to make complex tasks such as crop recommendations where the effect of previous conditions and environment is significant.

In the idea of crop recommendation, LSTMs can, in fact, handle chronic data such as climate changes, soil conditions, previous crops, and others. They are essential in modeling the complicated relationships that define crop development and yields. The procedure starts with the accumulation of large amounts of data from different sources, namely local weather stations, satellite data, agricultural census and IoT devices on farmlands/fields. These sources give a variety of information on environmental conditions, including temperature, rainfall, humidity, moisture of soil, and nutrients, all of which are relevant while giving crop advice.

After the accumulation of a set of standard data, this data must pass through a procedure called data preprocessing so as to qualify for analysis. The first stage is to deal with missing data in the raw data set and normalize and scale the data before selecting the features that will enhance the growth of crops. Feature selection is quite significant in the reduction of dimensionality, which in turn improves the efficiency and performance of the LSTM model.

What gets into building a good LSTM model for recommending crops includes the following: First, the architecture of the model has to be decided to include the correct number of layers, and for each layer, the correct number of memory units (neurons) has to be chosen in order to incorporate the complexity of the data. Correct parameters, such as the learning rate, batch size, and number of epochs to train, must be set appropriately since they influence the model's performance. Also equally important is the issue of overfitting, whereby the model becomes too 'parochial' with the training data set so that it fails to perform well with new data sets. Some of the methods used for overfitting are dropout regularization, detection of the level of overfitting through early stopping, and cross-validation. Another essential consideration made in the model design is the ability to set the sequence length, which defines to what extent the model looks back to the past in order to make a prediction. While this parameter introduces an essential feature of memorising long-term dependencies, it has to be tuned with respect to the other parameters to optimise the computational complexity of the model.

Training of the LSTM model entails feeding historical data to the model along with the use of optimizers such as the Adam or RMSprop to reduce the loss function, which depicts the closeness of the model forecast to actual data. Other aspects that should be considered in the learning process include managing the learning rate because it defines how fast the model updates. If it is chosen to be too high, then the model may converge prematurely; if the learning rate is set too low, then it will take a long time for any convergence to occur.

Model training is a crucial step just before its implementation, but it is equally important to validate and evaluate the trained LSTM model. This is a process that aims to evaluate the model with the help of new data that have not been used during model training. This can be done using different measures of accuracy, which are root mean squared error (RMSE), mean absolute error (MAE) as well as the R-squared statistic. The correctness of the used model's ability to generalize to other conditions and datasets is essential for its implementation in actual practice. Other methods, like the k-fold cross-validation, are yet other validation methods that could be used to validate the model further through the different splits of the data.

After validation, LSTM helps to offer crop yield predictions which farmers and other stakeholders in the agricultural value chain can make use of in making their decisions. Such recommendations may be helpful in scheduling planting activities, choosing the suitable crop varieties and in efficient use of the resources. The idea of using LSTM-based recommendations with other data-driven decision-support systems can play a very instrumental role for the agricultural sector to achieve further improvements in several aspects, such as efficiency and productivity as well as sustainability. The use of such sophisticated models does not only enhance recommendations on matters performance of crops, but also is a step towards enhancing food security, given that it also enhances the aspect of accuracy in carrying out agricultural practices.

Finally, the implementation of LSTM networks in crop recommendations leads to a significant contribution towards precision agriculture. This technique is a powerful tool in addressing the more nuanced problem of predicting agricultural outcomes, taking advantage of how LSTMs model complex temporal dependencies. This type of LSTM models integration with extensive environmental data set provides a powerful domain towards improving plant recommendation thus promote sustainable and productive agriculture practices in the long run.

4.2 BI-LSTM DEEP LEARNING MODEL:

Bidirectional Long Short-Term Memory (Bi-LSTM) networks are more advanced than the traditional LSTM networks that are further equipped with the bidirectional processing, thereby, increasing accuracy in sequence prediction tasks. Unlike the LSTMs The

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

standard LSTMs that work with data in a single, forward direction, Bi-LSTMs are specifically designed for looking at and examining data sequences from both the past and the future. Bi-LSTMs are better at understanding the context which surrounds each point by using the two processing mechanisms. A Bi-LSTM indeed constitutes two LSTM layers, one taking the input sequence to the end and the other processing it backwards. The outputs of the two layers are typically either concatenated or summed and the final result is the full representation of the input sequence on each time step. It is how a comprehensive picture is being gotten, which in turn, in the prediction process, enables Bi-LSTMs to get the full context of each element in the sequence.

One of the main reasons for BILSTMs that are very effective is those that are designed to learn the representation at very deep and local levels, are well-suited for these. As for Bi-LSTMs, they can be tremendously efficient in such environments due to the fact that they are particularly equipped with the ability to capture the intricate dependencies and interactions of the data points in a sequence, which in turn, brings up more precise and contextual situations.

BiLSTMs are particularly useful for crop recommendations as they support the analysis of sequential data, such as weather conditions, soil quality, and the historical crop yields. By processing the sequences in bidirectional fashion, BiLSTMs discern more nuanced patterns or relationships that might have been hidden in a unidirectional model. Consider the situation where a weather event impacting crop growth may not be due to historical conditions, but to future environmental conditions. The yesteryear effect of rainfall may be considered alongside the future condition of temperature. Such an analysis would result in better crop yield predictions. This representation requires more memory and processing power than a unidirectional LSTM because sequences are processed in two directions so that the BiLSTM reaches an analysis of past and future states of the sequences. This model is substantially harder to tune than a uni-LSTM due to the need to balance the forward and backward layer contributions, notably when passing information further down the chain. While there may be disadvantages to the Bi-LSTM model, like the LSTM's tendency to overfit, there is certainly something to gain from the richer context of the BiLSTM model in the world of crop recommendations.

4.3 TRANSFORMER ENCODING MODEL:

The method based on the Transformer Encoding Model handled the sequential data very successfully with the easy inlaying of an innovative self-attention mechanism. The current transformer model, unlike traditional RNN or LSTM, makes parallel processing of complete sequences instead of sequential processing. This is allowed by dynamically weighing the importance of different parts of the sequence—a technique generated from its self-attention mechanism. That is, for each element in some sequence, the queries, keys, and values are transformed into vectors of the same size; then the self-attention mechanism outputs scores showing the importance of one element relative to all others in the sequence. Sorry, but it is only through this weighted sum that a Transformer naturally grasps dependencies and patterns whose interactions are pretty crucial in attempts at modeling ordered data such as time series, language, or indeed most other sequence information.

Some essential decisions and constructions are needed to help a model handle sequential data more efficiently within a Transformer architecture. The critical one is positional encoding, which retains the order of elements in the sequence. This has been the case since the Transformer's design made it clear that they were not ordered sequence processors. Unlike RNN and LSTM, they don't see an order of elements. These types of architectures thus make use of positional embeddings summed up with input embeddings to feed the model with relative and absolute positions of elements within a sequence. By doing so, every position will gain unique identification, thus keeping the order of the sequence and letting the model learn properly the relationship depending on the position.

The other important feature of the Transformer model is multi-headed attention. It contributes an added ability of the model to consume multiple features of a sequence at the same time, hence giving it full understanding of the data. The attention mechanism is broken into different "heads", which are focused on different parts of the sequence. This will enable the model to understand complex patterns and relationships in the data. This proves very useful in the case when elements of a sequence have very complex and sophisticated interactions, so complex that one attention mechanism could never model them. Feed-forward neural networks provide much-needed nonlinearity that helps the model learn more complex representations of data. These layers transform the output from the attention mechanism into a final sequence embedding to make a prediction or classification. It is through the implementation of layer normalization and residual connections that stability and acceleration of the training of the Transformer model occur, making gradient flow during back-propagation much more efficient. These results are critical for training very deep models.

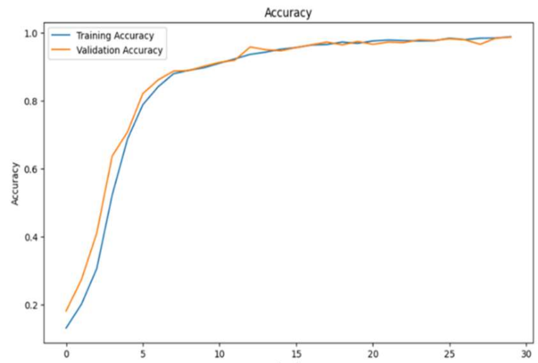
One of the strengths of using Transformer models for crop recommendations lies in their efficient learning from large-scale data, which is collected from multiple sources. Agricultural datasets typically include satellite images, weather information, soil moisture content, and historical crop yields. The Transformer architecture fits well with the fusion of a large number of different sources of data due to mechanisms for embeddings, while position encodings enable spatial and temporal dependencies. Of the most impactful attributes of the Transformer model in relation to crop recommendation, there is the multiheaded attention mechanism that can capture very diverse patterns in its data. Feedforward networks further refinance the feature representation and give an accurate crop yield prediction that will, in ways, optimize how we produce our food to increase food security.

While using Transformer models for crop recommendations, the metrics that shall form a basis for evaluating models will be accuracy, precision, recall, F1 score, and the confusion matrices. Accuracy pertains to the whole rightness—is the model correct in recommending? Precision and recall relate to how correctly it identifies a specific crop type. The F1 score creates a balance between precision and recall, thus offering a holistic view of how well the model is performing when the dataset is imbalanced. Confusion matrices go down to the granularities of how the model is performing across different classes and can be used to point out areas for fine-tuning of the model. These metrics combined provide a strong evaluation framework that not only predicts yields as best as possible but also identifies relevant crops, which is very important information in practical agricultural decision-making. The most obvious benefit of the encoding model in crop recommendation is simplicity.

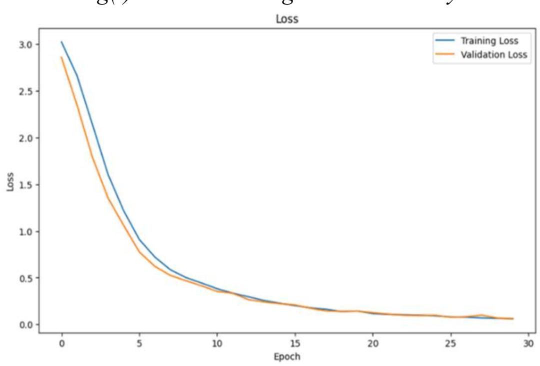
This model is inherently capable of handling large datasets, capturing long-range dependencies, and merging diversified sources of data. Therefore, this encoding model is the most vital in boosting power in agricultural decision-making. In using the transformer, optimal resources are made to achieve better sustainability in agricultural practices through accurate crop recommendations and better crop management. Adoption of better models like Transformers is important for real improvements in predictive power over future innovations dealing with global food security, considering the fact that climate change and resource scarcity continue challenging the agricultural industry, among other productivity-enhancing advances.

5. RESULTS:

Figure (i) shows the accuracy of the LSTM training model regarding epochs. Figure (ii) depicts the loss value per epoch. The error rate for the LSTM training model is in the decreasing stage for these plots. These show that the LSTM model is getting trained and learned on the data set.

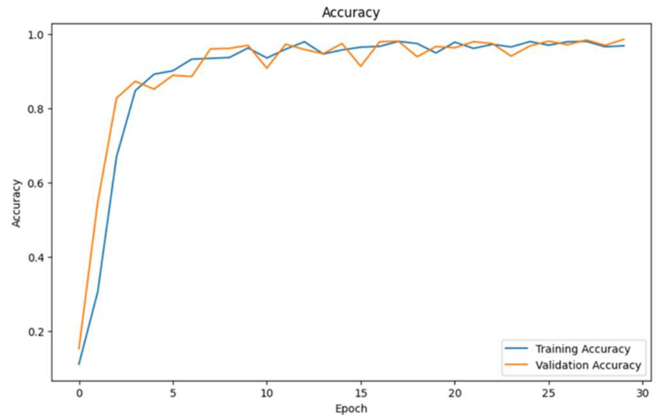


Fig(i): LSTM Training Model Accuracy.

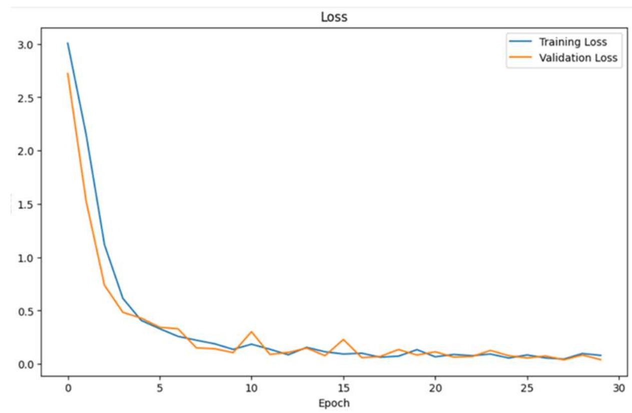


Fig(ii): LSTM Training Model Loss.

Figure (iii) displays BiLSTM model accuracy as a function of the training epoch, where there is performance improvement. Figure (iv) displays graphically how the BiLSTM model loss and epoch are depicted to realize the error reduction rate. These are descriptive in showing the rate at which error reduction has been realized. The epochs traverse in the x-axis, and the desired changes in the values of the two parameters are plotted regarding their values/amounts on the y-axis. These figures show the details of the BiLSTM model training processes in relation to accuracy and loss in the course of training.



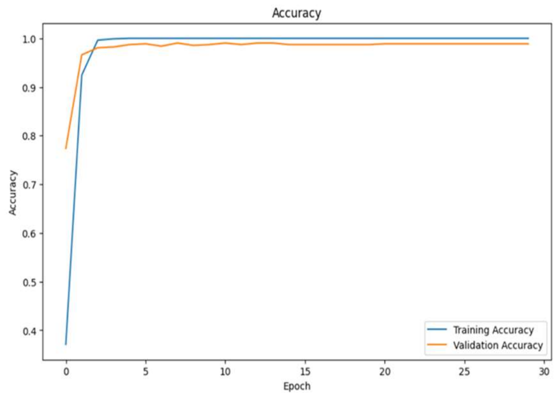
Fig(iii): BiLSTM Training Model Accuracy.



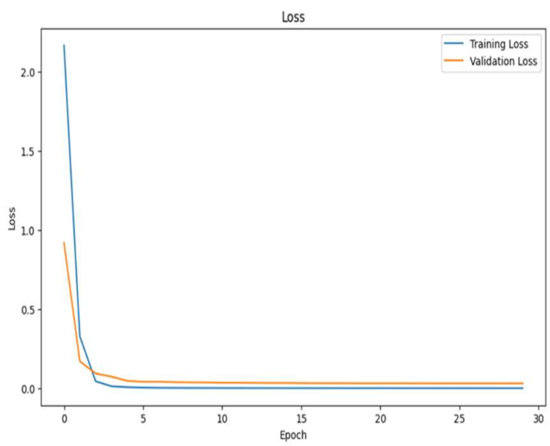
Fig(iv): BiLSTM Training Model Loss.

Figure (v) Measure of Accuracy of the Transformer Encoding Model at each Epoch
Figure (vi) Loss per epoch which the model has for training

A view of improvements of the Training Process on the Transformer Encoding Model is shown in Figure v., which presents values of Measure of Accuracy at each Epoch. In Figure (vi), one can get a loss per epoch that the model has for training, and can notice it is decreasing; therefore, an error rate is reduced. These historical metrics will help in understanding the model's training rate and performance gain.



Fig(v): Transformer Encoding Model for Accuracy.



Fig(vi): Transformer Encoding Model for Loss.

The study's strengths include the use of advanced deep learning models like Transformers and Bi-LSTM, robust data integration, and alignment with precision agriculture goals. Weaknesses involve computational complexity, limited real-time data integration, regional specificity, and challenges in scaling to larger datasets or broader applications.

6. CONCLUSION:

This study demonstrates that deep learning can significantly address critical challenges in agriculture. The proposed framework optimizes crop selection based on soil characteristics, weather patterns, and historical agricultural data using advanced algorithms such as LSTM, BiLSTM, and Transformers. Among these, the Transformer model stands out, achieving remarkable results with 100% training accuracy and 98.87% test accuracy, outperforming both LSTM and BiLSTM as well as traditional machine learning techniques.

By addressing gaps in handling complex agricultural data, the Transformer Encoding Model improves the management of large datasets, leading to enhanced predictive accuracy. Our framework empowers data-driven decision-making, increases farm profitability, and optimizes resource allocation. It provides precise crop recommendations that contribute to sustainable farming practices, boosting productivity, and ensuring the model's effectiveness in real-world agricultural contexts. This research has significant practical implications for the agriculture industry. It enhances crop recommendation accuracy, optimizes resource usage, and supports the development of digital farming tools. Additionally, it helps insurers manage crop failure risks and offers scalable AI solutions. These advancements align with the growing demand for precision farming and sustainable agricultural practices.

REFERENCES

[1] T. Blesslin Sheeba', 'L. D. Vijay Anand', 'Gunaselvi Manohar', 'Saravana Selvan', ' C. Bazil Wilfred', 'K. Muthukumar', 'S. Padmavathy', 'P. Ramesh Kumar', and 'Belete Tessema Asfaw'- Machine Learning Algorithm for Soil Analysis and Classification of Micronutrients in IoT-Enabled Automated Farms Hindawi Journal of Nanomaterials Volume 2022, Article ID 5343965, 7 pages

[2] K. Ranjini, A. Suruliandi, and 'S. P. Raja' - An Ensemble of Heterogeneous Incremental Classifiers for Assisted Reproductive Technology Outcome Prediction. ieee transactions on computational social systems, vol. 8, no. 3, june 2021.

[3] Hong-Xia, Wang. "An improved collaborative filtering recommendation algorithm." In 2019 IEEE 4th International Conference on Big Data Analytics (ICBDA), pp. 431-435. IEEE, 2019.

[4] Sheri Mahender Reddy, Naini Sriman and thatipamula saikiran, An Optimized Machine Learning Approach for Predicting Various Crop Yields (March 19, 2021).

[5] Prof. Pritesh A. Patil, Mr. Pranav Athavale, Mr. Manas Bothara, Ms. Siddhi Tambolkar, Mr. Aditya More'- An Overview of Crop Yield Prediction using Machine Learning Approach.Volume: 10 Issue: 02 | Feb 2023

[6] CH. Rakesh D. 2Vishnu Vardhan, 3Babu Bhavani Vasantha, 4G. Sai Krishna Crop Recommendation and Prediction System, 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS).

[7] Sivanandam k, Prasanth m m.kumarasamy Saran s, Naveen b , An Efficient Machine Learning Approaches for Crop Recommendation based on Soil Characteristics, Proceedings of the Second International Conference on Electronics and Renewable Systems (ICEARS-2023). IEEE Xplore Part Number: CFP23AV8-ART; ISBN: 979-8-3503-4664-0.

[8] Daneshwari Mod, Ashok V. Sutagundar, Vijayalaxmi Yalavigi, Anupama Aravatagimath, Crop Recommendation Using Machine Learning Algorithm, 2021 5th International Conference on Information Systems and Computer Networks (ISCON) GLA University, Mathura, India. Oct 22-23, 2021.

[9] Ramachandra A C, Garre Venkata Ankitha, Idupulapati Divya, Parimi Vandana, H S Jagadeesh, Crop Recommendation using Machine Learning 2023 International Conference on Data Science and Network Security (ICDSNS)

[10] Jaichandran R, T.Murali Krishna , Sri Harsha Arigela, Ramakrishnan Raman , Dharani N , Ashok Kumar, "Light gbm algorithm based crop recommendation by weather detection and acquired soil nutrients." In 2022 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS), pp. 1-5. IEEE, 2022.

[11] Parameswari, Ponnusamy, N. Rajathi, and K. J. Harshanaa. "Machine learning approaches for crop recommendation.", international conference on advancements in electrical, electronics, communication, computing and automation (ICAECA), pp. 1-5. IEEE, 2021.

[12] Raj Kumar, Neha Shukla, "Plant Disease Detection and Crop Recommendation Using CNN and Machine Learning." In 2022 International Mobile and Embedded Technology Conference (MECON), pp. 168-172. IEEE, 2022.

[13] Rakesh Kumar Ray, Saneev Kumar Das, and Sujata Chakravarty. "Smart crop recommender system-a machine learning approach." In 2022 12th International Conference on Cloud Computing, Data Science & Engineering (Confluence), pp. 494-499. IEEE, 2022.

[14] 'Prema T. Akkasaligar',' Sunanda Biradar',' Manjula C. Gudgeri',' Sana Mohammadi A Mulla', "Soil Mineral Prediction of Crops Using Machine Learning." In 2022 Third International Conference on Intelligent Computing Instrumentation and Control Technologies (ICICICT), pp. 1027-1032. IEEE, 2022.

[15] Tapas Kumar Mishra*, Sambit Kumar Mishra, Kanaparthi Jeevan Sai, Bachu Sai Alekhya, Athukuri Rama Nishith, "Crop recommendation system using knn and random forest considering indian data set." In 2021 19th OITS International Conference on Information Technology (OCIT), pp. 308-312. IEEE, 2021.

[16] Senapaty, Murali Krishna, Abhishek Ray, and Neelamadhab Padhy. 2024. "A Decision Support System for Crop Recommendation Using Machine Learning Classification Algorithms" *Agriculture* 14, no. 8: 1256. https://doi.org/10.3390/agriculture14081256