International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

IJISAE

ISSN:2147-6799 www.ijisae.org Original Research Paper

Cnn-Bilstm With Attention Mechanism for Crop Yield and Fertilizer Recommendation

S. Vasanthanageswari[1], P. Prabhu[2]

Submitted: 27/01/2024 **Revised**: 05/03/2024 **Accepted**: 13/03/2024

Abstract: Agriculture has been an incredibly significant means of wealth for humans from our ancient periods. Among the most significant economic sectors in India is agriculture. Approximately 70% of Indians are employed in agriculture, either directly or indirectly. Today's conventional agriculture is inefficient due to a number of issues, including global warming and climate change. Crop yield has decreased as a result of climate change's detrimental effects on the crop production cycle. However, the development of other technologies, such as machine learning and artificial intelligence, has shown great promise in other fields, and incorporating these methods into farming is also a wonderful step forward. Statistical models are usually employed to predict agricultural yield, although this is time-consuming and labor-intensive. The surge in popularity of machine learning and deep learning is a major milestone in the discipline However, in today's complex world, it is a gazillion challenges, because of the climate variations, limited resources, and the need to feed a growing global population. To overcome all these immense and enormous issues, we introduce a novel method, where an ensemble model that combines an attention mechanism, bidirectional long short-term memory networks (BILSTM) with convolutional neural networks (CNN). Through the identification of complex patterns in numerical agricultural statistics, our approach seeks to improve forecast accuracy and offer useful information to stakeholders and farmers. The ensemble model outperforms the individual networks in crop yield prediction by combining their capabilities. The performance rate is measured in terms of accuracy, precision, recall and f-measure.crop dataset scores 98% and fertilizer dataset gives the accuracy of 99%

Keywords: *Agriculture, Bidirectional Long Short-Term Memory Networks, Crop Yield Prediction, Convolutional Neural Networks, Deep Learning.*

1. Introduction

The foundation for world food stability is agriculture, which supplies food for the planet's expanding population. It contributes significantly to national economies and is essential to reducing poverty. According to statistics 60% of the Indian economy depends on agriculture whereas 10% farmers are educated, only 2% use modern agricultural techniques [3]. Agriculture supports a number of businesses outside food production, such as bio fuels, textiles, and medicines. Furthermore, it affects climate change, water supplies, and biodiversity, among other environmental issues. In our state, agriculture is a key component of social and economic development. It influences the general well-being of communities, creates jobs, and moulds rural livelihoods, Land usage, irrigation, and environmental preservation. Food sustainability, financial stability, and the general standard of living in a region all depend on the prosperity of agriculture. Conventional approaches can place a great deal of emphasis on previous data, which might not fully capture the dynamic character of agriculture that is influenced by shifting climatic conditions and changing farming techniques. Less accurate projections for certain places and time periods may result from conventional techniques' lack of the temporal as well as spatial precision required to capture fine-scale differences in crop development. Due to their heavy reliance on past weather patterns, traditional models are prone to errors when it comes to changes in weather brought on by climate change. A new method of predicting crop yields is crop prediction utilizing deep learning models, such as Convolutional Neural Networks (CNNs) along with Bidirectional Long Short-Term Memory networks (BILSTMs). Convolutional Neural Network (CNN) can be used to extract climatic parameters because it is suitable to process data with multiple array formats such as one-dimensional data (Signals and Sequences), two-dimensional data (images), and three-dimensional data (videos). The temporal dependencies of the climatic dataset used were captured by one – dimensional convolution. With CNN all the necessary information and the high-level features from the input data can be extracted by CNN when paired with the pooling operation [9] .These models use neural network technology to examine intricate patterns in agricultural data, offering insightful information for decision-making. This methodology, which considers both temporal and spatial fluctuations in agriculture, attempts to improve the precision and reliability of crop production and fertilizer

1 Ph.D., Research Scholar, Department of Computer Applications,Alagappa University, Karaikudi, Tamilnadu, India
ORCID ID : 0000-0001-5044-0977
E-mail: [1]smilingsudhamathi@gmail.com
2 Associate Professor in Information Technology, Directorate of Distance Education, Alagappa University, Karaikudi, Tamilnadu, India
ORCID ID : 0000-0002-5960-4103
E-mail: [2]prabhup@alagappauniversity.ac.in

predictions by merging CNNs with BiLSTMs with attention mechanism. The constraints of conventional approaches can be addressed with flexibility and adaptability due to machine learning techniques, which can offer insightful information for productive and sustainable farming practices. Fig.1. describes the General architecture of machine learning models.

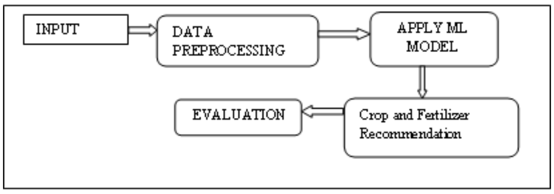

Fig.1. General Architecture of Machine Learning Models.

An organized procedure is followed in the creation of machine learning models, which begins with problem definition and pertinent data collection. Cleaning and preparing the data before dividing it into testing, training, and test sets is essential. Choosing an appropriate model, training it using training data and assessing its performance with test and validation datasets are the following steps. Hyper parameter tuning is done to optimize the model based on the evaluation. The model is then used to forecast new data, and its efficacy is continuously checked and updated to maintain it that way. Refinement and adjustment must be made iteratively in this process, taking into account changes in data attributes and model performance. To enable farmers to receive the profit from the same area of land without causing soil degradation, productivity must be raised. Indian farmers are unable to select the best crop for their soil due to a variety of elements, including pH, temperature, humidity, and rainfall, N, P, and K.Most farmers don't know which type of fertilizer—standard or organic—to use depending on the needs of their soil. Soil degradation brought on by insufficient and uneven fertilization is causing nutrient extraction and the emergence of second-generation issues with nutrient management. To solve the issue by proposing a recommendation system through an CNN-BILSTM with Attention mechanism model with majority voting technique crop for the site specific parameters with high accuracy and efficiency. To recommend fertilizer on the basis of N, P, K values and crop.

A CNN-BiLSTM-Attention mechanism for crop and fertilizer recommendation is a sophisticated machine learning model that integrates Convolutional neural networks, also called neural networks, (CNNs), Bidirectional Long Short-Term Memory (BiLSTM) networks, and Attention mechanisms. The goal of this model is to analyze multiple data sets linked to agriculture, such as soil properties, meteorological information, and historical crop performance.

The following is the order of the remaining paragraphs; Chapter II covers related works. Chapter III provides a full discussion of the several phases of the suggested methodology. Chapter IV covers the argument and outcomes.

2. Related Works

The goal of the literature review regarding crop yield prediction is to understand and evaluate the vast amount of knowledge that currently exists regarding crop yield prediction. It is a crucial part of agricultural research. In the context of agricultural yield prediction research, a number of important goals are sought to be accomplished by this thorough review of earlier studies, approaches, and results. First, by outlining the current level of understanding in the field of agricultural production prediction, the literature review builds the contextual environment. It explores the several elements and variables that affect crop yield, emphasizing how intricate the agricultural environment is. In order for researchers and readers to comprehend the environment in which new models for prediction are created and evaluated, this contextualization is essential. In addition, the literature review is essential for pointing out weaknesses, restrictions, and difficulties in the current body of research. Through a critical evaluation of previous approaches and results, researchers can identify areas in which further progress is required. This approach advances agricultural science in general and provides information for the creation of more reliable prediction models.

Iniyan Shanmugam, Jebakumar Rethnaraj, Srinivasan Rajendran, Senthilraja Manickam [1], proposed the application of long short-term memory (LSTM) time series analysis to field crop yield prediction. In this work long short-term memory (LSTM) time series analysis is used and the usage of deep learning models is also mentioned. The LSTM algorithm outperformed the RNN method in terms of accuracy (93%).

C.K.Srinivas [2], suggest project estimates the amount of agricultural output depending on variables such as field area, soil moisture content, temperature, and humidity The Random Forest algorithm works well for predicting crop productivity. Appropriate fertilizer ratio recommendations can increase crop productivity. Machine learning-method is used for agricultural yield prediction: Assists farmers in choosing the right crops and applying the right amount of fertilizer.

B.G. Chaitra, B.M. Sagar, N.K. Cauvery. Padmashree[3], the study examines the importance of deep learning to predict crop productivity and compare deep learning's performance with various other machine learning approaches. Additionally, it discusses the use of three algorithms, Deep Neural Network (DNN), Random Forest

(RF), XGBoost (extreme gradient boosting). One of which is the Deep Neural Network, to accurately anticipate crop yield. The Deep Neural Network (DNN) produced the best crop yield forecast accuracy, 96%. The accuracy of Random Forest (RF): 92.9%- 93.3% accuracy rate for Extreme Gradient Boosting (XGBoost), 96% accuracy of Deep Neural Networks (DNNs).

Pardeep Kaur, Preeti Singh, Charu Madhu, Nidhi Garg [4], present research addresses the prediction of wheat yield using a CNN-LSTM model. Deep learning model CNN-LSTM for wheat yield prediction CNN-LSTM model performs better than current deep learning Predictive models that use deep learning are appropriate.

Zheng Li, Ruosi Xu, Xiaoru Luo, Xin Cao [5], Used an improved CNN structure and BiLSTM network,. The proposed hybrid model enhances the accuracy of wind power forecast. The paper's hybrid model performs better in terms of predictions. Deep learning, signal decomposition, and data processing are all combined in the wind power forecast model.

Malika Kulyal, Parul Saxena [6], discusses the application of CNN and DNN, two Deep Learning techniques, Crop yield prediction uses supervised machine learning methods like Random Forest. Additionally, deep learning methods like CNN and DNN are used. Crop yield prediction is done by machine learning algorithm. Deep learning techniques and random forest are mostly used. The study examines machine learning techniques for predicting crop productivity.

Preeti Saini, Bharti Nagpal, Puneet Garg, Sachin S.Kumar[7],proposes a hybrid CNN-Bi-LSTM deep learning-based technique for sugarcane yield prediction using ARIMA Conventional Stacked-LSTM, Holt-winter Time-series, and GPR approaches are combined with a hybrid CNN-Bi-LSTM_CYP deep learning methodology. The CNN-BILSTM_CYP method performed better than conventional methods

Dilli Paudel, Allard de Wit, Hendrik Boogaard, Diego Marcos, Sjoukje A. Osinga, Ioannis N. Athanasiadis[8], In Germany, examined LSTM and 1DCNN models over soft wheat LSTM and 1DCNN models are evaluated for their performance and interpretability in agricultural production forecasting. Performance compared between GBDT and linear trend models.

S.S. Olofintuyi, E.A. Olajubu, DejiOlanike[9], In this study, a deep learning method for predicting cocoa yield using a CNN and RNN with LSTM (long short-term memory) is discussed. Predicting cocoa yield with a deep learning strategy (CNN-RNN + LSTM) is effective. The lowest mean absolute error was obtained by the suggested CNN-RNN with LSTM model. The model is compared with other machine learning techniques.

S. Archana, K.G.Saranya, [10], proposed a survey on Crop Yield Prediction and made discussion of factors influencing crop productivity; analysis of performance indicators and deep learning techniques. Among the deep learning algorithms used to forecast agricultural yield are CNN, RNN, LSTM, and MLP.The evaluation metrics that are frequently employed in the examined studies include RMSE, MAPE, R2, MSE, and MAE.

Dr. B. Srinivasa Rao,Dr. Seemantini Nadiger, Dr. Khushal N. Pathade[11], CNN, a Bidirectional Long Short-Term Memory network, with an Attention Mechanism are combined to create the CNN-BiLSTM-ECA model, which is offered for training (AM),. It combines CNN and a BILSTM network with a lightweight ECA attention module. When comparing to other models, such CNN and BILSTM, which have an accuracy of roughly 98.5%, the suggested method performs well.

P.S.S.Gopi & M.Karthikeyan [12], suggested that an automated crop suggestion system, Red Fox Optimization using Ensemble Recurrent Neural Network for Crops Recommendation with Yield Prediction (RFOERNN-CRYP) model is the main topic of this study. In order to achieve better prediction performance than the individual classifier models, the RFOERNN-CRYP model that is being presented uses an ensemble learning process that employs three distinct DL models: long short-term memory (LSTM), bidirectional LSTM (BiLSTM), along with gated recurrent unit (GRU).

S.Iniyan, M.SenthilRaja, R.Srinivasan, C.Santhanakrishnan, Aravindan Srinivasan [13], proposed the initiative tackles the pressing need for customized fertilizer and crop advice in India's many agricultural areas. With the use of machine learning methods including outlier identification, k-nearest neighbors, random forests, and linear regression, our approach seeks to give farmers precise recommendations based on soil and environmental conditions. The association with input and product yield is established using the linear regression algorithm, and K-Neighbors provide recommendations by determining how similar farm attributes are to those of nearby farms.

D.Anantha Reddy, Bhagyashri Dadore and Aarti Watekar [14], proposed a research work the main goal is to assist farmers in choosing crops that make sense for their particular situation and the state of the environment at large. The development of forecasting techniques that take into account a variety of parameters, including soil nutrients, pH values, humidity levels, even rainfall patterns, can help achieve this goal. A variety of machine learning models are applied, including Gaussian Naive Bayes (GNB), decision-tree (DT), Support Vector Machines (SVM), and Logistic Regression (LR).

Neetu Mittal, Akash Bhanja[15],The objective is to create a predictive model based on machine learning that, taking into account a number of variables including soil type, climate, rainfall, and resource availability, suggests the crop that would be best for a certain area. Using natural language processing (NLP) techniques, the model is going to be trained to assess and extract relevant information about the traits, growing circumstances, and prospective yield of a variety of crops from text data.

P.Prabhu & N.Anbazhagan,[16] In this paper we have proposed FI-FCM algorithm for Business intelligence based on frequent item sets and Fuzzy C Means clustering to extract the intelligence from the dataset this algorithm to help customers to find ,recommend products they wish to purchase by producing the list of recommended products.

S.Vasanthanageswari, P.Prabhu [17], In order to forecast crop production using fertilizer amount, this research suggests applying the Tree-structured Parzen Estimator (TPEOSM), an Improved Support Vector Machine, to an agricultural dataset. Evaluation is done on performance criteria like accuracy, recall, f1 score, and precision.

M.Sivakami and P.Prabhu [18], the work on the aforementioned issues utilizing a variety of data mining approaches has been extensively documented in the literature and is critically examined here. The accuracy level indicated made it evident that more precision is required with the arrival of distinctive decisions in order to provide doctors with better support when making decisions.

Shehab Mohamed Beram, Harikumar Pallathadka, Indrajit Patra and P Prabhu [19], Image processing using a digital system was utilized to build a novel technique by including many our suggested effort involved preprocessing to produce clean photos and remove noise. This process would help to eliminate erroneous segments and enhance image quality. Image datasets are classified using KNN, CNN, along with ANN classification algorithms.

Prabhu P [16-19] et.al.Proposed various machine learning algorithm for improving the accuracy of classification using various real world dataset and achieved high accuracy up to 99.5%.

3. Proposed CNN-BiLSTM with attention mechanism Model

To overcome the difficulty such as suitable crop prediction and the amount of fertilizer that required by the crop we have proposed a new CNN-BILSTM with attention mechanism model. It predicts both the suitable crop and optimal fertilizer. Fig.2 describes the proposed CNN-BILSTM with Attention Mechanism architecture used to predict the suitable crop and optimum use of fertilizer.

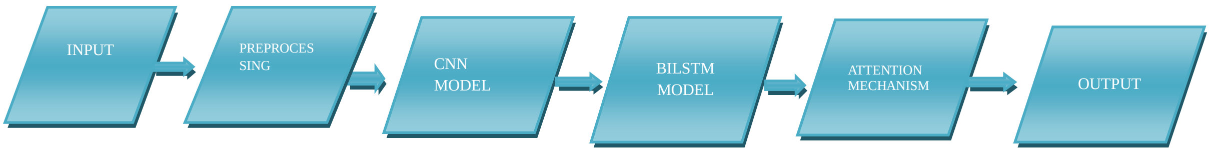

Fig. 2. Proposed CNN-BILSTM with Attention mechanism.

Table 2. Fertilizer Dataset

Crop	N	P	K	pH
Rice	80	40	40	5.5
Barley	70	40	45	5.5
Maize	80	40	20	5.5

Crop and Fertilizer data are input data which comprises of temperature, humidity, rainfall, ph value and Nitrogen, phosphorus and potassium. Next step is preprocessing stage data is preprocessed by label encoding and Normalization using standard deviation. In order to improve our forecast accuracy, we are implementing CNN BiLSTM with Attention mechanism. Convolutional filters are applied to the input data by CNNs, which results in feature maps that highlight the occurrence of particular patterns throughout the data. These features that have been derived from the data can draw attention to significant geographical linkages that may be related to yield results. BiLSTMs combine these findings to generate a more comprehensive view of the temporal dynamics at work by processing data in every direction. The attention mechanism helps the model to de-emphasize less important information and focus on factors that are very predictive of the result, like important weather occurrences or ideal growth circumstances. This mechanism allows the model to dynamically concentrate on specific elements of the input sequence, providing a means to identify the most influential factors that contribute to crop yield. The ensemble model combines these individual networks in order to utilize their complementary strengths. Through the utilization of diverse forecasts produced by many models, the ensemble improves forecast accuracy overall, reduces the likelihood of over fitting, and increases stability.

3.1 Exploratory data analysis

Dataset is collected from the Github repository[12]. Agricultural and fertilizer dataset comprises of temperature, humidity, rainfall, and ph value Nitrogen, phosphorus, potassium and ph. The Table.1 explains the features of crop dataset for the crop recommendation and Table. 2 the appropriate fertilizer the crop required. It will guide the farmers to select the crop for the particular temperature, humidity, rainfall and ph value and the appropriate fertilizer such as nitrogen, phosphorus.

Table 1. Crop Dataset

Temperature	humidity	pH	rainfall	label
20.87	82	6.502	202.93	Rice
26.08	56.06	6	152.09	wheat
28.95	89.07	6.42	57.65	Mung Bean

Temperature represents the average that relevant for crop growth. The range of the value lies between 8.82 to 54.98.It is measured in terms of degree or Celsius. The average humidity level in the air it affect the crop growth. It usually represented in percentage. The min value is 10.03 and max value is 99.98.Soil ph value is the availability of nutrients in the soil for the crop growth. It ranges from 3.50 to 9.93 that is acidic to basic. Amount of rainfall during a specific period used for crop irrigation. It is measured in millimeter or centimeter.min value is 20.2 and max value is 397.3. For each crop the temperature, humidity, ph values vary. A variety of data analysis, statistical, and machine learning activities, correlation values are used to determine the direction and strength of the relationship between variables. Fig.3. represents the correlation matrix of the features in the dataset.

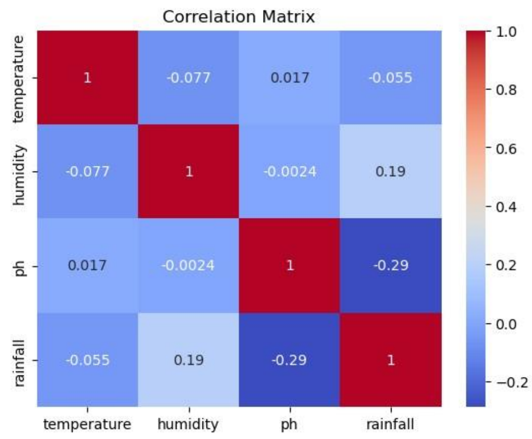

Fig.3. Correlation Matrix of crop dataset

The association between several environmental parameters, such as temperature, humidity, pH, and rainfall, is displayed in the correlation matrix that you have provided. The coefficient of correlation between two variables is represented by each of the cells in the matrix, which ranges from -1 to 1. An inverse relationship is indicated by a negative value, whereas a direct relationship is indicated by a positive value. For instance, there is a small negative correlation (-0.077) between pH and rainfall, which means that while one grows, the other slightly lowers. On the other hand, there is a large inverse relationship between temperature and humidity, as indicated by their strong negative correlation (-0.8). Fig.4 represents the temperature, humidity, pH, rainfall values for each crop in the crop dataset.

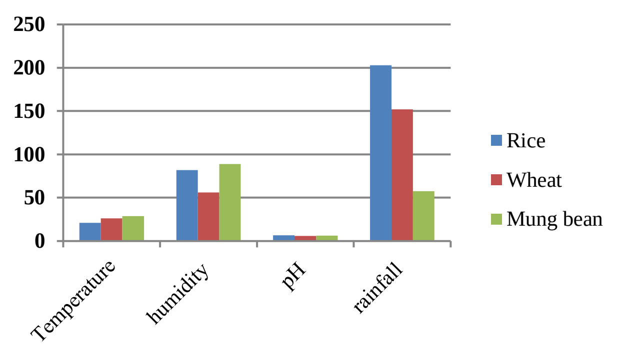

Fig.4. Various factor values for each crop

Fig.5 represents the fertilizer amount required for each crop in the fertilizer dataset.

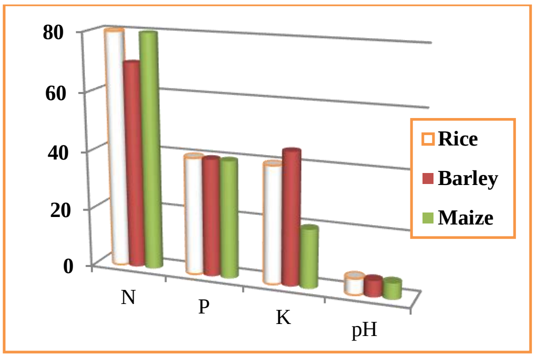

Fig.5. Fertilizer amout for each crop

3.2 The Preprocessing Stage

The process of converting unprocessed data into processed information that a machine learning model may use is known as data preprocessing. As the first and most crucial stage in creating a machine learning model.

3.2.1 Label Encoding

The process of converting labels, or data kinds, into a numerical format so that machines can read them is known as label encoding. One popular encoding technique for managing categorical information is label encoding.

The use of those labels can then be more intelligently determined by machine learning algorithms. It is an essential pre-processing stage for the structural dataset in supervised learning. The crop names in this work's label column are categorical values that are encoded to numerical values. For example, wheat and rice are organized as 0, 1 via the label converter. Table 3 displays sample label encoding for different types of crops.

Table 3. Label Encoding

Label	0	1	2	3
Name of the crop	Adzuki Beans	Black gram	Chickpea	Coconut

3.2.2 Normalization

A basic data preprocessing method called normalization is used in statistics and machine learning to modify the magnitude of data components. Because normalization enables characteristics with different scales and ranges to contribute equally to the analysis, it is especially crucial for datasets including these kinds of features. For many machine learning and statistical modeling tasks, using StandardScaler as a preprocessor is essential. By subtracting the mean and scaling to the unit variance, StandardScaler normalizes features. Known as z-score normalization, this procedure modifies the data so that each characteristic has a mean of 0 and a standard deviation of 1.

$$z = \frac{(x-\mu)}{\sigma} \qquad (1)$$

Where Z is the standardized value, x is the original value,µ is the mean of feature column. σ is the standard deviation of the feature column.

3.3 Modeling Phase

In the ever-evolving landscape of agriculture, predicting crop yield demands a sophisticated dance between spatial and sequential understanding. Our Hybrid Deep Learning Model choreographs this intricate ballet by seamlessly integrating Convolutional Neural Networks (CNNs) and Bidirectional Long Short-Term Memory (BILSTM) networks with Attention mechanism. An advanced machine learning model that combines Convolutional Neural Networks (CNNs), Bidirectional Long Short-Term Memory (BiLSTM) networks, and Attention mechanisms is called a CNN-BiLSTM-Attention mechanism for crop and fertilizer recommendation. This model is intended to examine a variety of agriculturally-related data sets, including soil characteristics, meteorological data, past crop performance. Fig.6. Shows the Proposed CNN-BILSTM with Attention Mechanism architecture to offer advice on the best crops for growing and the right fertilizers to employ in order to achieve the highest possible yield.

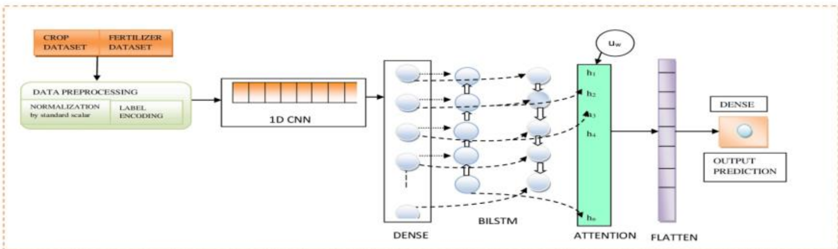

Fig.6. Proposed CNN-BiLSTM with Attention mechanism architecture

▪

3.3.1. Convolution Neural Networks (CNNs):

CNNs take the stage for spatial feature extraction. In our model, they meticulously scan through agricultural data, discerning patterns and extracting spatial features. It captures and canvas the agricultural factors like temperature, humidity, and rainfall. The CNN's keen eye ensures that spatial intricacies are woven into the fabric of our predictive tapestry.

1D Convolutional Neural Network (1D CNN):

Time series and other sequences are good candidates for 1D CNNs. To collect patterns and features, they use convolutional techniques along the temporal dimension.

$$(f * g)(t) = \sum_{a=1}^{m} f = 1.g(t-a) \quad (2)$$

Conv1D Layer: The layer here applies the input sequence's 1D convolution technique. In this case, f represents the input sequence, g is the kernel, or filter, The location is tand the filter size is m.

3.3.2. Bidirectional Long Short-Term Memory (LSTM) Networks:

Recurrent neural networks (RNNs) of the BILSTM type combine data from previous and next time steps. Tasks with temporal dependencies and sequences are a good fit for it.

LSTM Cells:

To regulate the information flow, these units contain gates and memory cells.

$$f(t) = \sigma(W_f . [h_{t-1}, x_t] + b_f \quad (3)$$

$$i(t) = \sigma(W_i . [h_{t-1}, x_t] + b_i \quad (4)$$

$$o(t) = \sigma(W_i . [h_{t-1,x_t}] + b_i \quad (5)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot c_t \widetilde{c_t} \quad (6)$$

Where, f_t is forget gate, C_t is the cell state, O_t is the output gate.

3.3.3. Attention mechanism

The Attention layer receives the outcome from the LSTM, which comprises data from the complete sequence processed in both directions. By giving certain sequence segments distinct weights, the Attention layer concentrates on various segments of the LSTM's output sequence. As a result, the model can highlight the characteristics that are most important for forecasting the target variable. The attention weights computation may be more stable by using the use scale parameter when the use scale parameter in the Attention layer is set to True, which indicates that the attention scores are scaled.

$$a_t = \frac{\exp(score(H_t, \ H'))}{\sum_{t'} \exp(score(H_{t'}, H'}) \quad (7)$$

where score(·) is a function that determines the relevance of the output of each time step, output H t with regard to some query H ′ (which may be the final hidden state, another learned vector, or a portion of the sequence itself) at are the attention weights.

This model combines CNN and BILSTM with Attention mechanism for modeling. It involves time and space dependencies, crop production prediction.

Combination:

The temporal sequence is used by the 1D CNN to extract features. The CNN's properties are used by the BILSTM to identify temporal dependencies.

$$\text{Hybrid Output} = \text{BiLSTM} + \text{Attention (1D Conv (Input))} \quad (8)$$

The above equation shows the combined model, Here the prediction done by feeding the 1D CNN's output into the BILSTM.1D CNN is used to extract features from numerical sequences, and the BILSTM receives its output to extract temporal dependencies and forecast crop yield. The model can concentrate on particular segments of the sequence which are most pertinent to the current job.

ALGORITHM CNN-BiLSTM-Attention

As our Deep Learning Model takes center stage, the Evaluation Phase becomes a critical act in gauging its prowess. Below is a step-by-step algorithm, akin to a musical score, guiding us through the intricate process of assessing the model's performance.

Algorithm : CNN-BiLSTM Attention Mechanism

Input:

//Crop and fertilizer dataset

CROP – (Ph, Temperature, Humidity, Class)

FERTILIZER – (Crop, Nitrogen, Phosphorus, Potassium, PH)

Output:

Suitable Crop and Fertilizer recommendation.

Phase I: Preprocess the data

1. Label encoding: Convert the categorical value to numerical data using Label encoding.

2. Normalize the data using method standard scalar.

3. Split the dataset into X train, X test, Y train, Y test.

Phase II Modeling

4. Configure and apply 1D CNN, Conv (X_train)= Relu(WC *Xtrain+bc).

5. Configure and apply BILSTM layers, BILSTM (pool(conv (Xtrain)).

6. Implement Attention layer, Attention (BILSTM=∑iαihi).

7. Compile the model specifies the loss function, optimizer and metrics.

Phase III Prediction

8. Test the model, Model = argmin £ (X_{test} , Y_{test}).

9. Evaluate the model performance,

$$\text{ACCURACY} = \frac{\text{Number of correct predictions}}{\text{Total Predictions}}$$

Calculate Precision, Recall and f measure

4. Experimental Setup

A. Software Tool

Python 3.9 software and Jupyter Notebook is previously known as IPython Notebooks is used for implementation in this research work. It is a general-purpose programming language with extraordinary capabilities and key features of python an easy-to-understand syntax; it includes keras, NumPy, pandas and several libraries.

B. Evaluation Measure

Accuracy: The ratio of correctly predicted Instances to the total instances.

Formula: $Accuracy = \frac{Number\ of\ Correct\ Predictions}{Total\ Number\ of\ Predictions}$ (9)

Interpretation: High accuracy indicates the overall correctness of the model.

Precision: The ratio of true positive (TP) predictions to the total predicted positives.

Formula: $Precision = True\ Positives(TP) + False\ Positives$ (10)

Interpretation: Precision measures the accuracy of positive predictions.

Recall (Sensitivity): The ratio of true positive predictions to the total actual positives

Formula: $Recall = True\ Positives + False\ Negatives$ (11)

Interpretation: Recall measures the ability of the model to capture all positive instances.

F-Measure: It produces a single rating that takes into consideration memory and precision issues in a single figure.

It is the harmonic mean of precision and recall.

Formula: $F = 2 \times \frac{Precision \times Recall}{Precision + Recall}$ (12)

Interpretation: F-measure value achieves high when precision and recall value are high

5. Result and Discussion

Agriculture datasets are utilized for forecasting crops and fertilizers. Calculations are made for the measured values and accuracy. An explanation is provided for the confusion matrix along with statistical measurement data.

5.1 Confusion Matrix

An N x N matrix is employed for the assessment of the efficacy of a classification model, with N denoting the quantity of target classes. The matrix juxtaposes the factual objective values with the prognostications of the machine learning model. A confusion matrix serves as a tabular representation utilized for the characterization of a classification algorithm's performance.

Fig.7. Visualize confusion matrix of crop dataset and Fig.8. Visualize confusion matrix for fertilizer dataset, rows represent the actual classes (ground truth) and Columns represent the predicted classes of the model. The diagonal elements, which extend from the top-left to the bottom-right, serve as a depiction of the precise forecasts made for each class. To illustrate, the presence of the number 22 in both the first row and first column signifies that there were 22 instances where the true class was correctly anticipated as the first class. In contrast, the off-diagonal elements signify the instances of misclassification. To exemplify, the inclusion of the value 1 in the first row and second column indicates the occurrence of a single instance where class 1 was predicted instead of class 0.

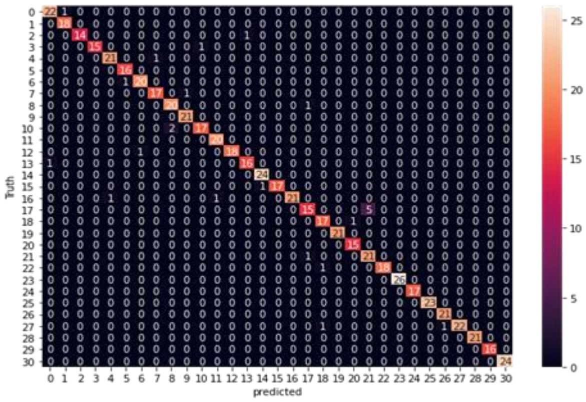

Fig.7.Confusion matrix for crop dataset

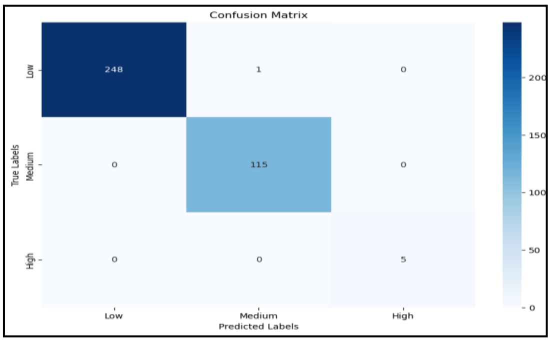

Fig.8. Confusion Matrix for fertilizer dataset

The dataset contains 31 distinct crops; hence the matrix's dimension is 31 × 31. The number of cases that were correctly anticipated for each crop type is represented by every cell along the diagonal. In this matrix, the first crop type had 22 cases that were accurately identified; the second crop type had 18, and so on. Erroneously classified cases are shown by cells that are off the diagonal. Here, one case of the first crop type was mistakenly identified as the second, and one case of the eighth crop type as the ninth. There was no incorrect categorization for those class pairs since there are a lot of zero cells outside the diagonal. This suggests that the model has excellent discriminative capacity for those particular classes.

5.2. Performance metrics

A better comprehension of the prediction findings is offered by the evaluation measures, which include accuracy, precision, recall, confusion matrix, and F1 score.

5.2.1 Classification Errors

Regression model performance is often assessed using the metrics known as Root Mean Squared Error (RMSE), Mean Squared Error (MSE) and Mean Absolute Error (MAE). Table.4 shows error values of various classifiers are discussed when compared to the conventional proposed ensemble CNN-BILSTM-ATTENTION classifiers error values measure low error value.

Table 4. Error Values Comparison of Proposed CNN-BILSTM-ATTENTION for crop dataset

Algorithm	MSE	RMSE	MAE
Linear Regression	0.35	0.59	0.42
Decision Tree	0.4	0.63	0.48
Random Forest	0.28	0.53	0.38
CNN	0.4	0.44	0.46
CNN-LSTM	0.31	0.32	0.3
CNN-BILSTM-ATTENTION	0.04	0.001	0.03

Figure 9 shows error values of various classifiers are discussed when compared to the conventional proposed ensemble CNN-BILSTM-ATTENTION classifiers error values measure low error value.

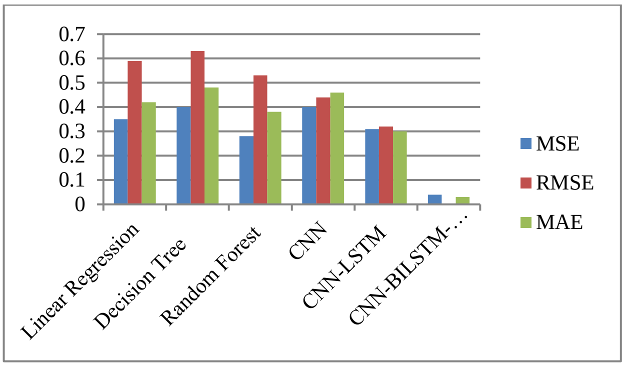

Fig. 9. Comparison of Error metrics with other conventional algorithms for crop dataset

Table.5 Error values of various classifiers are discussed when compared to the conventional proposed ensemble CNN-BILSTM-ATTENTION classifiers error values measure low error value.

Table 5. Error Values Comparison of Proposed CNN-BILSTM-ATTENTION for fertilizer dataset

Algorithm	MSE	RMSE	MAE
Linear Regression	0.25	0.50	0.35
Decision Tree	0.30	0.55	0.40
CNN	0.20	0.45	0.35
CNN-LSTM	0.17	0.41	0.29
CNN-BILSTM-ATTENTION	0.02	0.01	0.01

Fig.10. shows visual Comparison of Error metrics with other conventional algorithms for fertilizer dataset when compared to the other conventional algorithms proposed CNN-BiLSTM-Attention measures low error value

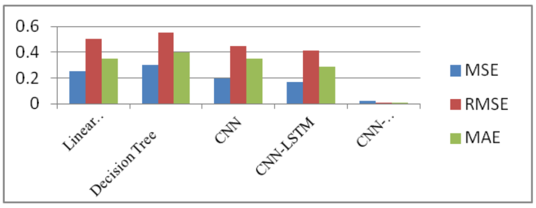

Fig.10. Comparison of Error metrics with other conventional algorithms for fertilizer dataset

Based on the relevant crop characteristics, the work uses neural network models like CNN and BILSTM-ATTENTION to forecast crop production. MSE, MAE and RMSE (root mean square error) were employed to evaluate the effectiveness of the neural network model.

5.4 ROC CURVE

A visual depiction of the binary categorization model's performance over various discrimination thresholds is called a Receiver Operating Characteristic (ROC) curve. For different threshold settings, it shows the True Positive Rate (Sensitivity) versus the number of false positives (1 - Specificity). A graphical depiction called the Receiver Operating Characteristic (ROC) curve. It offers a visual comparison of true positive rates and false positive rates across various thresholds, making it a vital tool for the assessment of diagnostic tests.

TRUE POSITIVE RATE

Plotted on the y-axis is the True Positive Rate (TPR), sometimes referred to as Sensitivity or Recall. It shows the percentage of real positives that the model successfully detected.

$$TPR = \frac{TP}{TP+FN} \quad (12)$$

FALSE POSITIVE RATE

On the x-axis, the False Positive Rate (FPR) is shown. It shows the percentage of real negatives that the model mistakenly classifies as positives.

$$FPR = \frac{FP}{FP+TN} \quad (13)$$

The figure 11 represents the roc curve of proposed CNN-BILSTM-ATTENTION mechanism .It contains 31 unique class.ROC values is calculated for each class.

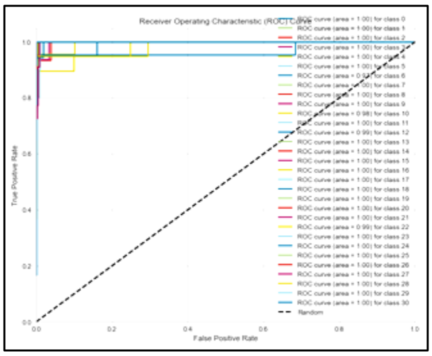

Fig.11. ROC curve for Proposed Work CNN-BILSTM-attention mechanism using crop dataset**.**

The Curve: At different threshold settings, the ROC curve shows TPR vs. FPR. Based on the output probabilities of the model, the threshold is a value that distinguishes between a positive and a negative forecast. The classifier's sensitivity and specificity are altered by adjusting the threshold, which shifts the point on the ROC curve. It comprises 31 crops, considered as 31 class then true positive rate and false positive rate are calculated.

Table 6 shows Performance values of proposed CNN-BILSTM-ATTENTION for Crop Dataset and Table 7 shows Performance values of proposed CNN-BILSTM-ATTENTION for fertilizer dataset.

Table 6. Performance values of proposed CNN-BILSTM-ATTENTION for Crop Dataset

Classifier	Precision	Recall	F-Measure
SVM	0.66	0.65	0.61
Naïve Bayes	0.78	0.69	0.68
Decision Table	0.66	0.73	0.67
AdaBoost	0.37	0.59	0.45
Random Forest	0.81	0.82	0.81
J48	0.68	0.7	0.67
CNN	0.9	0.91	0.9
CNN-LSTM	0.94	0.92	0.92
CNN-BILSTM	0.96	0.95	0.96
CNN+BILSTM+Attention	0.98	0.98	0.98

Table 7. Performance values of proposed CNN-BILSTM-ATTENTION for Fertilizer Dataset

Classifier	Precision	Recall	F-Measure
SVM	0.7	0.68	0.69
Naïve Bayes	0.75	0.72	0.73
Decision Tree	0.65	0.7	0.67
AdaBoost	0.8	0.77	0.78
Random Forest	0.85	0.83	0.84
J48	0.68	0.66	0.67
CNN	0.88	0.89	0.88
CNN-LSTM	0.92	0.9	0.91
CNN-BILSTM	0.93	0.92	0.92
CNN+BILSTM+Attention	0.99	0.99	0.99

Fig.12. shows precision, recall and F-measures of crop datasets and Fig.13. Shows precision, recall and F-measures of fertilizer datasets.

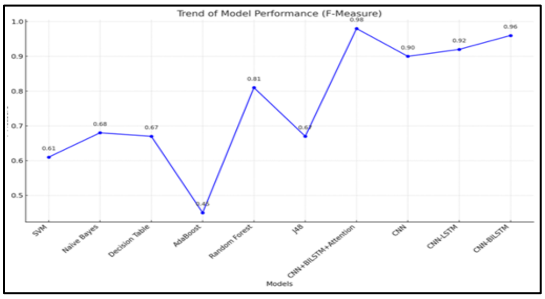

Fig.12. Comparison proposed method with conventional methods of f-measure crop dataset

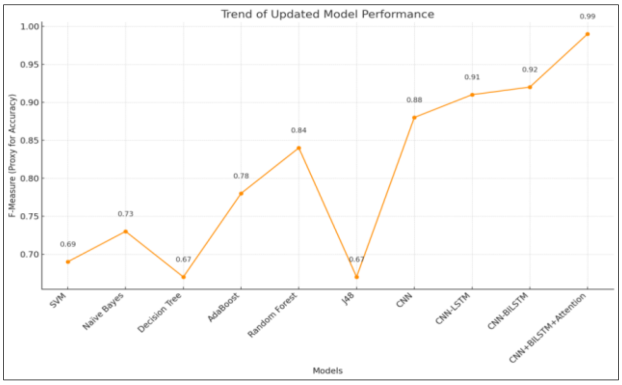

Fig.13. Comparison proposed method with conventional methods of f-measure for Fertilizer dataset.

It reveals that Proposed hybrid model gives higher accuracy when compared to other conventional algorithms. Table.8 describes the performance analysis of each crop. Each crop is considered as a class precision, recall and f1 score and support values are calculated.

Table.8 Performance analysis of each class for crop Dataset

S.no	Crop	precision	recall	f1-score
1	Adzuki Beans	0.96	0.96	0.96
2	Black gram	0.95	1.00	0.97
3	Chickpea	1.00	1.00	1.00
4	Coconut	1.00	1.00	1.00
5	Coffee	1.00	1.00	1.00
6	Cotton	1.00	1.00	1.00
7	Ground Nut	1.00	1.00	1.00
8	Jute	1.00	0.95	0.97
9	Kidney Beans	0.95	0.95	0.95
10	Lentil	0.95	1.00	0.98
11	Moth Beans	1.00	0.95	0.97
12	Mung Bean	0.95	1.00	0.98
13	Peas	1.00	1.00	1.00
14	Pigeon Peas	1.00	0.94	0.97
15	Rubber	0.96	1.00	0.98
16	Sugarcane	1.00	0.94	0.97
17	Tea	1.00	0.96	0.98
18	Tobacco	0.91	1.00	0.95
19	Apple	1.00	0.94	0.97
20	Banana	1.00	1.00	1.00
21	Grapes	0.94	1.00	0.97
22	Maize	1.00	0.95	0.98
23	Mango	0.96	1.00	0.98
24	Millet	1.00	0.96	0.98
25	Muskmelon	1.00	1.00	1.00
26	Orange	1.00	1.00	1.00
27	Papaya	1.00	1.00	1.00
28	Pomegranate	1.00	1.00	1.00
29	Rice	1.00	1.00	1.00
30	watermelon	1.00	1.00	1.00
31	Wheat	1.00	1.00	1.00
	Accuracy	0.98	0.98	0.98
	macro avg	0.98	0.98	0.98
	weighted avg	0.98	0.98	0.98

The above result prescribes that proposed work shows better accuracy of 98% for crop dataset and 99% for fertilizer dataset when compared to other conventional algorithm. This proposed method gives 7% better accuracy for crop dataset and 10% better accuracy for fertilizer dataset when compared with conventional algorithms.

Conclusion and Future Enhancement

In order to recommend suitable crop and estimate amount of fertilizer, we investigated the possibilities of ensemble modeling in this work by merging Convolutional neural network (CNN), Bidirectional Long Short-Term Memory networks (BILSTM), and attention mechanisms to enhance Crop Yield Prediction through Ensemble Modeling. By integrating these cutting-edge methods, a reliable and precise model for estimating crop production was created by attempting to capture complex patterns in statistical agricultural statistics. Performance metrics such as confusion matrix, Precision, Recall, F1 score, ROC curve have been examined and measured for CNN-BILSTM ATTENTION mechanism which gives better accuracy. Where, Proposed CNN-BILSTM with ATTENTION mechanism scores 98% and 99% accuracy for crop and fertilizer data respectively.

Hyper parameter tweaking and sophisticated feature engineering approaches may improve model accuracy even more. Furthermore, investigating domain adaptation and transfer learning techniques may help the model more successfully adapt to various agricultural settings and geographical areas, increasing the system's versatility and scalability.

References

[1] Shanmugam, J. Rethnaraj, and G. Mani, "Selection of crop varieties and yield prediction based on phenotype applying deep learning," International Journal of Electrical and Computer Engineering (IJECE), vol. 13, no. 6, pp. 6806-6816, 2023.

[2] K. C., Srinivas, "Crop Yield Prediction Using Machine Learning," International Journal for Research in Applied Science and Engineering Technology, vol. 9, pp. 553-555, 2021, doi: 10.22214/IJRASET.2021.37413.

[3] B. G. Chaitra, B. M. Sagar, N. K. Cauvery, and T. Padmashree, "Forecasting Crop Yield with Machine Learning Techniques and Deep Neural Network," in International Conference on Intelligent Vision and Computing, Cham, Springer Nature Switzerland, Nov. 2022, pp. 532-542.

[4] S. Bharti, P. Kaur, P. Singh, C. Madhu, and N. Garg, "Crop Yield Prediction Using CNN-LSTM Model," in 2022 IEEE Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI), IEEE, Dec. 2022, pp. 1-5.

[5] Z. Li, R. Xu, X. Luo, X. Cao, and H. Sun, "Short-term wind power prediction based on modal reconstruction and CNN-BiLSTM," Energy Reports, vol. 9, pp. 6449-6460, 2023.

[6] M. Kulyal and P. Saxena, "Machine Learning approaches for Crop Yield Prediction: A Review," in 2022 7th International Conference on Computing, Communication and Security (ICCCS), IEEE, Nov. 2022, pp. 1-7.

[7] P. Saini, B. Nagpal, P. Garg, and S. Kumar, "CNN-BI-LSTM-CYP: A deep learning approach for sugarcane yield prediction," Sustainable Energy Technologies and Assessments, vol. 57, p. 103263, 2023.

[8] D. Paudel et al., "Machine learning for regional crop yield forecasting in Europe," Field Crops Research, vol. 276, p. 108377, 2022.

[9] S. S. Olofintuyi, E. A. Olajubu, and D. Olanike, "An ensemble deep learning approach for predicting cocoa yield," Heliyon, vol. 9, no. 4, 2023.

[10] K. Archana and K. G. Saranya, "Crop yield prediction, forecasting and fertilizer recommendation using Data mining algorithm," International Journal of Computer Science Engineering (IJCSE), vol. 9, no. 1, pp. 76-79, 2020.

[11] B. S. Rao et al., "A Novel Approach for Crop Yield Prediction based on Hybrid Deep Learning Approach," in 2023 8th International Conference on Communication and Electronics Systems (ICCES), IEEE, June 2023, pp. 1310-1315.

[12] https://github.com/Gladiator07/Harvestify.[13] P. S. S. Gopi and M. Karthikeyan, "Red fox optimization with ensemble recurrent neural network for crop recommendation and yield prediction model," Multimedia Tools and Applications, vol. 83, no. 5, pp. 13159-13179, 2024.

[13] S.Iniyan, M.S.Raja, R.Srinivasan, C.Santhanakrishnan and A.Srinivasan,"Crop and Fertilizer Recommendation System Applying Machine Learning Classifiers," in 2023 International Conference on Emerging Research in Computational Science (ICERCS),pp.1-6,doi: 10.1109/ICERCS57948.2023.10433972,2023.

[14] D.Anantha Reddy, Bhagyashri Dadore and Aarti Watekar, "Crop recommendation system to maximize crop yield in ramtek region using machine learning", International Journal of Scientific Research in Science and Technology, vol. 6, no. 1, pp. 485-489, 2019.

[15] N.Mittal and A.Bhanja,"Implementation and Identification of Crop based on Soil Texture using AI," in 2023 4th International Conference on Electronics and Sustainable Communication Systems(ICESC)pp.1467-1471, doi:10.1109/ICESC57686.2023.10192937,2023.

[16] P. Prabhu and N. Anbazhagan, "FI-FCM Algorithm for Business Intelligence," in Mining Intelligence and Knowledge Exploration, R. Prasath and T.

Kathirvalavakumar, Eds. Cham: Springer, 2013, pp. 518-528.

[17] P. Prabhu, "Improving SVM Classifier Model Using Tree Structured Parzen Estimator Optimization for Crop Prediction," Journal of Theoretical and Applied Information Technology, vol. 100, no. 22, pp. 6808-6818, Nov. 2022, ISSN: 1817-3195.

[18] M. Sivakami and P. Prabhu, "A Comparative Review of Recent Data Mining Techniques for Prediction of Cardiovascular Disease from Electronic Health Records," in Intelligent Data Communication Technologies and Internet of Things, D. Hemanth, S. Shakya, and Z. Baig, Eds. Cham: Springer, 2019, pp. 477-484.

[19] S. M. Beram, P. Harikumar, P. Indrajit, and P. Prabhu, "A Machine Learning Based Framework for Preprocessing and Classification of Medical Images," ECS Transactions, vol. 107, no. 1, April 2022, ISSN: 1938-6737, doi: 10.1149/10701.7589ecst/.