Mapana - Journal of Sciences
2023, Vol. 22, Special Issue 2, 147-169
ISSN 0975-3303 | https://doi.org/10.12723/mjs.sp2.9

Crop Prediction and Recommendation Using Ensemble of DL Models

Dr B Subbulakshmi*, Dr.M Nirmaladevi†, RRithani‡

Abstract

Indian agriculture is the primary source of livelihood for millions of people in India It contributes significantly to the country's economy and is a crucial sector for the overall development of the country. Agriculture in India is characterized by a diverse range of crops, soil types, and agro -climatic conditions due to the country's vast size and varied topography The agriculture sector in India employs over 50% of the workforce, and it is the primary source of income for rural households The main idea is to suggests an additional ensemble model that can be implemented for effective and speedy prediction and recommendation of crops To employ this system, data is collected having nearly 8 distinct features from various databases and finalized 2201 instances. The data collected primarily focuses on climatic conditions such as temperature, rainfall,

* Assistant Professor, Department of Computer Science and Engineering, Thiagarajar College of Engineering, Madurai, India , email - bscse@tce.edu

† Assistant Professor, Department of Computer Science and Engineering, Thiagarajar College of Engineering, Madurai, India, email - mnit@tce.edu

‡ Student. Department of Computer Science and Engineering, Thiagarajar College of Engineering. Madurai, India, email-rithanirameshl234@gmail.com

crop type, and soil features, particularly the ratio of nitrogen, potassium, and phosphorous levels Extensive research indicates that algorithms such as Neural Networks and XGBoost are highly effective and accurate in developing crop yield prediction models After conducting thorough research, it is concluded that the ensemble of XGBoost and MLP Classifier algorithms provides the most significant accuracy of 99.39 %. This strategy aims to predict crop yield based on historical data, which helps in recommending which crop should be cultivated depending on the soil type and weather conditions of the field location

Keyword: crop prediction, XGboost, MLP classifier. Ensemble

1. Introduction:

The agriculture sector in India employs over 50% of the workforce, and it is the primary source of income tor rural households The mam crops grown in India are rice, wheat, pulses, oilseeds, cotton, jute, sugarcane, tea, and coffee The country is the second-largest producer of rice, w[r]heat, and sugarcane in the w'orld However, Indian agriculture faces several challenges, such as inadequate irrigation facilities, low productivity, land degradation, and insufficient credit facilities Additionally, small and marginal farmers face difficulties in accessing markets due to inadequate transportation infrastructure and a lack of access to modern technologies. Despite the challenges, Indian agriculture has show'n resilience and has the potential to achieve sustainable growth in the future. The adoption of modern technologies, including precision agriculture, smart farming, and the use of biotechnology., can help overcome the challenges faced by Indian agriculture and ensure food security for the growing population Various elements have affected the fitness of agriculture in India One such method is using pre-trained Deep Learning models and applying ensemble on them so as to increase their accuracy and predictive ability. Recommendation crop one most important area in precision agriculture

One popular approach to crop prediction is ensemble modeling, which involves combining the predictions of multiple models to improve overall accuracy Ensemble models can be created using a variety of techniques, including bagging, boosting, and stacking Ensemble modeling has been shown to be effective in crop prediction, particularly in situations where the data is noisy or where there is a high degree of variability in the crop yields. By combining the predictions of multiple models, ensemble modeling can help improve the accuracy of crop yield forecasts, leading to better decision-making and increased agricultural productivity.

In summary, crop prediction is a crucial aspect of agriculture that can benefit from the use of ensemble modeling techniques By combining the predictions of multiple models, farmers and policymakers can make more informed decisions that can lead to higher crop yields and greater efficiency in the agricultural sector

2. Literature Survey:

Crop prediction and recommendation is an important area of research that aims to provide farmers with accurate predictions of crop yields and recommendations on crop management practices. A literature survey of this field reveals that there have been several studies conducted to address this issue Researchers have employed different approaches, including machine learning, data mining, and remote sensing, to develop models for crop prediction and recommendation. Some studies have focused on specific crops, while others have developed models that are applicable to multiple crops. Overall, the literature suggests that the use of advanced Technologies and data-driven models can improve crop prediction and recommendation accuracy, which can ultimately lead to better crop management and higher yields The detailed inferences from each literature are well briefed below.

G Abirami, et al(2022) This journal article provides an in-depth analysis of various machine learning models that support crop recommendation prediction. The study recommends using ensemble techniques to predict the crop type based on a set of given parameters using either supervised or unsupervised learning techniques In order to obtain the desired output parameter, an appropriate feature must be generated from a set of variables that will depict the output using the given input parameters This involves combining two or more machine learning algorithms to create an ensemble. w'hich enhances the accuracy of crop yield prediction Limitations The work has never ensembled w'ith DL models like MLP,CNN,ANN involving neural networks .

Priyadharshini A,et al(2021) The accuracy of crop forecasting in agriculture relies heavily on soil and environmental factors, such as temperature, humidity, and rainfall. However, with the rapid changes in the environment, it has become challenging for farmers to manage crop selection, development monitoring, and harvest timing. Traditional prediction methods have been replaced by machine learning approaches, which have been utilized in this study to calculate agricultural yield To ensure high precision in machine learning models, effective feature selection techniques are required to transform raw' data into a clear and understandable format Only the most relevant data aspects should be used in defining the model's final output to decrease redundant data and improve accuracy.

Optimal feature selection is critical to ensure that only the most important features are included in the model, preventing needless complexity and increased time and space complexity. The research findings indicate that ensemble techniques provide more accurate predictions compared to current categorization techniques **Limitations:** The didn't summarize about type of ensemble modelling used in their model

Janmejay Pant, R.P. Pant, et al (2021) This study utilizes Machine Learning techniques to develop a model trained to identify patterns in data for crop prediction, specifically for four of the most cultivated crops in India: Maize, Potatoes. Rice (Paddy), and Wheat. After data pre-processing, the final dataset consists of crop, country, year, yield value, average rainfall, pesticides, and average temperature as features. The study applies several Machine Learning models for crop yield prediction, including Gradient Boosting Regressor, Random Forest Regressor, SVM, and Decision "t ree Regressor, and compares their performance using the Rooted Square value The pre-processed data is split into training and testing datasets, with a 80:30 ratio. The evaluation matrix is calculated based on the $R^{A}2$ (coefficient of determination) regression score function The $R^{A}2$ score indicates how^{r} well data points fit in a curve or line The Decision free Regressor model achieves the highest $R^{A}2$ score of 96%, indicating the best fit to the observed data. Limitations: The study does not take soil properties like N,P_ an K levels to make prediction more efficient, and the study takes only minimal crop item

Kalaiselvi et al(2021) The proposed system in this study recommends crops based on attributes such as N, P, K_7 pH, Temperature, Humidity, and Rainfall The dataset consists of 2200 instances and S attributes, and nearly 22 different crops are recommended based on different attribute combinations.

The system utilizes supervised learning methods to identify the optimal model and employed machine learning algorithms in WEKA for classification, including multilayer perceptron rules-based classifier JRip and decision table classifier The primary objective is to develop a model that can predict high-yield crops and support precision agriculture The proposed system incorporates loT technology and essential agricultural measurements to achieve this goal The collected data is converted it into comma-delimited excel format (CSV) for compatibility with WHKA Then the system use supervised learning methods to preprocess the data and perform classification with selected classifiers The performance characteristics are noted and tabulated after completing the classification process. Then the time taken to build the model using the nominal to binary attribute selection preprocessing method is reduced The optimum model accurately predicts crop recommendations based on the dataset and seven different attributes, achieving an accuracy performance percentage of 98 2283% to 88.5909% for multilayer perceptron and JRip classifier The receiver operator characteristics of 0 991 to 998 for the Lazy category classifier decision table and multilayer perceptron are noteworthy **Limitations:** The paper never gave a detailed explanations about the methodology and working S. P Raja et a I (2022) In this work, crop forecast in agriculture is important and heavily dependent on the soil and environmental factors, such as temperature, humidity and rainfall. In the past, farmers had control over the selection of the crop to be grown, the monitoring of its development and the timing of its harvest. It is now challenging for the farming community to carry on as a result of the quick changes in the environment As a result, machine learning approaches have increasingly replaced traditional prediction methods This work has used a number of these methods to calculate agricultural yield Using effective feature selection techniques to transform the raw data into an easily understandable format is essential to ensuring that a particular machine learning (ML) model operates at a high level of precision Only data aspects that are significantly relevant in defining the model's final output should be used, which will decrease redundant data and improve the model's accuracy. In order to ensure that only the most important features are included in the model, it is important to be using optimal feature selection The model will become needlessly complex if combine every characteristic from the raw data without first examining their function in the model-building process. Also, the time and space complexity of the ML model w'ill increase w'ith the addition of new characteristics that have little impact on the model's performance The findings show[7] that an ensemble technique provides more accurate prediction than the current categorization technique **Limitations** : The limitation of this system is that explanation about the feature extraction step is not detailed.

Archana, ket al(2020) In this paper The accurate prediction and forecasting of crop yield are crucial tor agricultural stakeholders and can be achieved through the application of machine learning techniques. In many cases, farmers may lack knowledge about soil nutrition and composition, leading to low crop yields To address this issue, The proposed system have developed a system that focuses on important factors such as macronutrients (NPK), pH, electrical conductivity in the soil, and temperature to provide appropriate crop recommendations The proposed system integrates crop rotation, crop yield prediction and forecasting, and fertilizer recommendation to create a collaborative system that supports informed decision-making Then utilized an ensemble classifier algorithm based on voting to suggest suitable crops using an agricultural dataset This system has the potential to increase agricultural production by enabling accurate yield prediction and forecasting while also improving soil fertility through periodic crop rotation By facilitating farmer-friendly fertilization decision-making, this system can significantly enhance agricultural productivity, system achieved an accuracy of around 92% **Limitations** : This system didn't use Deep learning models to enhance the prediction and recommendation

Ujjainia, Shikha et al(2021) The integration of technology with crop yield prediction has brought significant advancements in global agricultural production. Machine learning has played a crucial role in optimizing the situation for farmers and the agricultural industrv. The use of various algorithms has increased the effectiveness of the prediction methodology, resulting in minimal deviation in crop production. The research suggests the integration of three models, which are typically programmed separately in agricultural devices Information technology has proven to be effective in improving the agricultural industrv through different functions The use of ensemble algorithms has made the agricultural industry more competent in maintaining the expected crop production levels Limitations: The model proposed in this paper never compared the developed ensemble model with the existing ones and never gave a detailed explanation about the type of ensemble technique used

M. keerthana, et al (2021)T his study explores the use of ensemble techniques for predicting crop yields based on area characteristics. Machine learning algorithms can predict outcomes based on input parameters using both supervised and unsupervised techniques The goal is to develop a function that accurately represents the output based on the given input variables or parameters. The study uses an ensemble of two machine learning methods, w'hich improves the accuracy of crop yield prediction. The study has collected data from multiple databases using a search approach, resulting in 28242 instances and nearly 8 features. The key factors considered are related to climatic factors, such as temperature, rainfall, and crop ty pe After analyzing various algorithms, the study found that decision trees and neural networks are the most popular for these types of models Decision trees can be adjusted using parameters such as maximum depth and n-estimators to improve results. The study concludes that an ensemble of decision tree and AdaBoost regressors provides significant accuracy Limitations : Ultimately, this approach provides recommendations for which crops to grow’ based on local weather conditions and historical data The system never use any soil properties for crop prediction and recommendation .

Barbosa, Alexandre et a!(2020) The study suggests that by adding uncertainty quantification to the previously suggested Convolutional Neural Network (CNN) model for yield prediction and then proposing a risk-averse optimization method on top of it, the study address both issues in this work. Under the Deep Ensemble framew’ork, the study modifies the CNN architecture so that the predictive model produces a probability distribution rather than a single value This model is then used by a gradient-based optimization method to identify crop input maps that maximize predicted net revenue while abiding by risk restrictions. The study demonstrate that the new model not only quantifies uncertainly’ but also outperforms its predecessor’s predictions in terms of performance The optimization algorithm s tests have revealed an up to 6 4% increase Limitations: The study doesn’t involve soil properties like N, P. an R levels to make prediction more efficient and takes only minimal crop item.

Methodology

Crop yield prediction is a system which is used to predict the suitable crop based on the parameters considered in the dataset.

Data Collection:

The dataset is extracted from Kaggle website for training The dataset in our system has 8 features of the soil like N, P. R. and climatic conditions like rainfall humidity, temperature, pH & crop type Ensemble of algorithms improves the accuracy rate rapidly The process consists of the following modules:

Pre-processing:

In the dataset there are one categorical column. Categorical data are the attributes that where the values are in labelled format. There is a fixed set which doesn't exceed the limit of the possible values, here in our case, crop name Multiple machine learning algorithms cannot operate on labelled data directly In order to operate the label, data should be converted to numerical data using any of the encoding techniques. So, we have applied Label Encoder to coven the categorial data to numeric ones by labelling them with a unique representative integer, For example we have about 21 types of crops present in the dataset from apple to water melon. The unique integer LCT represents LLApple“ ,T represents “Banana^ & so on

Data Exploration:

This is done by exploring the relationships between the columns of the dataset A best way to check correlations among columns is by visualizing the correlation matrix using a heat map. Pearson Correlation is used to visualize the correlation matrix as a heat map

Scaling:

The attributes in the data frame are highly different in range, magnitudes, and units. The features w ith high values will result in huge complexity and confusion rather than features with smaller values To overcome this, we need to bring all features to same range of magnitudes This can be done by scaling the data with MinMaxScaler or StandardScaler.

Training and Testing Phase:

The crucial step on data pre-processing is the training and testing data. Here, we have considered 70/30 ratio which is the most considered ratio This predicts how w'ell the test data is trained giving the most accurate outcome.

Model Selection & Comparisons:

Before deciding to choose an algorithm to use, evaluation should be done to choose the best one that fits for the specific dataset In our dataset we have 7 features including N, P. K. humidity, rainfall, pH, temperature and crop type Of these feature the target label is the crop type. From the base papers collected we have selected 3 ML and 1 DL model for ensemble they are Logistic regression. Decision tree, Xtreme -Gradient Boosting and Multi-layer perceptron (MLP). And after selecting the models next step is to decide whether to use a classifier or a regressor. Regressor refers to predictive modeling problems that involve predicting a numeric value given an input. Classifier refers to predictive modeling problems that involve predicting a class label or probability of class labels for a given input Hence, The selected classifier for ensemble are :

- Logistic Regression
- Decision free

Here in Table 1 depicts the metric comparison of the selected models, We have taken accuracy metrics like : R2_Score

Mean Absolute brror(MAL)

- Mean Squared Error(MSE)
- Model's Accuracy

Table 1: Accuracy of the trained individual models

ML Models	R2_Score	MAE	MSE	Accuracy **
Logistic Regression	0 92419	0.31969	3.10151	0.96060
Decision free Classifier	0.89439	0.89696	8.4121	0.91212
XGBoost Classifier	0 98326	0 06969	0.6848	0.9909
MLP Classifier	0.98688	0.09848	0.95	0.9888
Gaussian Naive Baves Classifier	0 9886	0.08282	0.88282	0.99393
Ada Boost Classifier	0.13860	48636	35.2848	0.21212

Prediction:

By considering the above accuracy metrics XGBoost Classifier shows R2_Score value of 98 3% which is the most highest and efficient than the other predictive models. Then the next two places are for i–MLP Classifier " and LLNaive Bayes Classifier" So wre can train Xgboost along with MLP or Naive haves for ensemble But we have chosen MLP classifier for ensemble it w ith Xgboost This is to show that the combination of a pow erful ML Models like XGBoost can be combined with a DL model MLP that uses back propagation and neural networks We have rejected Ada Boost classifier for ensemble because it show's a very low7 r2_Score and also accuracy The main reason for its low' efficiency is that the models is not suitable for our chosen dataset

1. Working of the algorithms :

Working of XGBoost Classifier:

XGBoost (Extreme Gradient Boosting) is a popular machine learning algorithm for classification, regression, and ranking tasks It is based on the concept of ensemble learning, where multiple weak learners {decision trees) are combined to create a strong learner.XGBoost uses a gradient boosting framework that builds an ensemble of decision trees sequentially. In each iteration, the algorithm tries to fit a new tree to the negative gradient of the loss function of the previous iteration The output of each tree is added to the output of the previous trees, creating a final prediction. The XGBoost classifier is a specific implementation of the XG Boost algorithm for classification tasks. It uses a binary logistic regression loss function to optimize the decision trees. During training, the algorithm adjusts the weights of each observation to prioritize those that were misclassified in previous iterations Overall, XGBoost is a powerful and flexible algorithm that can achieve state-of-the-art performance on a w ide range of machine learning tasks, especially in the domains of structured data and feature engineering.

Working of MLP Classifier

The MLP (Multilayer Perceptron) classifier is a type of artificial neural network that is commonly used for classification tasks It consists of multiple layers of interconnected neurons, where each neuron is a simple computational unit that processes its inputs and produces an output .The MLP classifier works by taking input data and passing it through the layers of neurons, where each layer transforms the data into a higher-level representation. The final output layer produces a probability distribution over the possible classes, allowing the classifier to make a prediction The training of the MLP classifier involves adjusting the weights of the connections between the neurons to minimize the difference between the predicted output and the actual output. This is done using an optimization algorithm such as backpropagation

Ensemble Learning :

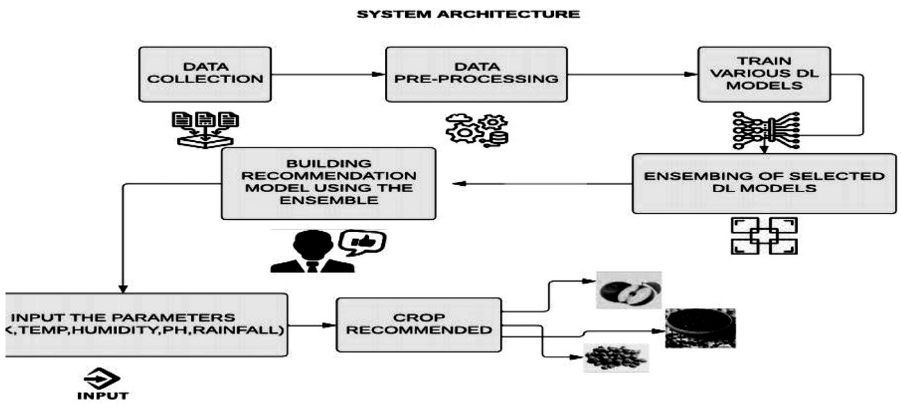
Ensemble learning is a machine learning technique that combines multiple models to improve the overall performance of the system Instead of relying on a single model to make predictions, ensemble learning creates a group or ‘’ensemble” of models that work together to provide more accurate results The ensemble then combines the predictions of these individual models to produce a final prediction, which is often more accurate than the predictions made by any individual model. There are 3 types of ensemble technique :

Boosting Bagging Voting

A voting classifier is an ensemble learning technique that combines the predictions of multiple individual classifiers to make a final prediction. The idea behind a voting classifier is to take advantage of the “wisdom of the crowd” by combining the predictions of several different models, each of which may have its own strengths and weaknesses. There are two main types of voting classifiers: hard voting and soft voting. In hard voting, the final prediction is based on the majority vote of the individual classifiers In soft voting, the final prediction is based on the average probability scores of the individual classifiers Here we used soft voting in our ensemble model

Workflow :

Fig No 1 depicts the workflow of the ensembled model :



In the work flow we have first collected dataset from kaggle site as a csv file. Then we analysed the dataset to select ML models that suits our data. The dataset has input data like N , P ,humidity, temperature, pHT rainfall and crop type In wrhich the target variable is crop type in our dataset, the dataset has 7 features and 2400 instances After spilting them into train and test data . The selected ML models are trained individually then we combine them using voting classifier In voting classifier technique , we then used soft voting to combine the predictions of XGBoost and MLPcIassifiers. Check the accuracy and if are outputs similar the system is ready for prediction else refeed into the DB

Implementation:

We have used Google Colab, sklearn. ensemble, matplotiib for visualization, Label Encoder and metrics packages are available in Python. These are the main packages which produced the desired results

Dominant packages used tor implementation:

Sklearn. Ensem ble library:

sklearn. ensemble is a module which includes two averaging algorithms, random forest and decision tree. This is used to combine the predictions of various base estimators built with a given algorithm to improve accuracy

MatplotLih:

is a comprehensive library used for visualizing the data. It makes use of NumPy and provides figures w'ith best quality and standard It is a library for visualizing 2 dimensional plots from data and also provides animated visualizations

7- Experimental Setup & Result:

Dataset: The Crop recommendation dataset was downloaded from Kaggle platform, the dataset is of CSV type having 8 columns & 2201 instances The 8 columns are:

- N - Nitrogen Level
- P - Phosphorous Level
- K. - Potassium Level
- Humidity
- Temperature
- Rainfall
- pH
- Crop Type

Experiment: This system considers weather patterns, soil properties, and crop type By observing all the prediction results of ensemble techniques, we can propose that ensemble of LXG boost Classifier with MLP Classifier ' gives the best results with 99% of Accuracy. This Lnsemble technique does not over fit the model nor under fit the model So, we can say this model is more suitable for the dataset The speed and efficiency of XG Boost w'hen combined with MLP classifier which excels in backpropagation and neural network the models become more efficient.

Recommendation Model:

The input data ie values of N,P.K. temperature, humidity, rainfall and pH are fed into the array in prediction system Before insering the input data we use scalar transformation to make the system understand the original values of the input rather the scale transformed value. Fig No 2 depicts a input data [34,76,80t20.65692,15 84573,7 985417,65.23811] refers the N, P, k temperature, humidity, rainfall and pH values respectively.For this input we get “”3” as output which refers to ‘Chickpea’

```
new_data = scaler.transform([[34,76,80,20.65692,15.84573,7.985417,65.23811]])
prediction = vc2.predict(new_data)
print("Prediction:", prediction)
```

```
Prediction: [3]
```

Fig No 2. The recommendation system

Combination Comparison Result:

Ensemble of Logistic Regression With Decision Tree Classifier :

Accuracy
0.9575757575757575

	precision	recall	f1-score	support
0	1.00	1.00	1.00	28
1	1.00	1.00	1.00	30
2	0.69	1.00	0.82	31
3	1.00	1.00	1.00	34
4	1.00	1.00	1.00	26
5	1.00	0.97	0.98	29
6	1.00	1.00	1.00	28
7	1.00	1.00	1.00	30
8	0.88	0.90	0.89	31
9	1.00	1.00	1.00	26
10	0.92	1.00	0.96	22
11	0.96	0.81	0.88	27
12	1.00	0.96	0.98	28
13	1.00	0.64	0.78	36
14	0.94	1.00	0.97	29
15	0.97	1.00	0.98	30
16	1.00	1.00	1.00	34
17	0.95	1.00	0.97	39
18	0.97	1.00	0.98	28
19	1.00	1.00	1.00	32
20	0.97	0.89	0.93	37
21	1.00	0.96	0.98	25
accuracy			0.96	660
macro avg	0.96	0.96	0.96	660
weighted avg	0.96	0.96	0.96	660

Fig No 3 E1

Ensemble of Naïve Bayes Classifier With Decision Tree Classifier

Accuracy
0.9863636363636363

	precision	recall	f1-score	support
0	1.00	1.00	1.00	28
1	1.00	1.00	1.00	30
2	1.00	1.00	1.00	31
3	1.00	1.00	1.00	34
4	1.00	1.00	1.00	26
5	1.00	1.00	1.00	29
6	1.00	1.00	1.00	28
7	1.00	1.00	1.00	30
8	0.90	0.90	0.90	31
9	1.00	1.00	1.00	26
10	0.96	1.00	0.98	22
11	1.00	0.96	0.98	27
12	1.00	1.00	1.00	28
13	1.00	0.97	0.99	36
14	0.97	1.00	0.98	29
15	0.97	1.00	0.98	30
16	1.00	1.00	1.00	34
17	0.95	1.00	0.97	39
18	1.00	1.00	1.00	28
19	1.00	1.00	1.00	32
20	0.97	0.92	0.94	37
21	1.00	0.96	0.98	25
accuracy			0.99	660
macro avg	0.99	0.99	0.99	660
weighted avg	0.99	0.99	0.99	660

Fig No 4 E2

Ensemble XGBoost Classifier With MLP Classifier:

(Best Achieved)

```
Accuracy
0.9939393939393939
              precision    recall  f1-score   support

           0       1.00      1.00      1.00        28
           1       1.00      1.00      1.00        30
           2       1.00      1.00      1.00        31
           3       1.00      1.00      1.00        34
           4       1.00      1.00      1.00        26
           5       1.00      1.00      1.00        29
           6       1.00      1.00      1.00        28
           7       1.00      1.00      1.00        30
           8       0.97      0.97      0.97        31
           9       0.96      1.00      0.98        26
          10       1.00      0.95      0.98        22
          11       1.00      1.00      1.00        27
          12       1.00      1.00      1.00        28
          13       0.97      1.00      0.99        36
          14       1.00      1.00      1.00        29
          15       1.00      1.00      1.00        30
          16       1.00      1.00      1.00        34
          17       1.00      1.00      1.00        39
          18       1.00      0.96      0.98        28
          19       1.00      1.00      1.00        32
          20       0.97      0.97      0.97        37
          21       1.00      1.00      1.00        25

    accuracy                           0.99       660
   macro avg       0.99      0.99      0.99       660
weighted avg       0.99      0.99      0.99       660
```

Fig No 5 E3

Ensemble Logistic Regression with MLP Classifier

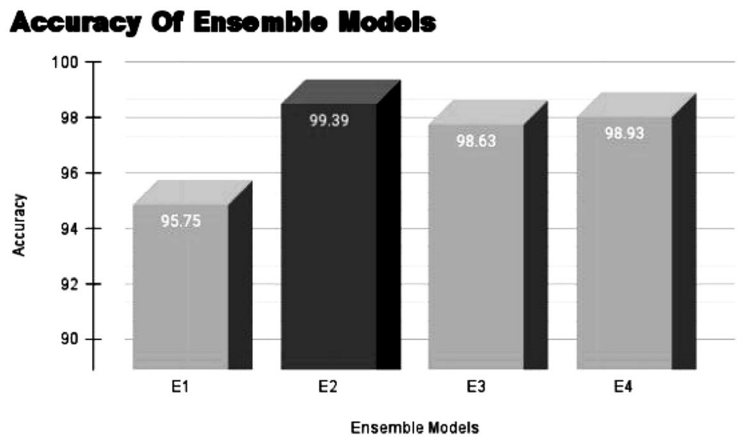
```
Accuracy
0.9893939393939394
              precision    recall  f1-score   support

           0       1.00      1.00      1.00        28
           1       1.00      1.00      1.00        30
           2       1.00      0.97      0.98        31
           3       1.00      1.00      1.00        34
           4       1.00      1.00      1.00        26
           5       1.00      1.00      1.00        29
           6       0.97      1.00      0.98        28
           7       1.00      1.00      1.00        30
           8       0.94      0.97      0.95        31
           9       0.96      1.00      0.98        26
          10       0.95      0.95      0.95        22
          11       1.00      0.96      0.98        27
          12       1.00      1.00      1.00        28
          13       0.97      1.00      0.99        36
          14       1.00      1.00      1.00        29
          15       1.00      1.00      1.00        30
          16       1.00      1.00      1.00        34
          17       1.00      1.00      1.00        39
          18       1.00      0.96      0.98        28
          19       1.00      1.00      1.00        32
          20       0.97      0.95      0.96        37
          21       1.00      1.00      1.00        25

    accuracy                           0.99       660
   macro avg       0.99      0.99      0.99       660
weighted avg       0.99      0.99      0.99       660
```

Fig No 6 E4

Table 2 : Gives the accuracy comparison of the models

Ensemble \lodeJs	**Accuracy r**	**Precision**	**Recall**	**FI Score**
Ensemble of Logistic Regression with Decision lree classifier (El)	0.95757	0 96	0.96	0.96
Ensemble of XG Boost classifier and MLP Classifier (E2)	0.99393	0.99	0.99	0.99
Ensemble of Naive baves classifier with decision tree classifier (E3)	0.93939	0.99	0.99	0.99
Ensemble of MLP classifier with Logistic regression (E4)	0.98939	0.99	0.99	0.99

Fig No 7 : shows the graph of the developed ensemble models



8, Conclusion & Future work :

To conclude we have collected the dataset from Kaggle having 8 columns and 22JO instances The 8 features are N s P , fC ,humidity,temperature.rainfall,pH and crop type in which crop type is a categorical type so applying it to Xgboost and MLP is challenging ?So . we have lahel encoded to label encode all the crop types from 0-21 Of which 0 implies Apple, 1 implies Banana and so on Then we use scalar transformation for standardization. And then we apply the ensemble of Xgboost and MLP that we have developed individually earlier. The accuracy that we reached by soft voting method is 99.3 %. In the realm of model efficiency, TP, TN, FP, and FN are commonly employed to assess the performance of a machine Seaming model, particularly in binary classification problems But the ensemble of XGBoost Classifier with MLP Classifier have lour values of FP ? FN like 1T or 0 So Even though the model is just 1% increased with accuracy than other ensemble models the efficiency and impact is justified respectively. After training we make the system for prediction by input the data in a array and fit the prediction system to the test data. As a future wrork we have set aside a set of powerful ML models like CADboost $_7$ Random Forest to be ensembled with DL models involving neural networks like ANN,LSTM etc. And the use of integrating the crop's genotype and weather variables can increase the prediction efficiency and accuracy.

References:

1. Keerthana* Mummaleti, K. J. \L Meghana, Siginamsetty Pravallika, and Modepalli Kavitlia. "An ensemble algorithm for crop yield prediction." In 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), pp. 963-970. IEEE, 2021.
2. G. Abirami, R. It H. Helan, k. Anandhan, G. V Reddy, S. N. Kumar, and V R. Karthick, "Crop Yield Prediction Using Ensemble Algorithm", 1JMDES, vol. 1, no. 6, pp. 54-59, Jun. 2022.
3. Privadharshini, A,* Chakraborty, S., Kumar, A., & Pooniwala, O. R. (2021, April). Intelligent crop recommendation system using machine learning. In 2021 5th international conference on computing methodologies and communication (ICCMG) (pp. 843-848). IEEE.
4. Pant, Jan me jay, K. P. Pant, Manoj Kumar Singh, Devesli Pratap Singh, and Hirnanshu Pant. "Analysis of agricultural crop yield prediction using statistical techniques of machine learning/1 Materials Today; Proceedings 46 (2021): 10922-10926.
5. Raja, s.P. & Sawicka, Barbara & Stamenkovic, Zoran & Ganesan, MariammaL (2022). Crop Prediction Based on Characteristics of the Agricultural Environment lising Various Feature Selection Techniques and Classifiers.

IEEE Access. 10. 1-1. 10.11Q9/ACCESS.2022.3154350.

6. BakthavatchaJam, KalaiseJvi, Balaguru Rarfhik, Vijayan Thiruvengadam, Sriram Muthal, Deepa Jose, Ketan Rotecha, and Vijavakumar Varadarajan. 2022. “IoT Framework for .Measurement and Precision Agriculture: Predicting the Crop Using Machine Learning Algorithms1’ Technologies 10, no. : 13. https://doi.Org/10.3390/techiioJogiesl0010013
7. IC Archana, Dr.K.G.Saranya«, “Crop Meld Prediction, Forecasting and Fertilizer Recommendation using Voting Based Ensemble Classifier,” SSRG International Journal of Computer Science and Engineering , vol. 7, no. 5, pp. 1-4, 2020.
8. Ujjainia, Shikha & Gautam, Pratima & Suraparaju, Veenadhari. (2021). A Crop Recommendation System to Improve Crop Productivity using Ensemble t echnique. International Journal of Innovative Technology and Exploring Engineering. 10. 102-105. 10.35940/ijitee.D8507.0210421.
9. M. Reert liana. k. J. M. Meghana, S. Pray a liika and M. Ravitha, “An Ensemble Algorithm for Crop Yield Prediction,” 2021 Third international Conference on Intelligent Communication Technologies and Virtual Mobile Networks (1CICV), I iruneJveJi, India, 2021, pp. 963-980, doi: 10.1109/JCICV50886.2021.9388489.
10. Barbosa, Alexandre ; Hovakimyan, INaira ; Martin, Nicolas F. /Risk-averse optimization of crop inputs using a deep ensemble of convolutional neural networks. In: Computers and Electronics in Agriculture. 2020 ; Vol. 178.

Artha
JOURNAL OF
SOCIAL SCIENCE

A Quarterly Inter-disciplinary peer-reviewed Social Science Journal
ISSN 0975-329X, doi:10.12724/ajss
from CHRIST (Deemed to be University), Bengaluru
published in January, April, July and October

Research articles, scientific papers, general articles, case studies, review articles, book reviews, comments, profiles, and books for review are invited. The contributions may kindly follow the "Guidelines for Contributors" provided in the journal.

All editorial correspondences are to be sent to:

Managing Editor, *Artha–Journal of Social Sciences*,
Centre for Publications, CHRIST (Deemed to be University),
Bengaluru 560029, India
Phone: (Office) 91 80 40129027
Email: artha@christuniversity.in or journals@christuniversity.in

---

Ushus
Journal of Business Management

A quarterly peer-reviewed Business Management Journal
ISBN 0975-3311, doi:10.12725/ujbm
from CHRIST (Deemed to be University), Bengaluru
published in January and July

Ushus-Journal of Business Management publishes original research articles based on primary or secondary data, book reviews, case studies, shared experiences, interviews with top business executives and the like in the areas of international business and management, marketing, human resources, finance and accounting, managerial economics, communications, management information systems, entrepreneurship and related disciplines. Contributors may kindly follow the "Guidelines for Contributions" provided in the Journal.

All editorial correspondences are to be sent to:

The Executive Editor, Ushus-Journal of Business Management,
Centre for Publications, CHRIST (Deemed to be University),
Bengaluru 560 029, India; Phone: (Office) 91 80 40129027
Email: ushus@christuniversity.in; journals@christuniversity.in

Mapana - Journal of Sciences
2023, Vol. 22, Special Issue 2, 171-201
ISSN 0975-3303|https://doi.org/10.12723/mjs.sp2.10

Introspection on the Research Avenues of Robotic Process Automation as a Service (RPAaaS)

Yashwanth Balan and [2] R Arokia Paul Rajan

Abstract

One of the newest business and technology developments is cloud computing, where several users approach the Cloud to complete various tasks. Cloud RPA is a technology that uses robotic process automation on Cloud-native using artificial intelligence. RPA-as-a-service: an automation software or bot that any user with an internet connection can use in the Cloud. It is an automaton self-service in cloud drag-and- drop actions and different GUI as a user-friendly software service. Cloud RPA ensures users automate any process via the Internet on the Cloud and can access it in their browser. RPA enables an intelligent agent to replicate typical manual decisions, such as rule based, well-structured ones involving vast amounts of data in a digital system, and eliminate operational errors. The highlights of the self-service aspect of cloud RPA, which allows any user with an

* Christ University, Bangalore 560029, India [1]yashwanth.balan@res.christuniversity.in [2]arokia.rajan@christuniversity.in

* Christ University, Bangalore 560029, India arokia.rajan@christuniversity.in