Farm2Kitchen AI: A Multi-Modular Machine Learning Approach for Self-Sustainable and Organic Farming

Jaydev Jangiti *, Ananthu Sajith, Sumedha Vadlamani*** and Yokesh Babu S†**

Abstract – In developing nations such as India, agriculture is the heart of the economy since it provides a significant category of the population with food and a means of subsistence. One of the major issues in farming is the need to increase crop yields to feed a growing global population while dealing with the challenges of climate change and limited resources. Moreover, Middlemen in agriculture create inefficiencies in the food supply chain by increasing prices for farmers and consumers, reducing transparency, and limiting direct connections between farmers and markets. Our Capstone initiative intends to optimize the farmer-to-consumer food supply chain by eliminating intermediaries and delivering fresh produce directly to consumers' kitchens. The project employs supervised machine learning models to recommend the most profitable crops for farmers to cultivate based on their individual expertise, environmental conditions, soil health, and other factors while concurrently balancing both supply and demand. By eliminating the need for middlemen, the project aims to lower costs for farmers and boost their self-sustainability. Farmers are provided with an Android application that stores all the data in a central cloud database. Based on the information provided, farmers are assisted to grow the highest yielding crop by analyzing their farming conditions and clustering individual farmers using unsupervised clustering algorithms, based on their expertise to ensure efficient crop distribution with respect to the market demand. In addition, the system promotes the growth of organic crops by assessing the health of the land's soil. The long-term objective of the project is to increase the profitability and efficiency of the agricultural business by enabling farmers to sell their goods directly and efficiently.

Keywords: Supervised Machine Learning, Big Data, Agri-Technology, Android Application, Cloud Database, Clustering Algorithms, Middlemen, Organic Farming.

1. Introduction

Agriculture has been the foundation of a strong economy and a sustainable world for generations. In developing nations such as India, agriculture is the heart of the economy since it provides a significant category of the population with food and a means of subsistence. However, Indian farmers face several obstacles that restrict their yields and make it difficult for them to achieve self-sufficiency. Numerous farmers are highly reliant on external inputs, such as seeds and fertilizers, which raises their expenses and makes them more susceptible to market changes. Inadequate infrastructure and limited market access make it difficult for farmers to sell their goods efficiently. The overuse of fertilizers is an additional major issue for Indian farmers. A study by the Center for the Study of Developing Societies (CSDS) revealed that a big proportion of Indian farmers are heavily reliant on fertilizers and pesticides, which can result in severe environmental repercussions and a lack of sustainability [1]. This is often the result of a lack of understanding regarding fertilizer consumption and the availability of alternative crop management techniques. According to the National Sample Survey Office (NSSO), more than half of India's farming households are in debt, and only about a third of them can earn enough to meet their expenses [2]. The existence of intermediaries, often known as "middlemen," is one of the biggest challenges in the agriculture sector. Typically, these intermediaries comprise wholesalers, processors, and distributors that purchase product from farmers at a low price and resell it to retailers or customers at a higher price, resulting in a significant markup. The average commission imposed by intermediaries in India for the sale of farm produce is 6.3%. In some cases, this fee can be as high as 30% [3]. This can make it difficult for farmers to receive a fair price for their produce, limit their ability to innovate, and conceal the true source of produce from customers. Moreover, they result in other negative outcomes for farmers, including decreased earnings, limited market access, a lack of control over price and distribution, a limited capacity for innovation, and a lack of transparency.

† Corresponding Author: School of Computer Science and Engineering, Vellore Institute of Technology, University, Vellore, India (yokeshbabu.s@vit.ac.in)

* School of Computer Science and Engineering, Vellore Institute of Technology, University, Vellore, India (kiee@kiee.or.kr)

** School of Computer Science and Engineering, Vellore Institute of Technology, University, Vellore, India (kiee@kiee.or.kr)

** School of Computer Science and Engineering, Vellore Institute of Technology, University, Vellore, India (kiee@kiee.or.kr)

Received ; Accepted

Agriculture has become a profession riddled with risks due to dwindling resources, shrinking land areas, rising input and labor expenses. To withstand the emerging challenges, the industry needs a major revision. Technological advancements can provide farmers with a wide range of resources that can help them improve their yields, reduce their costs, and become more self-sustaining. By utilizing these technologies, farmers can overcome their lack of expertise and become more competitive in the marketplace while also contributing to sustainable and organic farming practices. Today, along with other industries, agriculture is turning digital, and data-driven Agri-technology and precision farming are the emerging approaches to sustainable agriculture for maximizing production and reducing environmental impacts. According to a report by TechSci Research, the Indian precision farming market is projected to grow at a CAGR of over 19% during the forecast period of 2020-2025 [4]. Advanced agricultural operations generate data from a variety of sensors that enable a better understanding of the operational setting via the interactions between dynamic crop, soil, weather conditions, and machinery data, resulting in more precise and efficient strategic planning. Machine learning (ML) is a rapidly advancing field of study that is revolutionizing a broad range of industries. ML is the scientific discipline that allows computers to learn without being explicitly programmed, per one of its many definitions [5]. It is the application of algorithms and statistical models to enable computers to learn from data. This has led to the development of revolutionary technologies that have the potential to transform industries including healthcare, banking, and agriculture. It is revolutionizing the agricultural industry by enabling farmers to make more informed decisions about crop selection, farming practices, and market trends, using data-driven algorithms and models, resulting in increased efficiency, profitability, and sustainability for the agricultural industry. ML has the potential to solve the problem of agricultural industry intermediaries by recommending the most profitable crops for farmers to cultivate based on their individual expertise, environmental conditions, soil health, and other factors, while simultaneously balancing supply and demand and enabling farmers to sell their goods directly and efficiently. Moreover, it can be used to assess the health of the soil and identify optimal conditions for organic farming practices. Our study intends to optimize the food supply chain by eliminating intermediaries and delivering fresh produce directly to the homes of consumers. By removing middlemen, we want to reduce expenses for farmers and boost their self-sufficiency. Farmers will be able to use our user-friendly Android app to store data in a centrally managed cloud database, and the ML models will use this data to assist farmers in growing the highest-yielding crops, distributing them efficiently based on market demand, and promoting the growth of organic crops by assessing soil health. Our primary objective is to boost profitability and efficiency in the agriculture business by facilitating farmers' direct and efficient distribution of their goods.

2. Literature Review

The problem of lack of self-sustainability among Indian farmers and inorganic production has been a major issue for decades. Several studies have been conducted to address this problem and improve the agricultural industry in India. In this section of the literature review, the authors analyzed numerous techniques adopted by researchers, such as the use of machine learning, data analytics, the Internet of Things, and other technology advances to aid farmers in crop planning, distribution, production, and organic agriculture. With this assessment, the authors intend to uncover research gaps and provide a novel approach that addresses the constraints of earlier work while expanding on its accomplishments.

Adebiyi, Marion, and colleagues [6] focused on a machine learning–aided mobile system for farmland optimization. In the study they employed ML random forest and BigML algorithms to analyze and classify datasets containing crop features that generated subclasses based on random crop feature parameters. The study concluded that the approach aided decision making and that the system led to users' optimization of information when implemented on their farmlands. However, the study lacked a comparative analysis to understand the large-scale applicability of the work.

M Ryan Haley et al. [7], focused on a numerically efficient Machine Learning technique premised on the notion that a farm manager wants the allocation of crops that reduces the likelihood of achieving a harvest return below a specified goal rate of return. Using two pruning procedures, the study closes by discretizing the choice space of the real shortage probability. However, further research may be integrated by setting limits on the ML technique, and it may be shown to be effective in various types of shortfall and non-shortfall decision processes.

Seeboli Ghosh Kundu, Anupam Ghosh and colleagues [8] developed an ensemble method that combines ML models for improved agricultural output forecasting in India. The Ensemble model's validation output exhibits superior AUC, classification, F1 score, Precision and Recall values in comparison to other ML models, demonstrating optimal model performance and eliminating any overfitting bias. However, the ensemble ML technique has tremendous potential for forecasting yield estimates on different data sets and expanding the framework of research.

S.Veenadhari, Dr. Bharat Misra and Dr. CD Singh [9] sought to review research findings on the use of data mining methods in the sector of agriculture utilizing the ID3 algorithm and the K-Nearest Neighbor methodology. Comparing the spectral rank order resulted in a classification accuracy of 60%. Although, there is a great deal of potential for improvement in this developing scientific topic, its growth is anticipated.

P. Priya, U. Muthaiah and M. Balamurugan [10] aimed at forecasting the yield of the crop using the Random Forest method and existing data. The results demonstrate that the Random Forest method can accurately estimate agricultural production, making it appropriate for large-scale crop yield prediction. However, it is possible to discover the most effective algorithm based on their accuracy measures, which will aid in selecting an effective algorithm for crop production prediction.

Sonal Agarwal and Sandhya Tarar [11] proposed a python-based system that employs AI and calculations for formulating expectations, such as Multiple Linear Regression, to recognize the model among information in order to predict the most productive harvest under given conditions while minimizing costs. The conclusion of the study is that the accuracy of this research is improved by 97% compared to previous crop forecast work that employed other methodologies. Nevertheless, the approach's potential development is substantial, and it can be applied and interfaced with a flexible and multi skilled application.

Navod Neranjan Thilakarathne et al. [12], suggested a unique cloud-based ML-powered crop recommendation platform to aid farmers in determining which crops need to be harvested based on the best-performing ML algorithm used to construct the platform. Based on training performance measures, the Random Forest algorithm outperformed all other ML models and served as an underlying ML model in the crop recommendation platform. However, the effectiveness of the recommendation platform is contingent upon the availability and accuracy of data. Inaccurate or incomplete data could lead to poor recommendations and potentially impact harvest yield.

Manik Rakhra, Amitabh Bhargava, and colleagues [13] examined the demand and supply algorithms used to determine market equilibrium, assessed, and classified the works based on decision levels, supply chain topics, and optimization techniques, and suggested future research directions. The analysis concluded that decision trees performed the best on the dataset, with a mean absolute error of 138.33, and were selected as the best model. One major concern with the work is the potential for the system to be biased, particularly if the data used to train the system is not representative of the population of farmers who will use it which makes it unfit for largescale deployment.

D Ramesh et la. [14], aimed to develop a farmer-friendly interface that provides an analysis of rice productivity based on existing data. Comparing the outputs of two approaches enhanced and confirmed the accuracy of yield prediction. Nevertheless, there is a further scope to compare the crop production projection with all the available data.

Marion Olubunmi Adebiyi et la. [15] provided a machine learning–aided mobile system for farmland optimization, using various inputs such as location, crop type, soil type, soil pH, and spacing. Random forest algorithm and BigML were employed to analyze and classify datasets containing crop features that generated subclasses based on random crop feature parameters. The system followed pre-defined thresholds. However, ML model's accuracy was not discussed, furthermore the system was not deployed to understand the system's efficiency. Furthermore, the machine learning algorithm needs to be embedded in the system and used to perform predictions.

Other notable works include, [16] an online farmer portal to sell crops direct to consumers. [17] proposed an IoT Based Agriculture Monitoring and Smart Irrigation System Using Raspberry Pi. [18] utilizes smart irrigation model that forecasts the water demand for a crop using a machine learning algorithm. [19] developed an android app to Connect Farmers to the food Processing Industry. [20] proposes a beneficial guide to all stakeholders towards enhancing awareness of the potential advantages of using machine learning in agriculture. In conclusion, the lack of self-sufficiency among Indian farmers and their reliance on inorganic production has been a key issue that researchers have been trying to address for many years. The proposed initiative addresses a several challenges discovered during the research study. The work utilizes ML models to optimize crop planning, which would assist crop yields. It also promotes the development of organic crops by analyzing soil health and addressing the aforementioned concerns.

3. Proposed Methodology

Our approach intends to optimize the food supply chain by eliminating intermediaries and delivering fresh produce directly to consumers. We aim to minimize costs for farmers by assisting farmers in growing high-yielding crops and clustering them efficiently based on market demand, while also promoting the growth of organic crops considering environmental factors. We used machine learning models to balance supply and demand and to develop an Android application that will enable producers to access the system and store data in a centralized cloud database.

3.1 System Architecture

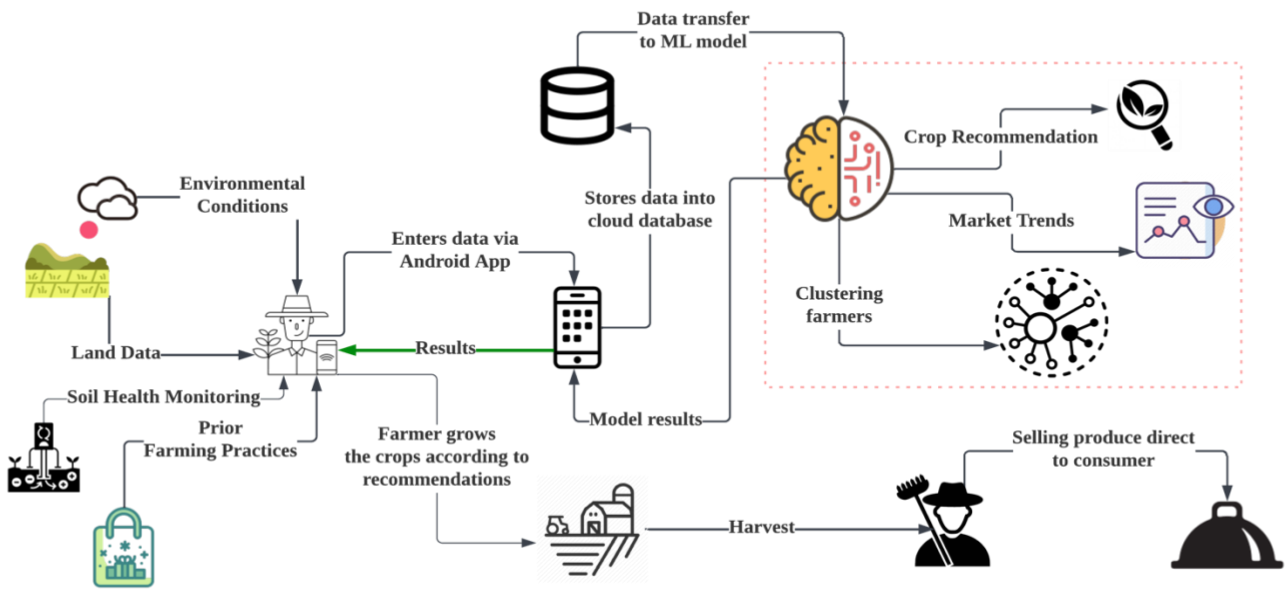


Fig. 1 Architecture Diagram

On basis of the data presented by the farmers, including in checking the health of the soil to ensure organic crop production, climatic conditions, and the crops they have produced in the past, an intelligent recommendation is devised to assist the farmer in achieving optimal yields. A simple user interface is created so that farmers can use the system, input data based on various factors, and receive recommendations on the best crops to cultivate. The farmer's data is stored in a centralized cloud database and transmitted to the machine learning model via the cloud service. In terms of data transfer, the machine learning model assists farmers in recommending the most profitable and high-yielding crops by placing together farmers with extensive field experience and balancing demand and supply trends. The results obtained by the machine learning model are transmitted to the mobile application, where the farmer seeks guidance on how to cultivate the crops in accordance with the recommendations. Based on the results, the farmer grows, and harvests crops and sells them directly to the consumer to avoid intermediaries, who are a significant factor of price hikes and the resulting steady decline in production.

3.2 System Modules

The system architecture is broken down into 3 distinct modules which are then integrated into a centralized android application.

3.2.1 Crop Planning Module

This module is responsible for analyzing the data entered by farmers and planning the most appropriate crops to be grown given the farming conditions. The authors have developed a novel deep learning-based ML model which is trained on the data collected to accurately provide the predictions. Based on the information provided by farmers, such as the health of the soil, the weather, the crops they have grown in the past, and the current market demand, an intelligent recommendation will be made to help the farmer get the best yields. This serves the purpose of lowering overhead costs for farmers and boosting their self-

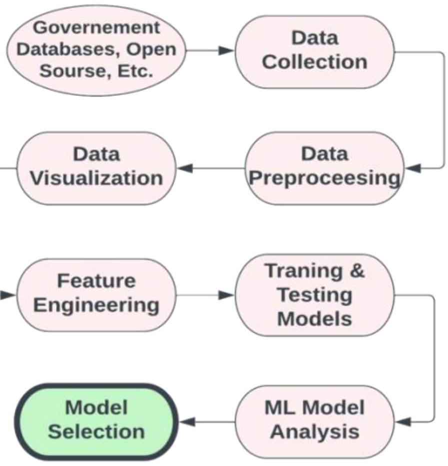


Fig. 2 Crop Planning Module Workflow

sustainability.

Fig.2 represents the workflow of the module and the step-by-step process followed.

3.2.1.1 Data Collection

The dataset utilized in this module was acquired from numerous sources including government and open-source resources. The International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), a non-profit organization that performs agricultural research for development in the semi-arid tropics of Asia and sub-Saharan Africa, was one of the primary sources for the dataset. ICRISAT provided valuable data on various crops grown in India, including their yield and nutrient requirements. In addition to ICRISAT, the dataset was also obtained from different open-source sites, including Kaggle. The "All Agricultural Related Datasets for India" dataset on Kaggle included a lot of information on various crops grown in India, in addition to climate and soil data.

Dataset Attribute	Attribute Description
N	Nitrogen Content of Soil (Integer)
P	Phosphorous Content of Soil (Integer)
K	Potassium Content of Soil (Integer)
Temperature	Temperature of Farm (Float)
Humidity	Humidity of Farm (Float)
pH	pH Soil of content (Float)
Rainfall	Rainfall of Farm (Float)
Crop	Crop Grown (String)
Farming Type	Organic or Inorganic Farming Conditions (String)
Demand	Market Demand of Crop (String)

Table 1. Dataset Attributes

Table 1. shows an abstract view of the attributes present in the dataset along with their data types and descriptions. This dataset proved particularly beneficial for detecting the association between agricultural yields and other environmental conditions. The dataset provides information on several agricultural characteristics like N, P, K, rainfall, temperature, humidity, pH, crop, demand, and farming type. N, P, and K indicate the nutrient content of the soil, whereas precipitation, temperature, and humidity represent the climatic conditions. pH represents the soil acidity, and crop, demand, and farming type represent the crop name, demand level (High, Mid, Low), and farming type (Organic or Inorganic), respectively. The collection covers a total of 22 different crop varieties, including rice, maize, chickpea, kidney beans, and more. The data provides valuable insights into the nutrient requirements and environmental conditions necessary for optimal crop growth, as well as the demand levels for various crops. This project's dataset was compiled from a variety of sources, including government and open-source repositories such as ICRISAT and Kaggle. The collection comprises a wide range of agricultural variables and gives significant information on diverse crops farmed in India.

3.2.1.2 Data Preprocessing

Data preprocessing is a key step in preparing data for machine learning models. In this project, the dataset was preprocessed using a variety of methods, including scaling, addressing missing values, and sampling.

1. To scale the numeric properties, the StandardScaler function from the scikit-learn library was utilised. Scaling is crucial because it guarantees that all variables are on the same scale, which can help prevent factors from having a disproportionate impact on the model's predictions. The StandardScaler function normalises the distribution of each variable so that its mean is zero and its standard deviation is one. Another key stage in data preprocessing is handling missing values.

2. In this study, missing values in the dataset were imputed using the SimpleImputer function from scikit-learn. The SimpleImputer function substitutes missing values with the column's mean, median, or most frequent value. Imputing missing values is significant since it ensures that the data is complete and can be used for model training.

3. Sampling is a crucial preprocessing procedure, especially when working with unbalanced datasets. In this project, the dataset was stratified sampled using the StratifiedShuffleSplit function from scikit-learn. Stratified sampling ensures that the distribution of the target variable (demand level) is preserved in both the training and testing datasets. This is essential because it helps prevent bias in the model's predictions, especially when working with unbalanced data sets. Scaling guarantees that all variables are on the same scale, whereas missing value handling ensures that the data is complete and can be used for training. Sampling, particularly stratified sampling, is critical when dealing with imbalanced datasets to reduce bias in the model's predictions.

3.2.1.3 Tensor Flow Based Deep Learning Module

TensorFlow is used to construct a neural network model capable of predicting the most suitable crop for a given region based on a variety of input variables including soil type, climate, and precipitation. Using the Sequential class from the tf.keras.models package, the model is constructed. The neural network architecture consists of three dense layers: the first layer contains 1000 neurons with a ReLU activation function, the second layer contains 500 neurons with a hyperbolic tangent (tanh) activation function, and the third layer contains as many neurons as crop categories, with a SoftMax activation function.

The ReLU activation function permits the model to introduce nonlinearity into the neural network, thereby enhancing its capacity to learn from complex data. The tanh activation function squeezes the output of each neuron within the range of -1 to 1, which might be advantageous when dealing with negative data values. The SoftMax activation function ensures that the outputs of the final layer are probabilities that sum to 1, making the results simpler to interpret. The model is trained with the training data and one-hot encoded labels using the fit function. The categorical cross entropy loss function measures the difference between the expected and actual output.

The Adam optimizer is used to modify the model's weights and biases to minimize the loss function. The model is trained for 100 epochs, during which the validation data is utilized to monitor the model's correctness. By the conclusion of the epochs, the model obtained an accuracy of 99.99%, showing that it can properly anticipate the most suited crop for a specific region.

To deploy the model on a mobile platform such as Firebase ML, the model must be transformed into a more efficient format, such as TensorFlow Lite. With the lite. TFLiteConverter from Keras model method, the Keras model is converted to a TensorFlow Lite model that can be utilized on mobile devices with minimal resources. TensorFlow plays a vital part in the crop recommendation project by enabling the development of a robust neural network model capable of making accurate predictions based on complicated data.

3.2.2 Android Mobile Application Module

To facilitate the utilization of the system by farmers across the country and minimize training time, an Android application with user-friendly features including multilingual capabilities and chat-to-speech was developed to deploy the system and integrate various ML modules.
The development of mobile applications for machine learning requires careful consideration of programming languages and SDKs. In this research, the Flutter SDK and Dart programming language were chosen for the development of the application due to several advantages. These include faster development time, cross-platform compatibility, and a native UI appearance that is more familiar to the user. The hot reload functionality and direct code generation also contribute to a quick and fluid program, which is essential for users with less powerful hardware. Furthermore, Flutter's interoperability with other Google services makes it easy to integrate databases and authentication systems.

To store and integrate the custom ML model developed in this study, Firebase was employed as a cloud-based database. The integration of Firebase in the application facilitated data accessibility and expedited processing times. The use of Firebase ML Kit enabled real-time forecasting of the crops that would generate the most profit for farmers. These features were integrated into an Android application that was distributed to farmers, offering a simple and user-friendly interface for entering data and receiving crop suggestions.

The developed mobile application for machine learning includes a chat-to-speech option, making it easier for farmers to listen to crop suggestions and input data without having to type. This feature is especially useful for farmers who may not be proficient in typing or reading. The application also includes form validations and authentication, ensuring that the data entered by farmers is accurate and secure.

To further enhance security, the entire application has been developed with security in mind. The Firebase cloud database used to store all data is secured with strong encryption protocols to prevent unauthorized access. Additionally, the application's user authentication and data transfer processes are protected with the latest security measures, including SSL/TLS encryption. The chat-to-speech feature, form validations, and strong security measures implemented in the application are aimed at providing farmers with a seamless and secure experience while using the application.

All these features are designed to assist farmers in making informed decisions and increasing crop yields. The integration of Firebase in the application ensures secure storage of all data in the cloud. Overall, the use of Flutter, Dart, and Firebase enabled the swift construction of a powerful and user-friendly mobile application for machine learning. By leveraging these features, farmers can make informed decisions regarding crop selection and management, leading to increased crop yields and profitability.

3.2.3 Farmers Distribution Module

The farmer distribution module employs the K-means clustering machine learning algorithm to effectively group farmers based on their unique farming conditions and prior experience. The clustering of farmers has been considered a viable strategy to optimize resource utilization, minimize competition, and foster collaboration among them.

The implementation of this approach has the potential to enhance the sustainability and efficiency of the agricultural ecosystem. This paper presents an implementation of the K-means clustering algorithm for the purpose of clustering farmers according to their respective farming conditions. In this module, the initial stage involves loading the dataset and subsequently partitioning it into distinct features. The features were subjected to scaling using the StandardScaler() function, which is a part of the sklearn library. The K-means algorithm is a widespread clustering technique that requires the determination of the optimal number of clusters.

To address this issue, the elbow method is commonly employed. This method involves plotting the within-cluster sum of squares against the number of clusters and selecting the number of clusters at the "elbow" of the curve, where the rate of decrease in the sum of squares slows down significantly. The present study involves the computation of the within-cluster sum of squares (WCSS) for varying numbers of clusters. The objective is to determine the optimal number of clusters, which is identified as the point at which the WCSS alteration begins to plateau.

Upon determining the optimal number of clusters, the K-means algorithm is subsequently trained with the identified number of clusters. The farmers are then grouped in accordance with the output of the algorithm. The clusters that were obtained were subsequently visualized through the utilization of a scatter plot. The color scheme employed in the present study involves the use of distinct hues to represent individual clusters. Additionally, the centroid of each cluster is visually indicated by a yellow point. The present study demonstrates the efficacy of the K-means algorithm in clustering farmers based on their farming conditions and experience.

This approach is expected to have a positive impact on the agricultural sector and contribute to the overall economic growth of the region. The implementation of this approach has the potential to enhance the productivity and profitability of farmers, while simultaneously advocating for sustainable and efficient agricultural practices.

3.2.4 Organic Farming Module

To attain the highest possible crop yields, organic agricultural practices must be promoted among farmers. In this research article, the authors acknowledged the significance of sustainable farming practices and advocated for organic farming techniques. To facilitate the implementation of these practices, the authors used a standard mathematical formula to evaluate the pH level of the soil and to analyze the fertilizer use history of farmers. This allowed for an objective evaluation of the producers' adherence to organic agricultural practices. In addition, to encourage the adoption of organic farming practices, the authors created an Android application that alerts farmers when chemical fertilizers are being used and encourages them to adopt organic farming practices instead. The authors seek to promote a healthier and more sustainable agricultural system by providing producers with real-time feedback and promoting sustainable farming practices.

3.2.5 Farmer Customer Auction Module

The authors have developed a novel approach referred to as the farmer-consumer auction module, which aims to establish a direct selling platform for farmers and consumers. This approach is designed to facilitate the sale of farm produce directly from farmers to consumers, thereby bypassing intermediaries. The module is expected to enhance transparency and fairness in the pricing of agricultural products, as well as promote the consumption of locally grown produce. The present system aims to eliminate intermediaries from the crop selling process, thereby empowering farmers to establish their own prices and vend their produce directly to end-users.

It operates by enabling farmers to enlist their available produce on the mobile application, along with the intended selling price. The app enables consumers to access the produce listings and engage in bidding for the crops they intend to procure. The establishment of a competitive marketplace for agricultural products is crucial in ensuring that farmers can receive equitable compensation for their crops.

This approach employed by the authors holds significant potential in transforming the agricultural industry towards the promotion of sustainable and equitable farming practises.

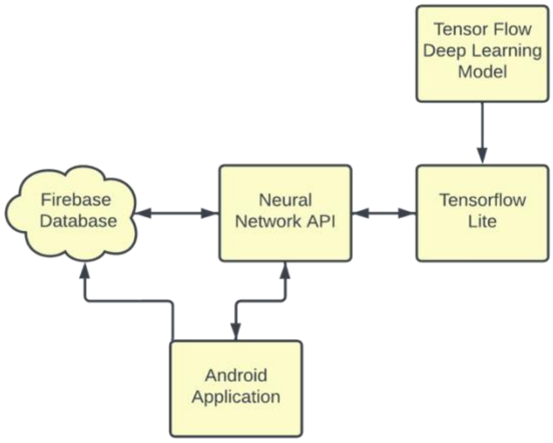
Working Algorithm:

1. The mobile application enables farmers to enlist their available crops along with the desired selling price.
2. The app enables consumers to access information on the crops that are currently available and place bids on the ones they intend to procure.
3. The bidding process in the agricultural sector involves the dissemination of bid notifications to the farmer, who is then empowered to either accept the highest bid or decline all bids.
4. Upon acceptance of a bid, the application serves as a communication platform for the farmer and the consumer to coordinate the delivery or pickup of the produce.
5. The direct payment system between the farmer and the consumer through the application obviates the requirement for intermediaries.
6. If the farmer is unable to vend the agricultural yield via the application, the farmer has the option to either re-list the crops or investigate alternative sales avenues.

This platform is designed to empower farmers and consumers alike, by facilitating direct transactions and promoting transparency in the agricultural market.

4. Implementation Results & Observations

To deploy the discussed proposed model, the authors have developed an android mobile application which can be easily accessed by farmers. The application has been integrated with several features to make it user friendly and has multilingual capabilities. The entire application's system modules have been integrated using the Firebase ML kit API. Fig. 3 below represents the implementation of the android



application.

Fig. 3 Firebase ML Kit

Firebase ML Kit is an effective machine learning solution for mobile applications. With its pre-built machine learning models and APIs, we can incorporate machine learning into our Flutter project fast and efficiently. But, for our unique application of crop recommendation, a new TensorFlow Lite model is required. To begin, we must first use the TensorFlow Lite Converter tool to convert our learned machine learning model into the TensorFlow Lite format. Once the model has been transformed, it may be stored in Firebase Cloud Storage or the device's local storage. Finally, we must configure the Firebase ML Kit API within our Flutter application. This requires adding the Firebase ML Kit dependency to our project, initialising the Firebase application, and configuring the Firebase ML Kit API to utilise our own TensorFlow Lite model.

As soon as the API is configured, we can begin generating predictions by utilising the dataset created. Then, we send this information to the Firebase ML Kit API, which utilises our proprietary TensorFlow Lite model to forecast the optimal crop to plant given the given soil and environmental circumstances. The forecast is returned to the app, where the recommended crop can be displayed to the user. Using Firebase ML Kit with a custom TensorFlow Lite model enables us to do crop suggestion rather easily in our Flutter app. By translating our learned machine learning model to the TensorFlow Lite format and integrating it with Firebase

ML Kit, we can propose crops to farmers based on their unique soil and environmental circumstances with speed and precision.

4.1 Crop Planning Module Results

Several well-known classification algorithms were evaluated as part of the authors' analysis of various machine learning (ML) models for the development of a produce recommendation system. Fig. 4 represents the accuracy results of the different ML models.

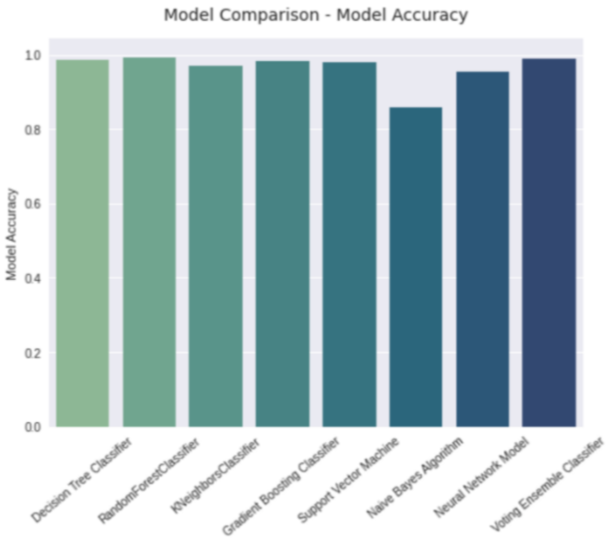


Fig. 4 ML Models Comparison

Among these were the Decision Tree Classifier, the Random Forest Classifier, the K Neighbors Classifier, the Gradient Boosting Classifier, the Support Vector Machine, the Naive Bayes Algorithm, the Neural Network Model, and the Voting Ensemble Classifier. The authors utilized the Python scikit-learn module to train and evaluate these models. Using a combination of Logistic Regression, Decision Tree Classifier, K Neighbors Classifier, Gaussian Naive Bayes, and MLP Neural Network, they then created an ensemble voting model. The Random Forest Classifier had the maximum accuracy of 98.3%, followed by the Voting Ensemble Classifier with 98.2%. The Decision Tree Classifier and Gradient Boosting Classifier had accuracies of 97.64% and 97.41%, respectively, while all other models had accuracies of less than 97%. Notably, the accuracy of the Naive Bayes Algorithm was the lowest at 85.91%. Both the Random Forest Classifier and the Voting Ensemble Classifier were viable options for the crop recommendation system, according to the authors. They noted that the Voting Ensemble Classifier was especially advantageous since it incorporated multiple models, thereby potentially enhancing overall precision. The authors cautioned, however, that there is a trade-off between precision and complexity, and while the Random Forest Classifier obtained the highest accuracy, it was also the most complex model and may not be appropriate for commercial deployment.

To further enhance accuracy, the authors investigated the possibility of fine-tuning Keras Neural Network models for their dataset and application. In addition, they considered employing transfer learning strategies and pre-trained neural network models for crop classification.

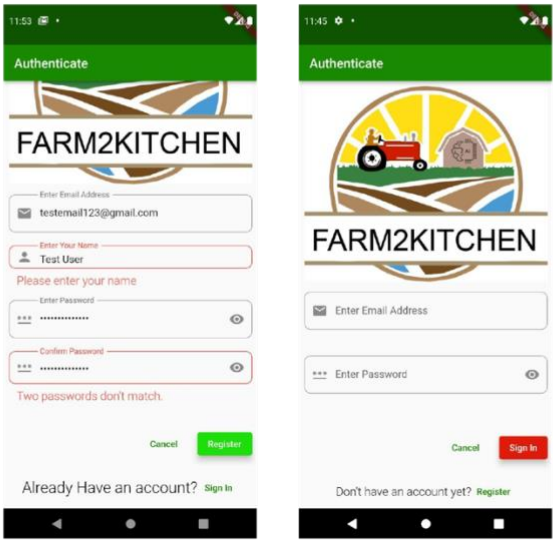
4.1.1 Tensor Flow Based Deep Learning Model Results

A custom neural network architecture comprising of three dense layers was developed by the authors to enhance the precision of the crop planning module. The neural network architecture consisted of three layers. The first layer comprised of 1000 neurons that utilized the Rectified Linear Unit (ReLU) activation function. The second layer consisted of 500 neurons that employed the hyperbolic tangent (tanh) activation function. Finally, the third layer contained several neurons equivalent to the crop categories and utilized the SoftMax activation function.

Upon conducting an evaluation of the custom neural network model, the authors observed that it exhibited superior performance in comparison to the standard models, as evidenced by its higher accuracy. The results indicate that the model attained a 98.6% accuracy score, rendering it the optimal choice for the crop planning module.

4.2 Android Application & Integrated Modules Results

An Android application was developed to enhance the accessibility of the system for farmers nationwide and reduce the duration of training. The application provides user-friendly features such as multilingual capabilities and chat-to-speech functionality, which enable the integration of multiple machine learning modules. The creation of mobile applications that utilise machine learning necessitates meticulous evaluation of programming languages and software development kits (SDKs). The author’s developed the android application using Dart programming language and Firebase cloud database. The figures below represent the final renders of the application. Fig. 5 shows the user registration and login screens. The authors have used various form validation techniques to ensure the data gathered is structured and reliable. Moreover, this ensures no duplicate users are created and integrates all the farmer’s details into the database based on their email address.

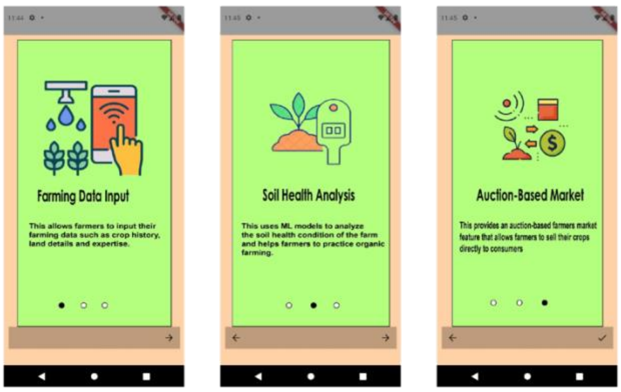


User registration with validations

Login page

Fig. 5 User Registration and Login

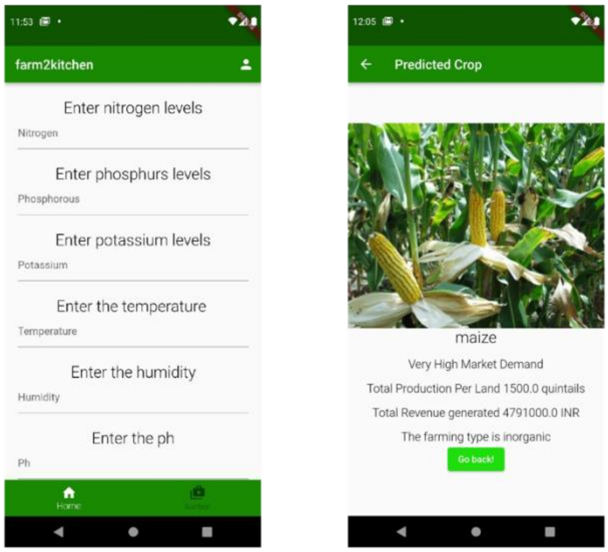
Once the user successfully creates an account by entering relevant details and logs in. They are greeted with few welcome screens with instructions on how to use the application and provides a simple guide. Fig. 6 represents the user onboarding screens.



User instructions with guide

Fig. 6 User Onboarding Screens

Once the user continues, they then have an option to select either to use the crop planning module to plan their next farming season or to sell their crops using the farmer auction module. Here the farmer is required to enter their information like soil health conditions, weather conditions, etc. which they are usually informed by the government officials who are responsible for providing the information. This information is then first stored in the Firebase database and then processed by the tensor flow based deep learning module before displaying the results to the farmers. Fig. 7 represents the data being entered by the farmers and the results shown by the application. Along with what crops to be grown the app also provides more information regarding the yield to be expected by growing the crops and the current market value for the crop at which they can potentially sell them at.

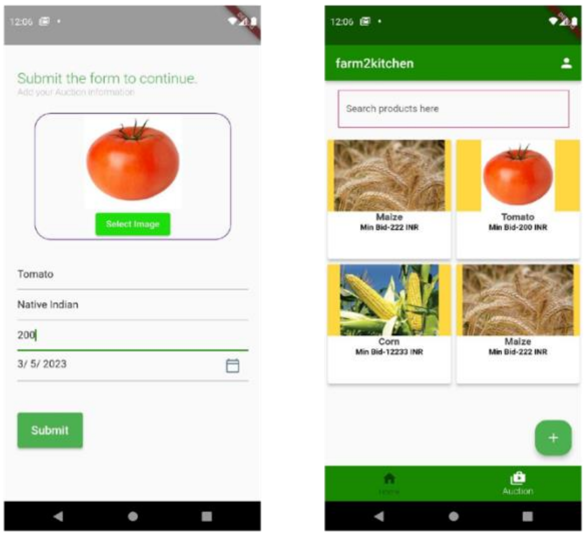


Entering details

Crop prediction

Fig. 7 Crop Planning & Predictions

Here the app also notifies the farmers if they are using inorganic practices. Fig. 8 represents the farmer-consumer auction module which allows the farmers to sell the crops at their desired rates.



Registering crop

Crop Auction Homepage

Fig. 8 Farmer-Consumer Auction Screen

The incorporation of the farmer-consumer auction module in the agricultural supply chain has been observed to have a positive impact on the consumers as well. This is attributed to the provision of an avenue for consumers to access fresh, high-quality produce at affordable prices. The module's ability to eliminate intermediaries from the crop selling process results in heightened transparency and a decreased probability of fraudulent activities and misrepresentation. It is an effective strategy for enhancing the autonomy of farmers in managing their businesses and fostering direct relationships with their customers. This approach has been shown to contribute to the development of a more sustainable and resilient food system.

4.3 Farmers Distribution Module

The implementation of the farmer distribution module has been found to be advantageous in fostering cooperation among farmers and optimizing resource utilization. The present study proposes a novel approach to enhance crop production and meet market demands by grouping farmers based on their farming conditions and prior experience. By working together, farmers can cultivate specific crops that are in high demand, thereby increasing their profitability.

This study employed the elbow method plot to ascertain the most suitable number of clusters for the crop advisory tool. The elbow method is a widely used technique in unsupervised machine learning algorithms, specifically in k-means clustering, to determine the optimal number of clusters.

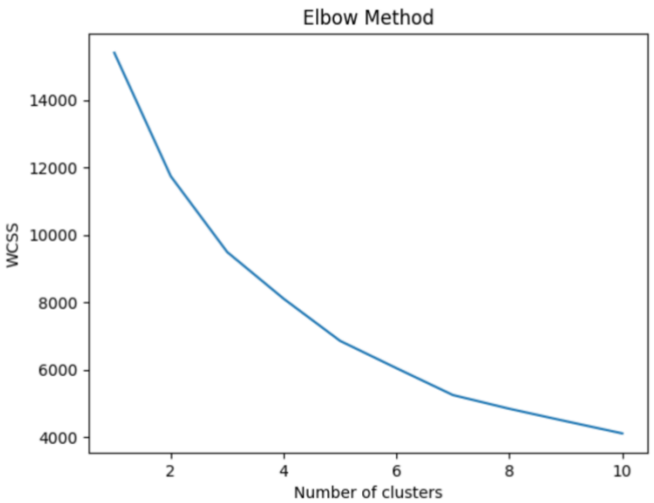


Fig. 9 Elbow Method Graph

The methodology of the elbow method involves the plotting of the within-cluster sum of squares (WSS) against the number of clusters. WSS is a mathematical metric that is computed as the summation of the squared Euclidean distances between every data point and its corresponding assigned cluster centroid. It has been observed that with an increase in the number of clusters, the within-cluster sum of squares (WSS) generally tends to decrease.

Upon reaching a certain threshold, the reduction in within-cluster sum of squares (WSS) exhibits a diminishing effect, leading to a distinct inflection point in the graphical representation resembling an elbow shape. The optimal number of clusters for the algorithm is commonly regarded as the number of clusters present at the elbow point. The present investigation employed the elbow method to determine the most suitable number of clusters for the crop recommendation system.

The optimal number of clusters for the crop recommendation system was determined to be three based on the analysis conducted by the authors. The data was utilized to train and assess the efficacy of the clustering algorithms implemented in the app. Fig. 9 represents the clusters formation by the K-means clustering algorithm.

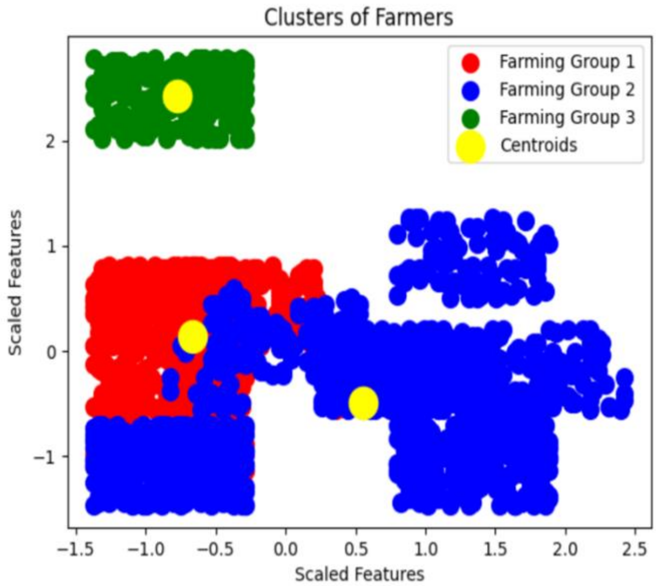


Fig. 10 Cluster Formation Based on Scaled Features

We can observe that the algorithm analyzed the dataset and formed 3 clusters based on the features correlations. By assigning each farmer a unique cluster, we can analyze the strengths and weaknesses of each cluster. This would allow the farmers to form unions and work together to increase the yield of the region. This also removes the dependence on growing a variety of crops by individual farmers and ensures there is less competition and more collaboration among farmers which ultimately improves the purchasing choices for the consumers.

5. Comparative Analysis

The literature survey analysis discussed earlier uncovered several prior attempts by researchers to create crop prediction systems, but none of them incorporated numerous elements targeted at empowering farmers and boosting their self-sufficiency in a single, multi-functional application. Table. 2 represents the side-by-side performance analysis of the accuracy measures for efficient crop planning module of prior work with the proposed deep learning model. By doing this, the authors show the originality and importance of their suggested approach in enhancing crop prediction technology.

Related Works	Methodology	Accuracy Score
Proposed Work	**Deep Learning Model**	**98.6%**
Reference [6].	Random Forest and BigML Model	97.5%
Reference [8].	Ensemble Model	98%
Reference [10].	Random Forest Model	95%
Reference [11].	Hybrid LSTM, RNN, SVM Model	97%
Reference [12].	K – Nearest Neighbors Model	96.3%

Table 2. Performance Analysis

From the above analysis it is evident that the proposed deep learning-based model for crop planning outperforms the recent works. Most of the prior work have had high success rates when combined with multiple other models. Work [8] encompassed a hybrid ensemble model however lacked the integration of market demand and price forecasts. Work [11] proposed a Multiple Linear Regression model with an accuracy of 97% compared to previous crop forecast work that employed other methodologies. Nevertheless, the approach's potential development is still substantial, and it can be applied and interfaced with a flexible and multi skilled application. Similarly, the other works lacked core features to enable farmers to be self-sufficient. These results indicate that the proposed model demonstrates superior accuracy compared to the referenced models, highlighting its potential effectiveness for the crop planning task. Moreover, the features provided in the proposed model such as multilingual capabilities, chat to speech, integrated cloud database and the unique farmer consumer auction system enables farmers to be completely self-sufficient and aims at easing their use of technology to improve their crop production.

6. Future Work

In future work, several areas will need to be explored to enhance the proposed initiative for optimizing the farmer-to-consumer food supply chain in developing nations. The integration of the farmer distribution module into the final application was not feasible due to limitations with the Firebase ML API, which restricted the testing and accurate clustering of individual farmers. Future work could explore alternative machine learning algorithms or frameworks that can handle the scale of data required for efficient farmer distribution, enabling a more comprehensive and effective implementation of the module. The integration of remote sensing technologies and Internet of Things (IoT) devices has the potential to facilitate more accurate monitoring of environmental conditions and soil health, thereby offering significant insights for crop selection and management. The validation of the effectiveness and scalability of the proposed solution in various regions and its wider adoption and impact can be achieved through conducting field trials and collaborating with government agencies. The initiative can further provide better assistance to farmers, encourage sustainable farming methods, and augment the overall efficiency and profitability of the agricultural industry by exploring these domains in future research.

7. Conclusion

In conclusion, the proposed work addresses the critical challenges faced by farmers to become more self-sufficient and reduce overhead costs By eliminating intermediaries and utilizing supervised machine learning models, the authors recommend the most profitable crops for farmers based on their expertise and farming conditions, while balancing supply and demand. The Android application developed by the authors, equipped with multilingual capabilities and chat-to-speech functionality, empowers farmers and promotes self-sustainability. The integration of a custom neural network architecture further enhances the precision of the crop planning module, resulting in superior accuracy. The proposed features, including the farmer-consumer auction system and cloud-based database, aim to improve transparency, lower costs, and boost the profitability and efficiency of the agricultural business. By enabling farmers to sell their goods directly and efficiently, the authors' initiative contributes to the long-term goal of ensuring food security and sustainable agriculture in developing nations. The integration of these features aims to alleviate the challenges faced by farmers, enabling them to leverage technology more effectively for agricultural advancements.

References

[1] Centre for the Study of Developing Societies (CSDS). "State of Indian Farmer: A Survey." 2018, http://www.csds.in/State_of_Indian_Farmers_Report.pdf. (Accessed on 24th January,2023)

[2] National Sample Survey Office (NSSO) - Key indicators of situation of agricultural households in India (January – December 2013) http://www.indiaenvironmentportal.org.in/node/403847/ (Accessed on 24th January,2023)

[3] National Bank for Agriculture and Rural Development. (2016). Report on Agricultural Marketing Infrastructure: A Case Study of Select Regulated Markets in India. https://www.nabard.org/auth/writereaddata/tender/2109171027Rpt%20on%20Agril%20Mkt%20Infra_24Nov2016.pdf

[4] https://www.techsciresearch.com/report/india-precision-agriculture-market/4029.html (Accessed on 22nd January,2023)

[5] Samuel, A.L. Some Studies in Machine Learning Using the Game of Checkers. IBM J. Res. Dev. 1959, 44,206–226.

[6] Adebiyi, Marion & Ogundokun, Roseline & Abokhai, Aneoghena. (2020). Machine Learning–Based Predictive Farmland Optimization and Crop Monitoring System. Scientifica. 2020. 1-12. 10.1155/2020/9428281.

[7] M. Ryan Haley,A Machine Learning approach for solving land allocation problems, Decision Analytics Journal, Volume 4,2022,100101, ISSN 2772-6622, https://doi.org/10.1016/j.dajour.2022.100101.

[8] Kundu, Seeboli & Ghosh, Anupam & Kundu, Avisek & Girish, G P. (2022). A Ml-Ai Enabled Ensemble Model for Predicting Agricultural Yield. Cogent Food & Agriculture. 8. 10.1080/23311932.2022.2085717.

[9] Suraparaju, Veenadhari. (2016). Data mining Techniques for Predicting Crop Productivity – A review article. 10.13140/RG.2.1.1755.0324.

[10] Priya, P., Muthaiah, U., & Balamurugan, M. (n.d.). Predicting Yield of The Crop Using Machine Learning Algorithm. International Journal of Engineering Sciences & Research Technology, 7(4),1-7.

[11] Agarwal, Sonal & Tarar, Sandhya. (2021). A Hybrid Approach for Crop Yield Prediction Using Machine Learning And Deep Learning Algorithms. Journal of Physics: Conference Series. 1714. 012012. 10.1088/1742-6596/1714/1/012012.

[12] Thilakarathne, N.N.; Bakar, M.S.A.; Abas, P.E.; Yassin, H. A Cloud Enabled Crop Recommendation Platform for Machine Learning Driven Precision Farming. Sensors 2022, 22, 6299. https://doi.org/10.3390/s22166299

[13] Rakhra, Manik & Bhargava, Amitabh & Bhargava, Deepshikha & Singh, Ramandeep & Bhanot, Astha & Rahmani, Abdul. (2022). Implementing Machine Learning for Supply-Demand Shifts and Price Impacts in Farmer Market for Tool and Equipment Sharing. Journal of Food Quality. 2022. 1-19. 10.1155/2022/4496449.

[14] Ramesh, D. (2015). Analysis Of Crop Yield Prediction Using Data Mining Techniques. International Journal of Research in Engineering and Technology. 04. 470-473. 10.15623/ijret.2015.0401071.

[15] M. A. Babar, S. M. Hasan, S. Hussain, A. Wahab, M. Imran, and M. A. Raza, "Smart Tillage: Improving Rental and Sharing of Agricultural Equipment Using Decision Trees and Google APIs," Scientifica, vol. 2020, Article ID 9428281, 12 pages, 2020. https://doi.org/10.1155/2020/9428281.

[16] Vighneshwar Boga, Varun Gaikwad, Shreeya Aswekar Prof. Saroja T.V. "An Android App for Online Auction of Agriculture Products", International Journal of Emerging Technologies and Innovative Research (www.jetir.org), ISSN:2349-5162, Vol.6, Issue 4, page no.195-201, April-2019

[17] Vineela, T., et al. "IoT based agriculture monitoring and smart irrigation system using raspberry Pi." Int. Res. J. Eng. Technol 5.1 (2018): 1417-1420

[18] K. S. Pratyush Reddy, Y. M. Roopa, K. Rajeev L.N. and N. S. Nandan, "IoT based Smart Agriculture using Machine Learning," 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 2020, pp. 130-134, doi: 10.1109/ICIRCA48905.2020.9183373.

[19] P. Shriram and S. Mhamane, "Android App to Connect Farmers to Retailers and Food Processing Industry," 2018 3rd International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 2018, pp. 284-287, doi: 10.1109/ICICT43934.2018.9034434.

[20] Seeboli Ghosh Kundu, Anupam Ghosh, Avisek Kundu & Girish G P (2022) A ML-AI ENABLED ENSEMBLE MODEL FOR PREDICTING AGRICULTURAL YIELD, Cogent Food & Agriculture, 8:1, DOI: 10.1080/23311932.2022.2085717.