WWW.IJRPR.COM

**International Journal of Research Publication and Reviews**

Journal homepage: www.ijrpr.com ISSN 2582-7421

# Cultimax – Web App for All Farming Needs

***K M Gaviprasad[1], Sailaja Thota[2], Keerthana Priya R[3], Shreya V[4], Suraj Kulkarni[5]***

***Dept. of Computer Science Engineering, Reva University, Yelahanka, Bengaluru -560064.***

**A B S T R A C T**

Agriculture is the main source of income in most developing countries. Modern agriculture is a constant change in agriculture and agricultural technology. It is important for farmers to respond to changes in the world, traders, consumers, etc. It becomes difficult to meet their expectations. Some of the problems faced by farmers are: (i) Avoidance of climate emissions due to soil erosion and industrialization (ii) Soil scarcity due to lack of essential nutrients such as potassium, nitrogen and phosphorus, slowing down the growth of crops. . (iii) Farmers make the mistake of planting the same crop the previous year without trying different varieties. They add random fertilizer without understanding its quality and quantity. The paper aims to find the best model for crop forecasting that can help farmers decide which crops to grow based on weather conditions and soil nutrients. With user interaction, the website provides a platform for farmers to guarantee higher revenue. The website serves as a distinctive and safe method for carrying out Agri-marketing. With some basic online navigational skills, farmers will be able to market their goods across the nation through e-farming. This solution enables customers to rapidly buy desired things by using online payment while examining the many products that are readily available.

Keywords: ***Light Gradient Boosting Machine, crop prediction, fertilizer recommendation, Agricultural sustainability, ML***

## 1.INTRODUCTION

Machine learning is an important decision to predict agricultural yields and decide which crops to plant and what to do during the growing season. Various types of machine learning are used to assist in crop forecasting research. The traditional agricultural market structure often disadvantages farmers. Dependence on middlemen, unpredictable demand, and limited access to consumers can result in unfair prices and difficulties in earning a fair living. E-commerce platforms designed specifically for farmers offer a groundbreaking solution.

Machine learning is used in many ways, from measuring consumer behavior in supermarkets to predicting consumer phone usage. Agriculture has been using machine learning for many years. Crop forecasting is one of the most difficult problems facing agriculture and many models have been developed and used to date. Since crop production is affected by many factors such as atmospheric conditions, fertilizer type, soil and seeds, this competition requires the use of various data sets. This means that estimating agricultural productivity is not a simple process; but it must have a complex process. Crop forecasting techniques may now be appropriate for using. Machine learning, with algorithms like Light Gradient Boosting Machine (LGBM), offers a powerful tool for crop prediction. By integrating LGBM predictions with e-commerce platforms, farmers can list their anticipated yield, adjust pricing strategies, and optimize resource allocation based on market demand, ultimately leading to a more robust and profitable agricultural ecosystem.

The synergy between LGBM predictions and e-commerce allows farmers to optimize production, set competitive prices, and minimize risk, paving the way for a more sustainable and profitable agricultural future. This combined approach has the potential to revolutionize the agricultural sector, leading to a more sustainable, efficient, and profitable future for farmers. It can also play a crucial role in ensuring a stable food supply for a growing global population, paving the way for a more food-secure future for all. This paper compares and defines how LGBM classifier is better for crop prediction based on nitrogen, phosphorous, pottassium, temparature etc.

It is reasonable to say that the process of shopping on the internet is becoming more common. Purchasing and selling of goods and services, as well as the transmission of funds or data, over an electronic network, most notably the internet. E-commerce and E-business are frequently used interchangeably. The primary goal of this project is to assist farmers in increasing their profitability through direct farmer-to-end user communication. Our project is concerned with the farmers' benefit of having their products sold at the best price online and to buy products at affordable prices.

Farmers, customers, and administrators are the primary users of this website. Farmers will have their own interface where they can perform tasks. By enabling various business models including multi-suppliers, e-sales, and different kinds of auctions, agricultural e-commerce creates favourable trading opportunities. Today's e-commerce still relies heavily on consumers' manual labour because it lacks completely automated business processes. So, the goal of our project is to fill in any gaps in the E-commerce business process.

## 2. Literature Survey

The author of [1] introduced a approach utilizing feed-forward back propagation Artificial Neural Network (ANN) to predict and forecast various crop yields in rural regions. By including a wide range of soil parameters including pH, nitrogen, potassium, alongside atmospheric factors such as rainfall and humidity, the model seeks to offer robust insights into crop yield dynamics.

The paper [2] aimed to determine the best and futuristic model for crop prediction, assisting farmers in selecting suitable crops based on climatic conditions and soil nutrients. The paper compared algorithms like KNN, Decision Tree, and RF Classifier using Gini and Entropy criteria. Results showed RF achieved the highest accuracy among them.

Naive Bayes, a supervised learning algorithm obtaining an accuracy of 97% was further improved by using boosting algorithm, which makes use of weak rule by an iterative process to bring higher accuracy [5] in that. To anticipate the yield, this employs advanced regression techniques such as ENet, Kernel Ridge, and Lasso algorithms [4].

Datasets from Karnataka and Maharashtra, different machine learning algorithms such as KNN, SVM, MLR, Random Forest, and ANN were deployed and assessed for yield to accuracy [9]. The accuracy of the above algorithms is compared.

## 3. Methodology

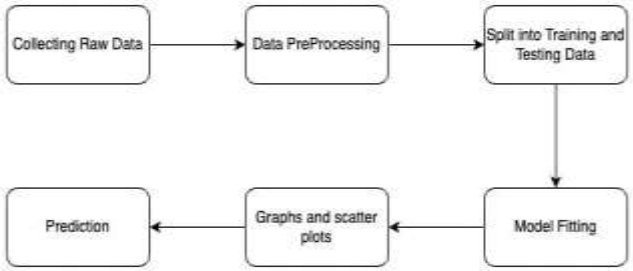


Fig-1: Methodology

**Data Collection** The practice of collecting and recording data is a way of tracking past events so that repeating patterns can be implemented. "Crop Re commendation" data is collected from the Kaggle website. These data include 22 different crops (N) (ii) Phosphorus content ratio in degrees Celsius (P) (v) Moisture percentage in mm.

**Data Preprocessing** The process of organizing raw data into the form of a learning algorithm to understand or predict results is called data preprocessi ng. The data processing process in this project is to find missing values. It is difficult to obtain all data pointing to every data in the data set. Blank cells, zero values, or certain characters (such as question marks) may indicate missing data. There are no missing values in the data used in this project.

**Train and test Split** is the process of exporting the dataset to training data and using the Train_test_split() function to split the dataset into training data and test data. Process of Scikit learning module. The 2200 files in the dataset were modified with 80% of the dataset as training dataset (1760) and 20 % of the dataset as test data (440) documents

**Fitting the Model Change the Model** Parameters that increase accuracy are called fitting. To create machine learning models, algorithms need to proc ess data for different purposes. The pattern is determined by comparing the output pattern with the face value of the target variable. Model fit is the abil ity of a machine learning model to generalize to data compared to training data. A good model fit is a model that accurately approximates the output giv en unknown input.

The **"Scatter Plot"** provides a visual representation of the model's performance. We can observe a good relationship between the prediction and the act ual production results, indicating that the LGBM model has a good result compared to the baseline model in the literature.

When a prediction model predicts the probability of a particular event, the "prediction" refers to the result of an algorithm that learns about previous dat a and applies it to new information. Use the test data feature set to predict the model using **the predict()** method. It outputs an array of predicted values.

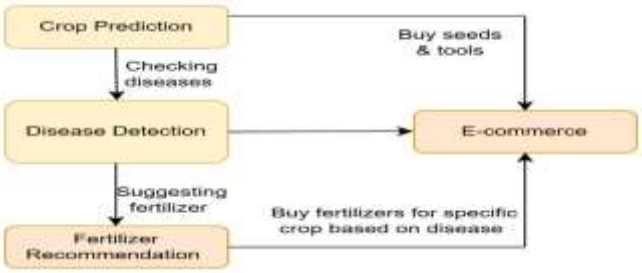


Fig – 2: Architectural Flow

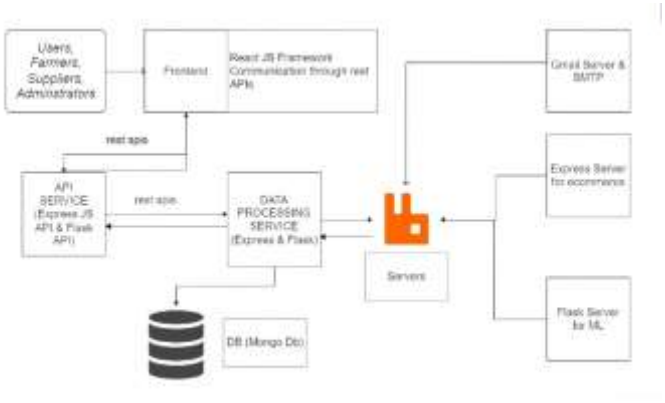


Fig – 3: Architectural Design

### *3.5.1 Mathematical Formulation*

Gradient boosting algorithms aim to minimize a loss function that measures the prediction errors of the model. Here's a simplified equation representing the LGBM objective function with L2 loss:

Obj = sum( $L(y_i, f(x_i))$ ) + λ * sum( $||w\_t||^2$ )

where:

- $L(y_i, f(x_i))$ represents the loss function for a data point i. It calculates the difference between the actual value ($y_i$) and the predicted value $f(x_i)$ for that point.
- λ is the regularization parameter that controls model complexity. A higher λ value reduces model complexity by penalizing large weight vectors (w_t).
- w_t denotes the weight vector for the t-th iteration of the model, indicating the influence of each tree in the ensemble.

### *3.5.2 Advantages of LGBM*

- Speed and Efficiency: LGBM's optimization techniques enable it to handle large datasets significantly faster than traditional gradient boosting methods.
- Improved Accuracy: LGBM often achieves competitive or superior accuracy compared to other algorithms, particularly for complex problems.
- Large Dataset Suitability: LGBM is well-suited for large datasets due to its memory efficiency and scalability.
- Categorical Feature Handling: LGBM has built-in functionalities for handling categorical features, eliminating the need for complex preprocessing steps.
- Customization: LGBM supports various loss functions, allowing you to optimize the model for specific tasks.
- Distributed Training: LGBM can be scaled across multiple machines to tackle even larger datasets

### *3.5.3 Algorithm: LGBM Training*

Input:

- Training data: X (features) and y (target labels)
- Learning rate: eta
- Number of trees: num_trees
- Other hyperparameters (e.g., max_depth, min_split_gain)

Output:

LGBM model: A model ensemble consisting of decision trees. Steps:

1. Initialize:
    - Initialize an initial prediction model (e.g., predict the mean of target labels).
    - Initialize an empty model ensemble model_ensemble = [].
2. for i in range(num_trees):
    - Calculate gradients: Compute the gradients of the loss function for each data point based on the current model in model_ensemble.
    - Build a new tree:
        - Use the gradients as targets and the features (X) as predictors to build a new decision tree.
        - Apply chosen hyperparameters (e.g., max_depth, min_split_gain) during tree construction.
    - Update predictions:

    Add the predictions from the newly built tree to the current model predictions.
3. Return the final model:

The final LGBM model is the ensemble of all decision trees built in the loop (model_ensemble).

## 4. Proposed Work

Our proposed web application “Cultimax”, integrates e-commerce and ML algorithms to recommend crops and fertilizers, factoring in both soil nutrients and weather conditions. This holistic approach aims to empower farmers with comprehensive insights for improved decision-making in crop selection and cultivation strategies. Cultimax aims to streamline agricultural trade by providing farmers with reliable solutions and offering customers a convenient platform for purchasing agricultural goods. The crop recommendation module, implemented by LGBM, utilizes input from farmers regarding soil and weather parameters to analyze and suggest the most suitable crops for cultivation

- Cultimax aims to increase farmers' income and success by providing personalized advice, a comprehensive platform for farming needs, and early detection mechanisms for crop diseases.
- Additionally, it facilitates direct farmer-to-customer connections, keeps farmers informed about government schemes, and offers a diverse range of agricultural products on its e-commerce platform.
- Ultimately, Cultimax strives to make a positive impact on the lives of farmers, transcending mere technological innovation.

Cultimax has following modules:

- E-commerce module
- Crop recommendation module
- Fertilizer recommendation module
- Disease detection module

LGBM Classifier: A Powerful Gradient Boosting Framework

In both the crop recommendation and fertilizer recommendation modules, LGBM algorithm has been employed.

LGBM is an open-source ML framework renowned for its speed, efficiency, and accuracy in handling large datasets. It excels at various tasks, including classification (predicting categories) and regression (predicting continuous values). LGBM differs from other tree-based learning algorithms by employing a vertical tree growth strategy instead of horizontal growth. In particular, LGBM expands trees leaf-wise, prioritizing leaves with the maximum delta loss for expansion. This approach allows LGBM to achieve greater loss reduction compared to level-wise algorithms when expanding the same leaf.

The core principle of LGBM is leveraging the power of gradient boosting, a technique that sequentially builds an ensemble of weak learners (often decision trees). Each new learner rectifies the errors of the preceding ones, progressively enhancing the model's overall performance. LGBM incorporates several innovative techniques to optimize training speed and memory usage, particularly for massive datasets:

- Gradient-based One-Side Sampling (GOSS): Focuses training on data points with higher gradients (errors from prior models), reducing computation by strategically selecting informative data.
- Exclusive Feature Bundling (EFB): Groups similar features before finding the optimal split point during tree construction. This improves processing speed and memory efficiency, especially for datasets with sparse features (many zeros).
- Histogram-based Splitting: Employs histograms to efficiently determine the best split points for continuous features. Histograms provide a summarized view of feature distribution, enabling faster identification of optimal split locations within the trees

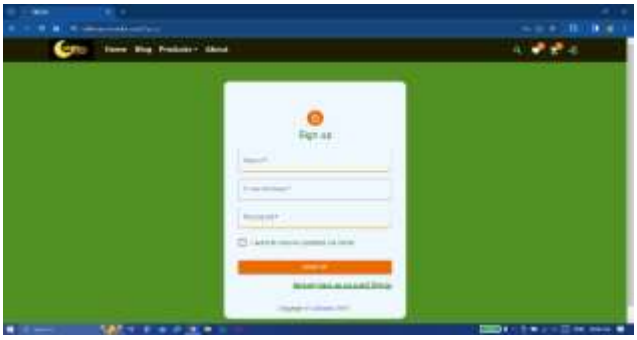

Fig 4 : Registration Page

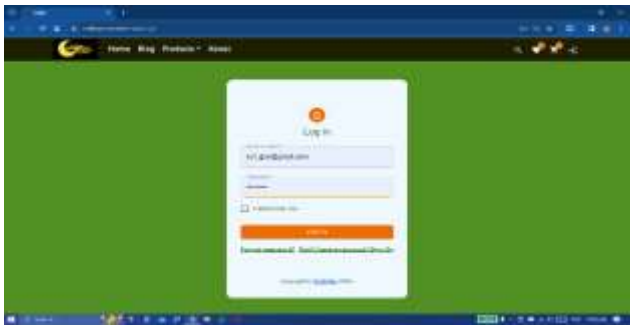

Fig 5: Login Page


Fig 6 : Home Page of the Cultimax


Fig 7: Crop Recommendation Module


Fig 8: Disease Detection Module


Fig 9: Fertilizer Recommendation Module

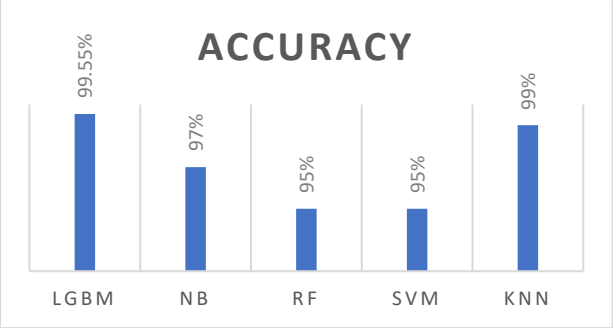


Fig 10 : Accuracy of various ML algorithms in Crop Recommendation

\ Fig 11: Crop Prediction Results

## 5. Conclusion

Cultimax is an innovative platform designed to help farmers overcome challenges they face, such as crop diseases and fertilizer suggestions based on diseases. Product Recommendations: Instead of spending time searching for good products, Cultimax lists highly rated items that other farmers have found effective. Crop prediction: LightGBM (Light Gradient Boosting Machine) is an ensemble model commonly used for crop prediction. It aids in smart crop recommendations by analyzing attributes like soil nutrients, temperature, and humidity. Additionally, LightGBM assists in crop type mapping from satellite imagery and predicts crop yields, contributing to efficient and sustainable agriculture practices. Disease Detection: Convolutional Neural Networks (CNNs) play a crucial role in revolutionizing agriculture by enabling early disease detection and informed decision-making for farmers. These networks are particularly effective for diagnosing plant diseases from leaf images. By training on diverse datasets containing various classes of plant leaves, CNNs can accurately identify diseases, helping farmers take timely actions to protect their crops. Fertilizer Recommendation: LightGBM is crucial for intelligent fertilizer recommendations in agriculture. It tailors' advice based on soil conditions, empowering farmers to enhance yields. Additionally, it contributes to sustainable farming practices by optimizing resource utilization and minimizing environmental impact. The combined power of LGBM and e-commerce platforms offers a glimpse into a future of intelligent and data-driven agriculture. However, for this vision to become reality, collaboration between researchers, agricultural stakeholders, and technology developers is crucial. By working together, we can create a more secure and sustainable food system for generations to come.

**References**

[1] Dahikar S and Rode S V 2014 Agricultural crop yield prediction using artificial neural network approach International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering vol 2 Issue 1 pp 683-6

[2] Suresh A, Ganesh P and Ramalatha M 2018 Prediction of major crop yields of Tamilnadu using K-means and Modified KNN 2018 3rd International Conference on Communication and Electronics Systems (ICCES) pp 88 93 doi: 10.1109/CESYS.2018.8723956.

[3] Medar R, Rajpurohit V S and Shweta S 2019 Crop yield prediction using machine learning techniques IEEE 5th International Conference for Convergence in Technology (I2CT) pp 1-5 doi: 10.1109/I2CT45611.2019.9033611.

[4] Nishant P S, Venkat P S, Avinash B L and Jabber B 2020 Crop yield prediction based on Indian agriculture using machine learning 2020 International Conference for Emerging Technology (INCET) pp 1-4 doi: 10.1109/INCET49848.2020.9154036.

[5] Kalimuthu M, Vaishnavi P and Kishore M 2020 Crop prediction using machine learning 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT) pp 926-32 doi: 10.1109/ICSSIT48917.2020.9214190.

[6] Geetha V, Punitha A, Abarna M, Akshaya M, Illakiya S and Janani A P 2020 An effective crop prediction using random forest algorithm 2020 International Conference on System, Computation, Automation and Networking (ICSCAN) pp 1-5 doi: 10.1109/ICSCAN49426.2020.9262311.

[7] Pande S M, Ramesh P K, Anmol A, Aishwaraya B R, Rohilla K and Shaurya K 2021 Crop recommender system using machine learning approach 2021 5th International Conference on Computing Methodologies and Communication (ICCMC) pp 1066-71 doi: 10.1109/ICCMC51019.2021.9418351.

[8] Bharath S, Yeshwanth S, Yashas B L and Vidyaranya R Javalagi 2020 Comparative Analysis of Machine Learning Algorithms in The Study of Crop and Crop yield Prediction International Journal of Engineering Research & Technology (IJERT) NCETESFT – 2020 vol 8 Issue 14.

[9] Mahendra N, Vishwakarma D, Nischitha K, Ashwini and Manjuraju M. R 2020 Crop prediction using machine learning approaches, International Journal of Engineering Research & Technology (IJERT) vol 9 Issue 8 (August 2020