ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11,S Iss 3, Dec 2022

FERTILIZER RECOMMENDATION SYSTEM USING CATBOOST ALGORITHM

Sri Silpa Padmanabhuni[1],J Lakshmi Narayana[2],Hema Lakshmi Bhavani Konjeti[3],Vudathu Venkata Krishna sai Poojitha[4], Boggarapu Rupa[5], Chirala Jaya[6]

[1]Associate Professor, Department of Computer Science and Engineering, PSCMR College Of Engineering And Technology, Vijayawada, A.P.
[2]Professor & Principal, Department of ECE, PSCMR College Of Engineering and Technology, Vijayawada, A.P.
[3,4,5,6] Student, Department of Computer Science and Engineering, PSCMR College Of Engineering And Technology, Vijayawada, A.P.
Silpa.padmanabhuni@gmail.com

DOI: 10.48047/IJFANS/S1/115

ABSTRACT:

Crop cultivation is crucial in the farming sector. Farmers strive hard in upbringing the yield by making their land fertile to meet the needs of the country. Due to years of farming, the soil has lost its fertility and results in low productivity. In order to overcome this factor chemical fertilizers came into lime light. Fertilizer is a natural or artificial substance that is applied on soil containing the chemical components that enhance plant growth and productivity. Monitoring the soil fertility and growth of the crop mainly depends on the pH of the soil, temperature and the moisture content in the soil. The proposed system utilizes the data set called Fertilizer Recommendation dataset in which all the above-mentioned factors are available. The class label considered in this dataset consists categorical values. As our class label contains categorical values Categorical Boosting algorithm is used in the proposed recommendation system. The Categorical Boosting Algorithm enhances the activity of recommendation. Our main aim is to help the farmers with a better and timely guidance by recommending the reasonable fertilizer for the land and yield considering the respective factors with the help of recommendation system. In this project the data of different parameters regarding the proper growth and their relative fertilizers pattern were collected, organized and summarized.

Keywords: Machine learning, CatBoost Algorithm, Fertilizer Recommendation,

1. INTRODUCTION:

This paper presents the fertilizer recommendation system using machine learning approach. Machine Learning is a subfield of artificial intelligence which uses the statistical models to provide predictions. It is frequently referred to as a type of predictive modelling or predictive analytics, and traditionally, It has been defined as a computer's ability to learn without being explicitly programmed. Programmed. Technically speaking, Machine learning uses algorithms to analyse empirical or historical data and produce results based on that analysis.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11,S Iss 3, Dec 2022

In some methods, the algorithms first operate on "training data," after which they learn, project, and seek ways to improve their performance over time. There are four types of machine learning, which are represented in figure 1.

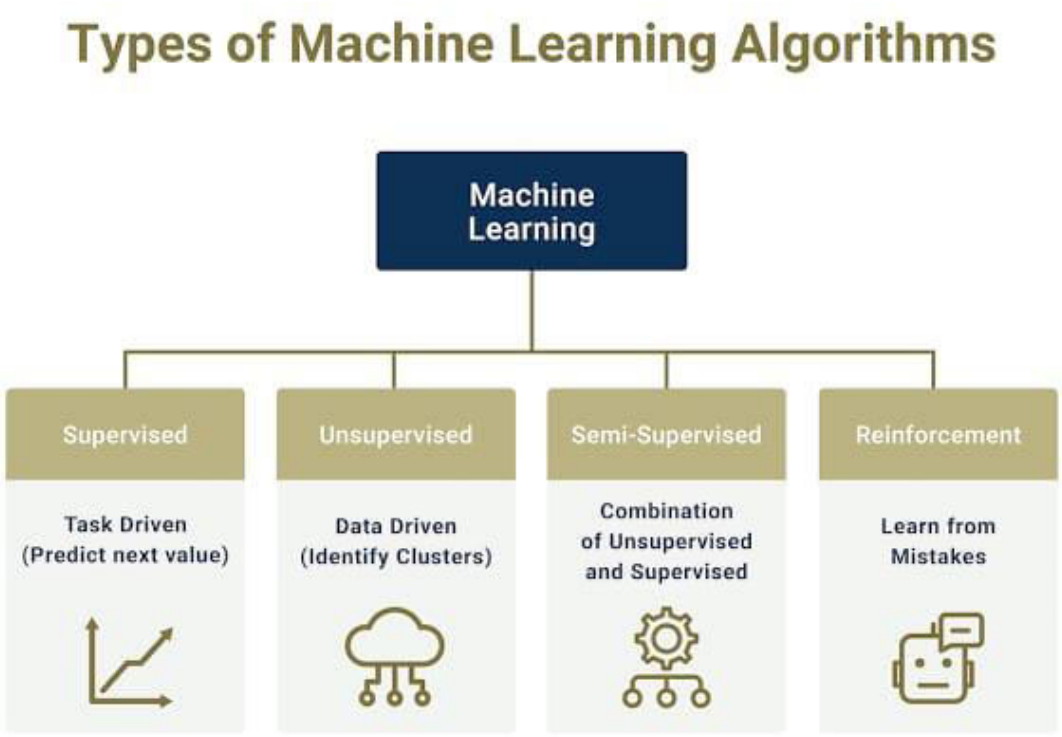

Figure 1: Types of Machine learning

Supervised Machine learning:

In this method we train our machines using the" labelled dataset" , and based on that training our machine learning model predicts the outcome. The main objective of this technique is to map the input variable with output variable.

Unsupervised Machine learning:

In this method we train our machines using the "Unlabelled dataset" , and based on that training our machine learning model predicts the outcome without any direction. The main objective of this technique is to categorize the dataset that is unsorted accordingly to the similarity.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11,S Iss 3, Dec 2022

Semi -Supervised Machine learning: This technique lies between the supervised and unsupervised learning. This concept is introduced to overcome the disadvantages of supervised and unsupervised learning.

Reinforcement learning:

Reinforcement learning is a feedback-based machine learning technique in which an agent learns to behave in an environment by executing actions and seeing the outcomes of actions. The agent receives positive feedback for each successful action, and negative feedback or a penalty for each unsuccessful action. In contrast to supervised learning, reinforcement learning involves the agent learning naturally from feedbacks without the need of labelled data. Since there is no labelled data, the agent can only learn by experience.

2. LITERATURE SURVEY:

[1] Zahra Hatam et al. proposed amodel based on gas exchange and responses to chlorophyll fluorescence; potassium fertilizer and zinc are recommended for cotton seedlings under salt stress. Using yield of photosystem II photochemistry describes Cotton as often grown in warm climates where salt and nutrient stress are pervasive and frequently affect plant growth simultaneously. The recommendation of fertilizer in salty environments is challenging since nutrient administration might affect a plant's resistance to salt, making crop output predictions more difficult.

[2] Govind Kumar Jha et al. proposed Fertilizers and Suitable Crops using Machine Learning Algorithms. In the proposed system, it is observed that samples of different soils are collected from the land and then sent to laboratories, checked and announced back to farmers. Many machine-learning approaches and algorithms have been used to analyse soil. It is observed that the proposed system achieved greater precision than all other algorithms. Finally, it is proved that the random forest is better than the all other algorithms regarding different indicators in performance. This algorithm has the highest precision and recall values and is more suitable for soil classification than other algorithms.

[3] Shiva Roshan et al. described an intelligent chemical fertilizer recommendation system for rice fields. The dataset selected for this study needed to be completed, and missing data were reported. The goal was to recommend the proper amount of supplementary urea, TSP, SOP, and ZnSO4 for the rice field. Random forest, XGBoost, The genetic algorithm was used to implement the SVM and SVM algorithms. The results of the experiment were compared to ANN methods based on the Particle PSO and C.S. SVM was found to be a better choice for predicting. The amount of soil moisture compared to the neural network artificial neural network that was implemented and experimented with for estimating the amount of phosphorous in the Kuhn drainage basin in Qazvin Province, Iran. Eighty-five samples were gathered from 1000 hectares of the selected location for experimenting with the phosphorous contents.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11,S Iss 3, Dec 2022

[4] Zujiaoshi et al. proposed a modified convenient method to maintain a better relationship between the crops; we need to understand the relationships between how much the crop-yielding capacity requirement of N so that we can minimize the levels of nitrate. We can also recommend the double cropping system, like wheat in winter and maize in summer. Here the wheat in the winter season will be linear, and the maize in the summer season will be exponential. Determine the goal yield and N.R. for following crop based on the long-term insight data to help determine the required fertilizer application levels.

[5] Dr. P. PandiSelvi et al. proposed a fertilizer recommendation based on soil using a Short- or Long-Term Memory Algorithm. It is observed that the proposed model was built to identify diseases in leaves and soil types and recommend a suitable fertilizer for the crop, which helps the farmers achieve more productivity. In the proposed model, the model was organized in different ways importing the data, soil analysis, leaf disease identification and fertilizer prediction is done through comparing and classification. In the existing systems, the process of identifying the disease in leaves and finding the soil type and fertilizer preferred were carried out manually, which is prone to various disadvantages. Hence the proposed system is done using short- or long-term memory algorithms.

Table-1: Existing System Analysis

S.NO	TITLE	AUTHOR	PROPOSED METHODOLOGY	MERITS	DE-MERITS	ACCURACY
1	Automatic Leaf Colour Chart (LCC) Recommendation of Nitrogen Fertilizer for Paddies	Zahra Hatam	Convolutional Neural Network (CNN), Decision Tree	The suggested filter and kernel size extract features cohesively, reducing misclassification risks while using the CNN Method.	Fertilizer recommendations in saline areas are highly challenging since nutrient administration may impact a plant's tolerance to salt, making crop yield projections more complex.	Decision Tree - 91.22%, CNN-94.22%.
2	A model to Recommend Appropriate fertilizers and the crops for the agriculture.	Govind Kumar	Random Forest	This methodology improves the predictive performance, which helps the farmers to choose the right crop to cultivate.	A recommendation is only based on the dosage of different nutrients, but other aspects must be considered.	80%

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11,S Iss 3, Dec 2022

3	Rice Fields' Fertilizer Recommendation System For Intelligent Chemical	Shiva Roshan	Support vector machine, Random forest, X.G. boost, Genetic algorithm.	Random forest is used to tackle the problem of missing data present in the dataset.	soil moisture becomes more or less than its defined allowed limit, the process of plant growth will be problematic	88%
4	An improved fertilizer recommendation for nitrate.	Zujiaoshi	Convenient method	N fertilizer improved crop output and rose even more when supplemented with P fertilizer. .	During the crop growing season, no supplemental irrigation was provided.	89%
5	A fertilizer recommendation based on soil using a Short- OR Long-Term Memory Algorithm	Dr. P. PandiSelvi	Short And Long Term Memory Algorithm	It is noted that the suggested model was constructed to recognize illness in leaves and type of soil and recommend an appropriate fertilizer for the crop, which significantly assists farmers in achieving greater output.	In previous systems, the process of recognizing illness in leaves, determining soil type, and selecting fertilizer was done manually, which had several drawbacks.	87.29%

3. PROPOSED SYSTEM

The proposed model utilized the Fertilizer Recommendation data set from Kaggle. The data set has some numerical values as well as categorical values. The class-label which is considered in this data set is categorical type. The ability to use categorical features directly with CatBoost without encoding is one of the distinguishing characteristics of CatBoost over other boosting techniques [6]. So this system employed CatBoost Algorithm and perform the activity of recommendation. The Figure 2 represents the block diagram of our proposed system.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11,S Iss 3, Dec 2022

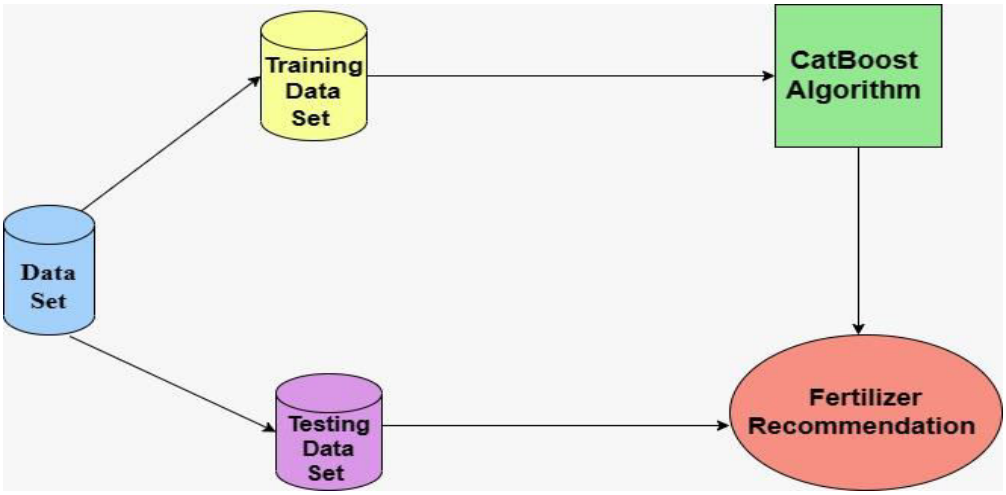

Figure 2: Proposed System Flow Diagram

3.1 Dataset Description:

This paper utilizes the Fertilizer Recommendation dataset, which is publicly available on Kaggle. This dataset consists of 9 columns and 100 sample records. There are eight attributes which are independent and act as significant factors in recommending the suitable fertilizer. It has a dependent variable called as Fertilizer name, which is used as a recommender for developing the proposed model. Since the class label contains seven unique values, such as DAP, Urea, etc.

3.1.1 Training Data Set

The initial dataset used to train or fit the machine learning model is the training data, which is the largest (in terms of size) subset. The ML algorithms are initially supplied the training data, which enables them to learn how to generate predictions for the given task.

3.1.2 Testing Data Set

After using the training dataset to train the model, it's time to put it to the test. The model's performance on this dataset is assessed, and it is confirmed that the model generalises effectively to new or untested datasets. The original data from the training dataset are divided up into the test dataset.

3.2 CATBOOST ALGORITHM:

CatBoost is one of the gradient boosting algorithms based on decision trees. During training, a sequence of decision trees is created. Each succeeding tree is created with less loss than the prior ones. One of the unique features of the Catboost algorithm is integration of working

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11,S Iss 3, Dec 2022

with many data kinds to tackle a wide variety of data challenges encountered by multiple enterprises.

Features Of Cat Boost Algorithm:

Figure 3: Features of CatBoost Algorithm

Robust: CatBoost has the potential to improve model performance while reducing overfitting and tweaking time. Despite the fact that CatBoost has many parameters to tune, the default parameters produce excellent results without the need for extensive hyper-parameter tuning.

Accuracy: The CatBoost method is a one-of-a-kind gradient boosting method that is both fast and greedy.

Categorical features support: CatBoost supports operating with non-numerical factors, which reduces time spent on pre-processing data and improves training results.

Easy Implementation: It is a very powerful machine learning algorithm and library that anyone can easy to use. It is designed in such a way that parameter tuning is rarely necessary, and the default parameters are optimal.

Faster Training and Prediction : CatBoost utilises distributed GPUs, which allows it to learn and predict 13-16 times faster than other algorithms.

Supporting Community of Users : The inability to contact a team when you have a problem with a product you use can be extremely inconvenient. CatBoost is not one of them. CatBoost has a large community, and the developers are always looking for feedback and contributions.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11,S Iss 3, Dec 2022

3.3 RECOMMENDATION:

The recommendation system will analyze the values in the parameters that is entered by the user and display the suitable fertilizer for the soil. This recommendation helps in timely guidance to improve the soil fertility which in turn increases the yield.

4. RESULTS AND DISCUSSIONS:

$$\text{Precision} = \frac{\text{True Positive}}{\text{Actual Results}} \text{ or } \frac{\text{True Positive}}{\text{True Positive + False Positive}}$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{Predicted Results}} \text{ or } \frac{\text{True Positive}}{\text{True Positive + False Negative}}$$

$$\text{Accuracy} = \frac{\text{True Positive + True Negative}}{\text{Total}}$$

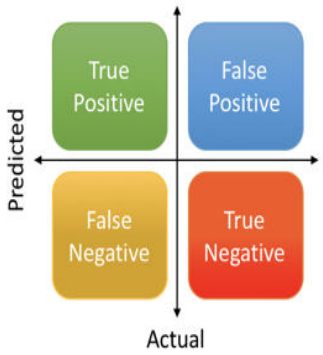

Figure 4: Computations of Metrics

Precision and Recall are the metrics that are used to evaluate the machine learning algorithms. Precision gives the proportion of the positive predictions that are actually correct and recall measures the proportion of the actual positives that are predicted correctly. Accuracy is used to understand the performance of the classification model.

	precision	recall	f1-score	support
0	1.00	1.00	1.00	3
1	1.00	0.75	0.86	4
2	0.50	1.00	0.67	1
3	1.00	1.00	1.00	1
4	1.00	1.00	1.00	5
5	1.00	1.00	1.00	2
6	1.00	1.00	1.00	4
accuracy			0.95	20
macro avg	0.93	0.96	0.93	20
weighted avg	0.97	0.95	0.95	20

Figure 5: Performance Metrics Of Proposed System

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11,S Iss 3, Dec 2022

According to the papers that have been studied, the accuracy provided by the KNN is 92%. The current model employs the CatBoost algorithm to recommend the best fertiliser for the soil. Using the CatBoost algorithm, the accuracy was increased to 95%.

5. CONCLUSION:

The following limitations are identified in the papers we reviewed; one of them is the system that recommends fertilizer and insecticides intelligently based on TPF -CNN for smart farming after displaying the results regarding the type of pesticides and insecticide in the web page that it cannot save the results in the system, other limitation in crop yield prediction, forecast and fertilizer recommendation does not consider other parameters like Climatic conditions such as light, Temperature, Water, humidity, and other limitation in the nutrient suggestion method based on evolutionary computing for soil fertilization is the optimization issue. Hence, the IoT sensors are need to be chosen because they can overcome the above problems and produce good results.

REFERENCES:

[1] Zahra Hatam, Mohammad SadeghSabet, MohammadJafarMalakouti, Ali Mokhtassi-Bidgoli,MehdiHomaee journal, 2020,https://doi.org/10.1016/j.sajb.2019.11.032

[2] Govind Kumar Jha1 PreetishRanjan, Manish Gaur, A Machine Learning Approach to Recommend Suitable Crops and Fertilizers for Agriculture, 2020

[3] Shiva HassanjaniRoshan, Intelligent Chemical Fertilizer Recommendation System for Rice Fields, AJAST, Volume 5, Issue 3,2021.

[4] ZujiaoShi, Donghua Liu, Miao Liu, Optimized fertilizer recommendation method for nitrate residue control in a wheat–maize double cropping system in dryland farming, journal 2021. https://doi.org/10.1016/j.fcr.2021.108258

[5] Dr.P. PandiSelvi, P. Poornima, Soil Based Fertilizer Recommendation System for Crop Disease Prediction System, Volume 8, Issue 02,2021.

[6]Sri Silpa Padmanabhuni and Pradeepini Gera, “Synthetic Data Augmentation of Tomato Plant Leaf using Meta Intelligent Generative Adversarial Network: Milgan” International Journal of Advanced Computer Science and Applications(IJACSA), 13(6), 2022. http://dx.doi.org/10.14569/IJACSA.2022.0130628