Iot Based Dynamic Operations to Automate the Machinery Tools for Agriculture

Sushma Gururaj Kulkarni

Department of Computer Science

RV College of Engineering

Banglore, India

sushmakulkarni4877@gmail.com

Prof. Ganashree K C

Department of Computer Science

RV College of Engineering

Banglore,India

Article Info
Page Number: 1294-1303
Publication Issue:
Vol. 71 No. 3 (2022)

Article History
Article Received: 12 January 2022
Revised: 25 Febuary 2022
Accepted: 20 April 2022
Publication: 09 June 2022

Abstract— *Climate change has had a negative impact on the performance of the huge percentage of India's crops over the past 20 years. Prior to harvest, crop yield predictions and fertilizer prediction would aid farmers and policymakers in deciding on the best course of action for marketing and storage. Before cultivating on an agricultural field, this project will help farmers determine the yield of their crop and will also suggest the right fertilizer for the crops, assisting them in making the right choices. It makes an attempt to solve the problem by creating a prototype of an interactive prediction system. The dataset is collected from the IOT sensors, data like soil moisture, temperature, humidity, soil type, NPK. The collected dataset is processed using data analytics technique. The cleaned and pre-processed data are trained on machine learning algorithms. The machine learning algorithm helps to analyse the crop yield for the next sowing and also suggest the fertilizer for the better yield. An easy-to-use web-based graphic user interface will be implemented in such a system as the farmer will be informed of the outcome of the prediction. As a result, there are numerous methods or algorithms for big data analysis in predict yield and fertilizer prediction, and with the aid of those algorithms, one can predict crop yield and suggest fertilizer using algorithms like the Random Forest algorithm(RF) and SVM.*

Keywords: Agriculture, Machine learning, Crop Yield Prediction, Fertilizer Prediction

Introduction

The agricultural industry is essential to human survival and the economy. Crop selection in traditional farming methods relied on farmers' basic knowledge. Most of the time, farm owners prefer to select the crop that is most in demand in their region or nearby. Land fertility is negatively impacted by a lack of crop rotation and a lack of scientific understanding of farming**.**[1] Soil nutrients, groundwater level, and fertilizer type are major determinants of crop quality. A traditional farmer deals with ongoing difficulties. The wrong crop selection and insufficient soil nutrients can result in increased soil acidity. The main factor affecting crop quality and yield is the unstable climate. For the health of the crop and the proper crop selection, soil fertility is crucial.[2]

The goal of data analysis is to find useful information and conclusions by inspecting, cleaning, and modelling data Analysis, information extraction, and prediction from a large amount of data are steps taken in order to spot patterns.Companies use this process to transform customer data from its raw form into information that is useful. The field of agriculture can also benefit from this analysis. Most farmers relied on their extensive field experience with a particular crop to predict a higher yield during the upcoming harvest season, but they still don't receive a fair price for their produce. Most often, it occurs as a result of improper irrigation, poor crop selection, or occasionally, a lower-than-expected crop yield. To estimate the total actual production functions for the yields of various crops in designated states while taking into account a variety of technological factors and a recently created weather index. To make a fair comparison between our actual result, also known as the target, and the prediction model, which has an approachable user interface for farmers and provides an analysis of yield predicting and fertilizer prediction based on available data, regression and coefficient of determination analysis as well as average error rate were conducted.

In India, there are 15 agro-climatic regions that are split up according to the type of land. Some particular crops can be grown in each agroclimatic region [3]. Based on that, you should advise the farmer on the kind of crop that would thrive in a given climate. Research aims to maximize crop yield while spending the least amount of money possible. Therefore, make better use of information technology to comprehend crop selection based on climate region. Hence, high crop yield production results in profitability and sustainability. So, one of the parameters in the proposed method for determining the best crop to cultivate is temperature.[4]

Motivation

India is the world's leading crop producer, surpassing the United States, and this makes sense given the country's massive population. Thousands of tons of crops are lost annually, as either a result of climatic factors or as a result of people not understanding the cultivation cycle. For the purpose of estimating his or her income, every farmer wants to understand how much yield can be expected from their farm. When choosing which farming method to do and in what crop to grow in which climate, Indian farmers must overcome numerous obstacles. The agricultural sector has a direct impact on the country's overall economic development. Indian land is adapted to a variety of climates and soil types, so little advance planning is necessary. Historical datasets can be used for this. Many studies on crop yield have been conducted in the past by various researchers, but each has its own flaw. As a result, this study can always be improved because the techniques can be adjusted to perform better.[5]

Related Work

Smart farming systems highlight the importance from using cutting-edge technologies in agriculture, such as the Iot technology (IoT), the Cloud, and machine learning. Conventional farming methods are being accelerated by the digital revolution in order to produce more crops of higher quality. Previous attempts to combine Machine learning and artificial intelligence with smart sensors and guide farmland practitioners in the correct direction were unsuccessful. We propose a four-layer architectural model with sensor, network, service, and application layers to deploy a smart farming system with low energy consumption. Furthermore, focusing on the application layer, we use deep learning to develop a fertilizer recommender that corresponds to expert advice. The results of the entire system are then presented to farmers in a single mobile application for their convenience.[6]

Using ML algorithms and Internet of Things sensors to gather data, Colombo-Mendoza et al. presented the concept for a smart farming method. 2 categories of sets of data data and harvest production data—are combined to predict crop yield using a novel data analysis methodology [7]. Khongdet et al. suggested a framework for control and tracing smart crops using real-time iot data. SVM is used to detect crop diseases. Fertilizer suggestions are also made based on prior land data [8]. Bakthavatchalam et al. presented a smart module for recommending an appropriate crop for farming in order to maximize crop yield. The WEKA tool is used to analyze data by ML algorithms. JRip, a multilayer perceptron rule-based classifier, and the decision table classifier were used for classification.[9]

R. Ghadge el concluded that by utilizing a few classification techniques and comparing various characteristics, this work aids in increasing agricultural production rates. Various data mining techniques were investigated in order to predict crop yield. Among the algorithms included for comparative analysis are artificial neural networks, support vector machines, K-Nearest Neighbors, decision trees, random forests, gradient-boosted decision trees, regularized greedy forests, and the proposed CSM technique (Crop Selection Method), which aids in predicting the order of crops that can be taken into account for planning in the upcoming seasons[10].

A significant development in the field of smart agriculture is the use of soil analysis attributes and real-time sensing data to select crops. Bhojwani et al. proposed a model with three modules: crop choice, cultivation practices, and crop maturity. They used variables like moisture in the soil, temp, humidity, air pressure, and quality of air along with weather conditions for higher yield selection and health monitoring. The Thing Speak application used the KNN algorithm to evaluate real-time sensor data. Patil et al. suggested a methodology for crop selection by combining various temp, soil, humidity, as well as infrared sensors with such a sensor to collect real-time data. For the purpose of predicting crops, certain data mining techniques are used to preprocess data and compare trained data with real-time data. Crop price increases mentioned on the National Commodity and Derivative Exchange were also considered for crop prediction.[11]

In all of 2017, Dakshayini Patil [12] stated that the rice crop plays an important role in economic growth. The yield of the rice crop was predicted using a variety of data mining techniques. India's sustainable security comes from its rice crop. In general, it makes up 40% of the overall yield. The right climatic conditions are essential for a crop to yield a high yield. Crop yield can be increased by developing a better plan for growing the crop in accordance with the climatic conditions. The reports use a variety of mining techniques based on historical information about yield of the crop and various geographic regions. In order to forecast crop yield, the authors used data from 27 regions of Maharashtra.

According to Eswari et al. 2018 [13], the perception, average, minimum, and maximum temperatures all affect crop yield. In addition, they added a new attribute called crop evapotranspiration. The weather and plant growth stage both affect the crop's evapotranspiration. The above element is taken into consideration when deciding how much the groups will yield. They all came together to compile the range of data with these attributes, fed it into the Bayesian network, and then divided it into the correct and incorrect classes. A confusion matrix was used to compare the observed categories in the model to the expected classifications in the model, bringing the accuracy up. They led to the realization that Naive Bayes and Naïve bayes (nb)offer high accuracy in crop yield prediction.

Proposed Methodology

The limitations of the primitive methods currently in use will be eliminated by the proposed system. According to various parameters, including humidity, gases, moisture, and temperature values, it compares the current data with historical data obtained from the Department of Horticulture and Agriculture. During the cultivation process, farmers can test the soil several times and take the necessary precautions to get good yield. Reports will be produced at the conclusion so farmers can track their fertility. To estimate the overall land's approximate fertility, the results of each test are averaged.

For crop yield prediction and fertilizer prediction, studies usually used.csv files of agricultural dataset. It is supervised learning for the dataset. It includes different characteristics like the name of the county, the state, the temperature, the soil moisture, the soil type, the yield, etc.

All test results will be shown on a display screen. A web page is created and the results are displayed as and when, the farmer physically enters the displayed value into his web application. Using the results of all the tests will reflect the ratio of the various nutrients, such as nitrogen, phosphorus, and potassium, present in the soil and will include the averaged result of all the tests. The list of appropriate crops and the amount of fertilizer needed for the land will also be included.

A.Dataset description

The sample data set for this project is as follows. Based on seven factors, the data in Fig 1 and Fig 8 are used to forecast crop yield. State, District, Crop, Area, Season, and Production are these 7 elements. By using this data, we can build a machine learning model, train it, predict yield production, and determine which kind of fertilizer should be applied to achieve the desired yield. The amount of nitrogen and phosphorus are the input parameters; the type of the appropriate fertilizer to use is the output.

B.Random forest algorithm

Random forest is a supervised learning algorithm that is used for two very different classification and regression. The random forest technique constructs decision trees from diverse datasets, estimates the information out of each subset, and thereafter uses user voting to determine which solution is right for the system. The data was trained in Random Forest using the bagging method. The bagging method was used to train the data in Random Forest. The bagging model employs the study of various models in enhancing the system's end result.[14]

The original datasets are initially passed as a random sample to each decision tree using a replacement called sampling. As a result, each decision tree will generate its own prediction. The outcome is determined by majority vote. Each subset of data is shared with each decision tree, resulting in highly accurate results. The algorithm performs well in classification when comparing regression and classification problems. Random Forest, another ensemble model, was chosen because it reduces overfitting issues and variance, both of which tend to boost accuracy. It can automatically handle missing values and is extremely stable, so any change to one decision tree won't have an impact on another. Furthermore, have less influence over the noise.[15]

The Random Forest algorithm, which provides accuracy for model-based predictions and the actual results of predictions in the dataset, was used to obtain high accuracy. A decision tree is constructed from a sample of data in the random forest, and the trees provide predictions from each family. Voting determines which solution is best, which raises the accuracy of the model. It delivers the system's best outcomes.

C. System architecture

The system architecture diagram for the proposed prediction system in Fig.1, Initially, datasets are collected using IOT sensors, then pre-processed and categorized into two categories: one for crop yield prediction and the other for fertilizer recommender. Following pre-processing, features such as Crop Type, Soil Type, Temperature, Season, States, Districts, and so on are available. The datasets are split and trained in an 8:2 ratio.80 percent of the data are used for training.

When crop yield prediction model is applied with algorithms like svm, linear regression, and random forest and the accuracy is compared, random forest performs better. The random forest algorithm is used for the fertilizer prediction model. when a user enters information about a crop type, an area, a state, a district, and so forth. The amount of yield produced based on the input provided is the output for the crop yield prediction. The fertilizer prediction model's output is the type of fertilizer that should be applied to the given input.

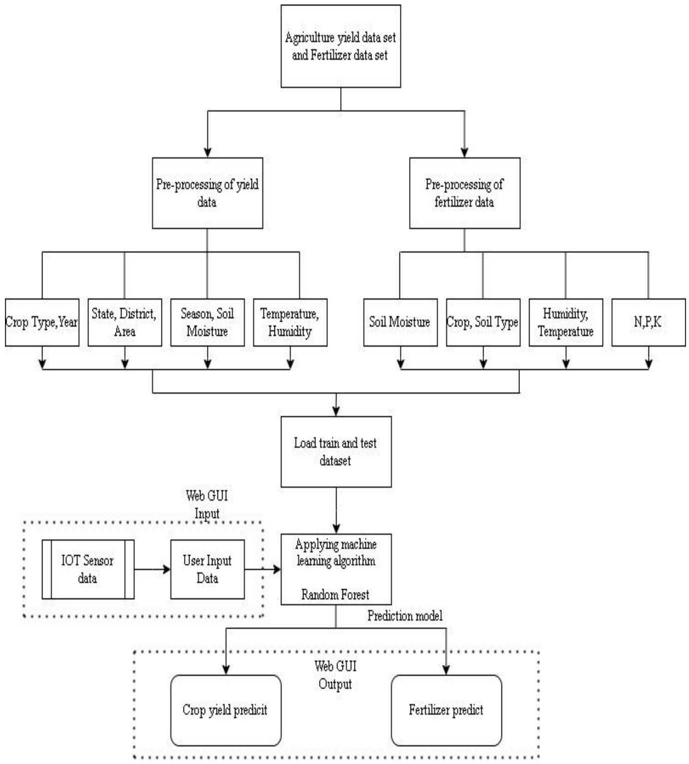

Fig 1 System Architecture for the prediction model

Results

The 2 prediction models are built here, one is crop yield predictor and other is fertilizer prediction model. The below are the results for the proposed methodology.

The Crop Yield Predictor

The crop yield predictor in the Fig 2 describes about the dataset of crop yield contains State name, District name, Crop year, Season, Temperature, Humidity, Soil Moisture, area, Production.

	State_Name	District_Name	Crop_Year	Season	Crop	Temperature	humidity	soil moisture	area	Production
0	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Arecanut	36	35	45	1254.0	2000.0
1	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Other Kharif pulses	37	40	46	2.0	1.0
2	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Rice	36	41	50	102.0	321.0
3	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Banana	37	42	55	176.0	641.0
4	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Cashewnut	36	40	54	720.0	165.0
...	...	...	...	...	...	...	...	...	...	...

Fig 2 Dataset for Crop Yeild

The crop yield interface Fig 3 allows the user to enter the data and get the output for the yield prediction which has to be sown.

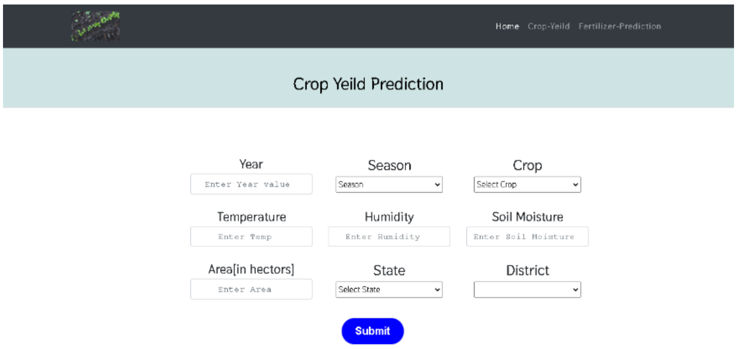

Fig 3 The web interface for crop yeild

In the Fig 4, it descibes about the heat map for the yeild prediction model, to realise the strength among the variables present in the dataset.The red area is the being most interacted region.

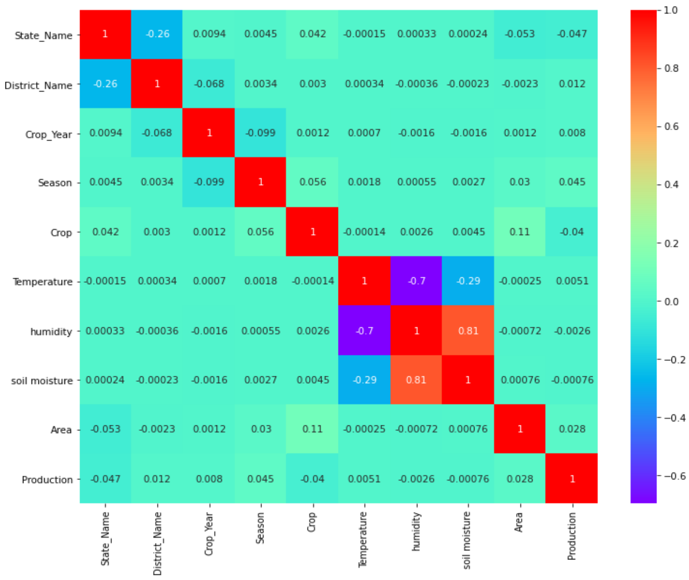

Fig 4 Heat map for Crop Yield

The Fertilizer predictor

The fertilizer predictor in the Fig 5 describes about the dataset of fertilizer contains Temperature, Humidity, Soil Moisture, Soil Type, Crop Type, Nitrogen, Potassium, Phosphorous, Fertilizer name.

	Temparature	Humidity	Moisture	Soil Type	Crop Type	Nitrogen	Potassium	Phosphorous	Fertilizer Name
0	26	52	38	Sandy	Maize	37	0	0	Urea
1	29	52	45	Loamy	Sugarcane	12	0	36	DAP
2	34	65	62	Black	Cotton	7	9	30	14-35-14
3	32	62	34	Red	Tobacco	22	0	20	28-28
4	28	54	46	Clayey	Paddy	35	0	0	Urea
...	...	...	...	...	...	...	...	...	...

Fig 5 Dataset for Fertilizer Prediction

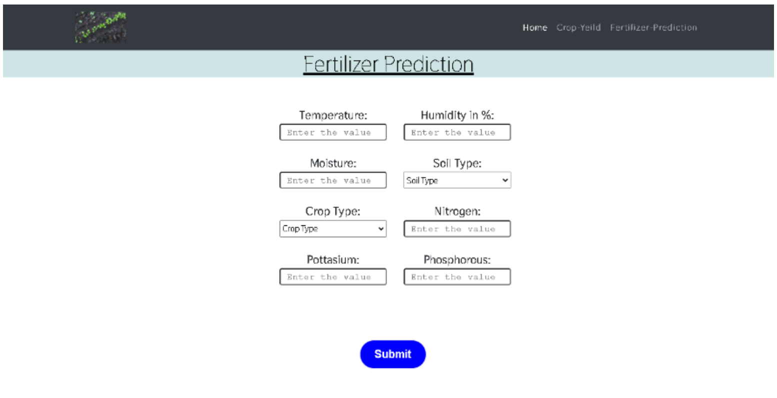

Fig 6 The web interface for feritilizer

The fertilizer prediction model interface Fig 6 allows the user to enter the data and get the output for the yield prediction which has to be sown.

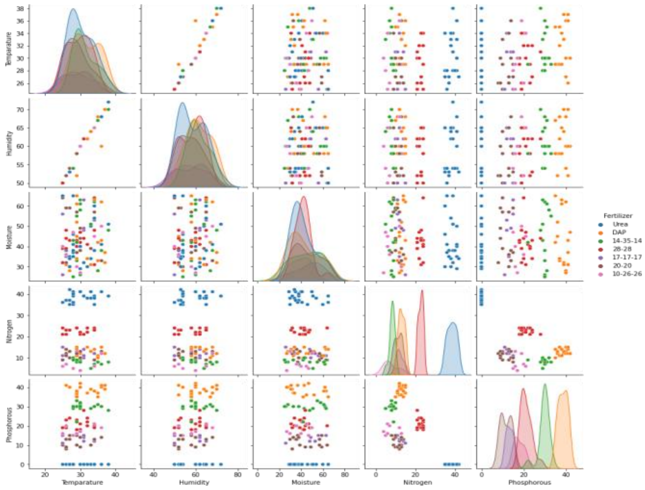

Fig 7 Plotting of categorical and continous data

In the Fig 7, plotting of categorical and continous data is carried out. The pair plotting helps to understand the pairwise relationship between the different variables present in the dataset.

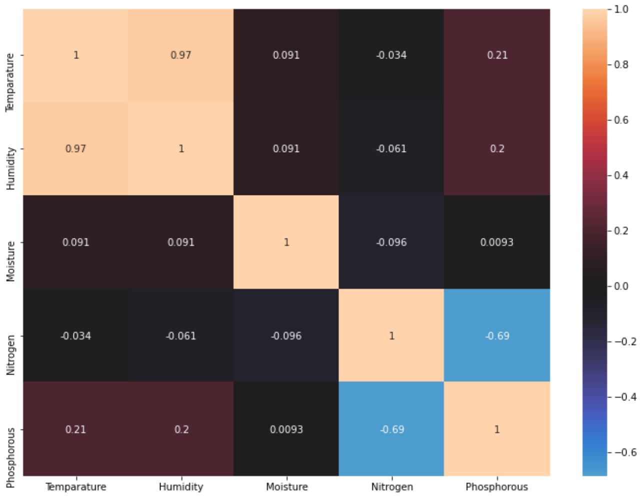

Fig 8 Heat map for Fertilizer

In the Fig 8, it descibes about the heat map for the fertilizer prediction model, to realise the strength among the variables present in the dataset.The light area is the being most interacted region.

In the above both models random forest algorithm is used as the accuracy for the systems is more than other algorithm like svm, linear regression. Hence random forest is better.

Conclusion

The proposed work focuses on crop yield prediction and fertilizer detection to forecast the best crop to cultivate in the field, as well as the necessary conditions for cultivation and yield. Farmers can use this to improve their plans and increase production, which boosts the overall profitability of the nation. Additionally, gave the farmer regarding the information on the fertilizer needs to help him or her get a better crop yield. Despite a high accuracy value provided by the accuracy score, the model is constrained by a small amount of data. Any climate disasters that have already occurred while gathering data or that might occur after prediction are not taken into account by the model. Geospatial analysis can be incorporated into the model in the future for better data and accuracy.

References

1. P. Kanchan and N. Shardoor, “Krashignyan: A Farmer Support System”, AJCT, vol. 7, no. 3, pp. 1-7, Dec. 2021.
2. F. Javed, M. K. Afzal, M. Sharif, and B. S. Kim, “Internet of Things (IoT) operating systems support, networking technologies, applications, and challenges: a comparative review,” IEEE Communication Surveys and Tutorials, vol. 20, no. 3, pp. 2062–2100, 2018.
3. S. Veenadhari, B. Misra and C.D. Singh, "Machine learning approach for forecasting crop yield based on climatic parameters", 2014 International Conference on Computer Communication and Informatics, pp. 1-5, 2014, January
4. A. k. Gajula, J. Singamsetty, V. C. Dodda and L. Kuruguntla, "Prediction of crop and yield in agriculture using machine learning technique," 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), 2021, pp. 1-5, doi: 10.1109/ICCCNT51525.2021.9579843.
5. S. Mishra, P. Paygude, S. Chaudhary and S. Idate, "Use of data mining in crop yield prediction," 2018 2nd International Conference on Inventive Systems and Control (ICISC), 2018, pp. 796-802, doi: 10.1109/ICISC.2018.8398908
6. B. Swaminathan, S. Palani, K. Kotecha, V. Kumar and S. V, "IoT Driven Artificial Intelligence Technique for Fertilizer Recommendation Model," in IEEE Consumer Electronics Magazine, doi: 10.1109/MCE.2022.3151325.
7. L. O. Colombo-Mendoza, M. A. Paredes-Valverde, M. D. P. Salas-Zárate, and R. Valencia-García, “Internet of Things driven data mining for smart crop production prediction in the peasant farming domain,” Applied Sciences, vol. 12, no. 4, 2022.
8. K. Phasinam, T. Kassanuk, and M. Shabaz, “Applicability of Internet of Things in Smart Farming,” Journal of Food Quality, Article ID 7692922, p. 7, 2022.

9. K. Bakthavatchalam, B. Karthik, V. Thiruvengadam et al., “IoT framework for measurement and precision agriculture: predicting the crop using machine learning algorithms,” Technologies, vol. 10, no. 1, p. 13, 2022
10. Ghadge, Rushika, et al. "Prediction of crop yield using machine learning." *Int. Res. J. Eng.Technol.(IRJET)* 5 (2018).
11. Y. Bhojwani, R. Singh, R. Reddy, and B. Perumal, “Crop selec-tion and IoT based monitoring system for precision agricul-ture,”International Research Journal of Engineering andTechnology (IRJET), vol. 4, no. 2, 2017
12. Dakshayini Patil et al., "Rice Crop Yield Prediction using Data Mining Techniques:An Overview", *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 7, no. 5, May 2017
13. K. E. Eswari and L. Vinitha, "Crop Yield Prediction in Tamil Nadu Using Baysian Network", *International Journal of Intellectual Advancements and Research in Engineering Computations*, vol. 6, no. 2, ISSN 2348-2079.
14. N. Valarmathi, N. Srija, S. Ananthi, B. P, E. T and N. N, "Amenity Agriculture App using Random Forest Algorithm," 2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), 2022, pp. 522-527, doi: 10.1109/ICSCDS53736.2022.9760901.
15. J. R, H. D and P. B, "A Machine Learning-based Approach for Crop Yield Prediction and Fertilizer Recommendation," 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI), 2022, pp. 1330-1334, doi: 10.1109/ICOEI53556.2022.9777230.