International Journal of Scientific Research and Engineering Development-– Volume 9 Issue 4, July – August 2026
Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Weather Based Crop Prediction

Sujit Chavan
Department of Master of Computer Applications (MCA)
JSPM University, Pune, Maharashtra, India
Email: chavansuji17@gmail.com

--************************--

Abstract:

Agriculture is a vital source of both food for lots of people in the world and income-generating economic activity. Changing weather conditions can dramatically alter this situation, however. The amount of rain, temperature, and humidity can all greatly affect overall crop production; soil conditions can also create variabilities that are often not in a farmer's control.

A weather/predictive model system uses historical data and machine-learning algorithms and will be able to accurately predict what kind of crops should be planted based on present weather conditions. It identifies patterns and correlations between a variety of factors that influence crop growth. It also classifies these patterns, so as a result of using the system, farmers will have better decision-making potential leading to less risk of failure and better use of their resources.

While I believe that using a model like this in making decisions regarding which crops to grow is beneficial, I still need to understand how all of the variables in this group of conditions are interrelated. There is also a difference between using historical data to make these predictions versus real-time data.

--************************--

I. INTRODUCTION

Background Study

Agriculture is a fundamental part of virtually all nations' economies, as well as an essential source of food security for the growing global population. The success of agriculture in producing enough food is directly correlated to environmental and weather factors such as temperature, precipitation, humidity, and soil characteristics. Farmers are encountering difficulties related to climate variability and extreme weather events. The change in climate has resulted in farmers not being able to accurately determine which type of crop will perform the best on their farm. Not selecting the appropriate crop can negatively influence the yield, profitability, and resource utilization of a farmer.

For many years, farmers relied on their own experiences, cultural norms related to cropping practices, and previous records from suppliers or their own historical records to make decisions regarding crop selection. Although these sources of information have been very useful and dependable in the past, the rapid pace of change in the environment may mean these sources will no longer provide an accurate representation of the productivity of a particular crop for any given farm. Thanks to ongoing technological improvements, especially in the fields of data science and machine-learning, there are numerous new opportunities for improving how information is gathered and utilized in agricultural operations.

Farmers can make decisions based on data analysis of large amounts of past weather data and crop history to better understand the types of crops that would be successful based on historical weather trends and crop growth patterns. Additionally, farmers can use current weather data (e.g., temperature, humidity, precipitation, soil moisture content) to determine the most appropriate crops to plant for the current growing season. Machine learning techniques are available that can assist farmers with identifying different weather-related problems that will ultimately lead to betterCrop-yield potential based on current weather conditions.

As more data has become available and computers have improved, the use of machine learning in agriculture has increased. In particular, machine learning will be used more frequently to classify environmental conditions that support certain crops using various classification (e.g., Random Forest, Decision Tree, Support Vector Machine (SVM), and K-Nearest Neighbors (KNN)) and regression techniques. In summary, machine learning can assist farmers with making more accurate predictions about which crops will perform well given the historical weather data they maintain along with the data they acquire each year for their growing season. A Weather-Based Crop Prediction System offers a means of implementation for precision agriculture through data-driven decision making by providing an avenue through which farmers are provided with crop recommendations prior to planting based on an analysis of the current weather conditions at that time which will improve the utilization of resources and support more sustainable practices in farming.

Problem Statement

Weather-related issues such as temperature, rainfall, humidity, etc., have a considerable influence on agricultural productivity, and most farmers experience substantial difficulty in determining their best possible crop selection options due to climate instability and the lack of reliable decision support systems that allow accurate forecasting/analysing prior to planting. The main methods of selecting crops have been based on farmer's experience and past practice, and consequently, they do not carry sufficient levels of dependability when crop production is influenced by continuously evolving weather patterns/conditions.

Inaccurate crop selection can result in lower yields, the poor utilization of resources, financial loss, and higher rates of crop failure. Farmers face increasing difficulties in accurately predicting which crop type will perform best through the changing types/conditions of weather that will occur this year, correlated with climate change. Therefore, it is necessary for farmers to have access to a sophisticated weather-related data analysis tool that provides accurate crop selection options based on forecasted weather related factors.

This project will solve the problem of predicting crops based upon weather parameters (such as humidity, temperature, and rainfall) (and) the environmental parameters that will affect the crop's ability to be grown. The agricultural industry has many elements that show how a crop will perform but not how the crop will do in its environment. Our objective is to provide a data-driven-weather based crop model, based upon accurate weather data, for farmers to utilize in determining the best crop to grow for their area to improve yield, reduce risk and encourage sustainable agricultural practices through implementing the correct crop with accurate information.

Research Gap

Many machine learning based crop prediction systems have been developed in the past; however, many of those studies are limited to a small set of environmental parameters or a small number of geographic areas, making them less useful across multiple agricultural scenarios. In addition, many of the common methods used for crop prediction models are based on historical data and aren't able to incorporate or accurately represent the rapidly changing climate and weather differences throughout their lifetime.

There are a variety of different crop recommendation systems today; however, many of these models have issues associated with them such as: not being accurate; not having enough features selected for predicting each crop type accurately; and, being unable to predict the various weather impacts on growing a given crop at the same time. Many systems will either be too complicated to understand without having experience using them or require a significant amount of training to use successfully. Finally, many of the existing models specifically predict yield but do not provide a recommendation for which crop to grow prior to the crop being cultivated.

Another major shortcoming exists in the development of a system that integrates various types of models into one comprehensive model to analyze different forms of weather parameters (temperature, humidity, rainfall, etc.) so that an accurate recommendation of crops can be made. Many existing predictive systems fail to reliably predict crop production based on variable climatic conditions and are therefore unable to generalize to new datasets.

Consequently, the need for a strong and dependable Weather-Based Crop Prediction System that incorporates machine learning techniques in analysing the various weather parameters to determine which crops are best suited, is necessary. Such a system can improve prediction accuracy, aid on making data-driven decisions relating to agriculture, reduce potential crop failures, and ultimately contribute to sustainable agricultural practices.

Objectives to be achieved through the proposed research are as follows:

the purpose of this research is to create a Weather Based Crop Prediction System that can recommend the optimal crops given a set of weather conditions and variables, based on the use of machine learning techniques.

1.Collect and review relevant datasets from agriculture and weather sources (e.g. temperature, humidity, rainfall, soil conditions, etc.) to identify all relevant parameters for crop predictions.

2.Clean and prepare the dataset by eliminating missing data, correcting inconsistencies and converting it to the appropriate format for input into machine learning models.

3.Determine how weather parameters and crop selection are related through both data analysis and the evaluation of features.

4. Create and train machine learning models to predict what type of crop would be best to grow given certain types of weather conditions.

Research Questions

1. What kinds of crops should be grown given certain temperatures, humidity levels and rainfall?
2. Can machines use past weather data to make predictions about the best types of crops to grow?
3. What weather data can be relied on most heavily when making a crop selection?
4. Which machine learning algorithms will provide the best results for predicting crops based on their past weather?

Significance of the Study

The purpose of this research project will assist with increasing crop production by helping farmers make better decisions regarding which type of crops they should grow. Farming, like all industries, relies on certain environmental conditions to be successful; thus, using machines to make this type of decision will aide farmers in maximizing their profits because they have followed the appropriate selection of crops based on weather. The weather-based crop prediction system will provide farmers with information on the weather in order to provide a basis for selecting a particular type of crop for planting based on their growing conditions

Farmers can use this research to help reduce uncertainty with which crops to plant while also minimizing the risk of planting crops that can be damaged due to weather conditions. The prediction framework can provide accurate crop recommendations that the user can use to improve their crop yield, make better use of their farmland and agricultural inputs, and lower their financial risks associated with selecting poor crops. These types of intelligent prediction models will increase the efficiency of farming practices and provide support for sustainable agriculture.

The research also adds value to the field of agricultural technology by illustrating how machine learning algorithms can solve real-world farming challenges. It provides an approach for integrating weather data with predictive analytics to support precision farming. Researchers can use this research as a resource for developing sophisticated agricultural prediction models and for building agricultural decision-making systems.

II. REVIEW OF THE LITERATURE

Agriculture is a key component in many countries' economic growth and is subject to a large influence from the environment and weather. Growth of machine learning and increased use of data analytics by researchers to create intelligent systems to predict the crop being harvested and to aid in agricultural decision making has increased the attention given to weather-centric crop prediction for the benefit of

farmers by helping farmers to select crop species that will be grown based on the weather conditions (i.e., temp, humidity, rainfall) encountered during the growing season.
Research studies have focused on the use of machine learning methods to predict the best crops to grow. Traditionally, farmers relied on previous years and their appearances to determine what crops to plant for many years. There was little scientific evidence indicating how changes in the weather would affect crop production, leading to poor crop selections at times due to the unpredictability of climate variability.

To address these shortcomings, researchers have begun to use data-driven methods that employ weather and agricultural datasets to predict "what crops to grow where" with much greater accuracy than traditional methods of problem-solving.

The Decision Tree algorithm is one of the machine learning methods commonly used for predicting the best crops to grow based on their weather and environmental conditions, by classifying the types of crops based on these same conditions.

The Decision Tree provides a very simple and interpretable model that can allow farmers to understand how their various input factors relate to the recommended crops that they should plant. However, when Decision Trees are used on large datasets training, they can suffer from overfitting.
As a result, many agriculture prediction systems have adopted Random Forest algorithms as an alternative to Decision Trees to mitigate the effects of overfitting. Random Forest algorithms are sets of Decision Trees that are combined to produce much higher prediction accuracy and lower levels of overfitting than would have been calculated using a single Decision Tree.
Numerous studies have shown that Random Forest algorithms produce better performance when predicting suitable crops than do other types of algorithms or classifiers that rely solely on weather factors. For this reason, this algorithm handles large datasets more effectively than other algorithms that have multiple input variables and are preferred for use in agriculture.

The Support Vector Machine (SVM) is one of the most common machine learning techniques that have been applied in crop prediction. SVM performs very well when working with high-dimensional datasets and is able to classify crops at a high level of accuracy. Researchers have used SVM models to evaluate data on temperature, rainfall and humidity in the context of recommending which crops to grow, demonstrating an ability to produce reliable results. However, due to the large datasets generated by agriculture, SVM can require careful tuning of its parameters as well as increased computational resources.
A K-Nearest Neighbor (KNN) has also been applied in agricultural prediction studies. The KNN algorithm classifies crops based upon how similar the new input data looks like compared to the historical records of that crop. While KNN is very simple to implement, its efficacy is very much reliant upon the user selecting the value of K and using quality training data. KNN studies show it performs well on smaller datasets, but can pose problems with efficiency when dealing with large-scale agricultural datasets.
Recent studies have looked at integrating multiple weather parameter inputs to improving the overall prediction performance. Studies have shown that by combining features for temperature, humidity, rainfall and soil-related characteristics may significantly increase the accuracy of crop recommendation systems. Additionally, feature selection methods have been used to determine the most influential parameters on crop suitability.

The use of deep learning solutions has become an advanced way to predict what will happen with agriculture. Artificial neural networks (ANN) and deep neural networks (DNN) have demonstrated a distinctly positive capability at modeling the complex interactions between crop yields and climate changes. DNN and ANN are classified as capable of learning from the large datasets they analyze, allowing them the potential to have more accurate predictions than earlier prediction models. Like many artificial intelligence (AI) applications, the challenge to develop effective DNN and ANN is that they require many more resources to compute and a large volume of inputs before they can provide useful information.

Comparative research has been done comparing performance of different machine learning algorithms for agricultural crop yield predictions. In the results of this research were the comparative results showing that ensemble methods (such as random forest) provided better prediction accuracy and performance than traditional classifiers. Even though experimental results demonstrate this, the most useful comparative factors in choosing which type of model to select are related to the dataset being worked with, computational requirements of the predictive modeling process, and business objectives related to crop yield prediction.

Despite all the advances that have been made in developing growing systems for crop yield predictions, there are still many continuing challenges in the development of these types of systems. Most crop prediction systems have been designed for specific regions and therefore will not perform well outside of those specific regions. Additionally, most of these systems also have been designed to predict crop yields only, rather than to recommend to the user a specific crop to grow. Finally, as most agricultural business are experiencing, predicting crop yields has become more difficult due to changes in weather caused by climate change. All of this represents the opportunity to develop better performing and more adaptable to use systems for prediction of crop yields.

III. METHODOLOGY

Research Design

To establish a Weather-Based Crop Prediction System utilizing Machine Learning techniques, this research implements a quantitative and experimental design and methodology. The project will also collect, process, and analyze weather-related agricultural data to predict which crops would best suit environmental conditions. Through an experimental approach, different machine learning algorithms will be assessed based on their ability to correctly predict crop yields.

The project consists of multiple phases of research, including (1) data Collection, (2) data Preprocessing, (3) Feature Selection, (4) Model Development, (5) Model Training, (6) Model Testing, and (7) Model Performance Evaluation. The sources of the data are historical weather (attributable to temperature, humidity, rainfall) and agricultural datasets, which will contain attributes (crops grown) corresponding to the weather during the farming period.

Once the agricultural/weather data has been collected, it will go through several preprocessing steps to remove inconsistencies, eliminate missing values, and prepare the dataset to be used for machine learning analyses.

Once preprocessing is completed, the dataset will then be divided into its training and testing sets. The training dataset will be used to train machine learning algorithms while the testing dataset will evaluate the prediction accuracy of those algorithms. Examples of classification algorithms that may be used to classify weather conditions and crops include random forests, decision trees, support vector machines (SVM), and k-nearest neighbors (k-NN).

Several methods of performance evaluation for developed models using Accuracy, Precision, Recall, F1-Score and ROC AUC analysis along with confusion matrix are established to measure the success of the proposed crop prediction system as well as its reliability.

A systematic approach provides an accurate means of achieving objective evaluations of the various models developed in a workflow manner to produce accurate crop predictions to ascertain whether agricultural productivity can be increased with the assistance of a decision support system by the implementation of Machine Learning through analysis of weather and crop datasets will assist growers to identify crops best suited for cultivation.

[1] Research Design Flow

1. Data Gathering
   - Weather and crop datasets will be obtained.
   - Attributes used will include: temperature, humidity, rainfall, etc…and crop type.
2. Data Preprocessing
   - Addressing missing values.
   - Removing duplicate and inconsistent records. Normalizing/transformation of data in

preparation for feature selection (this will also aid testing after model development).

3. Feature Selection
• Identify important weather parameters impacting crop prediction.

4. Model Development
• Utilize Machine Learning classification algorithms.
• Train Models on historical data.

5. Model Testing
• Testing the previously trained models using unseen data.

6. Performance Evaluation
• Calculating performance metrics, including: Accuracy, Precision, Recall and F1-Score.
• Generate accompanying Confusion Matrix and ROC-AUC Curve for model performance evaluation.

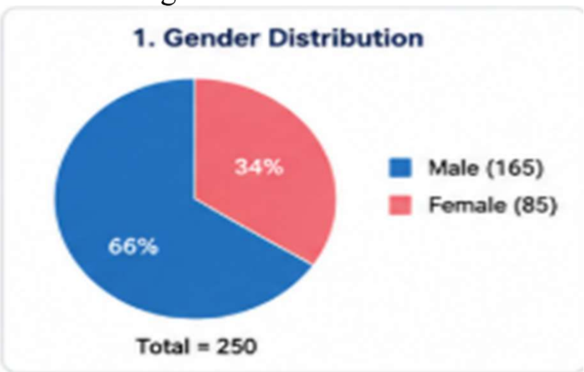
7. Crop Prediction
• Assisting in the selection of the most appropriate crop based on prevailing weather conditions.

Demographic Analysis

Figure 1: Gender Distribution

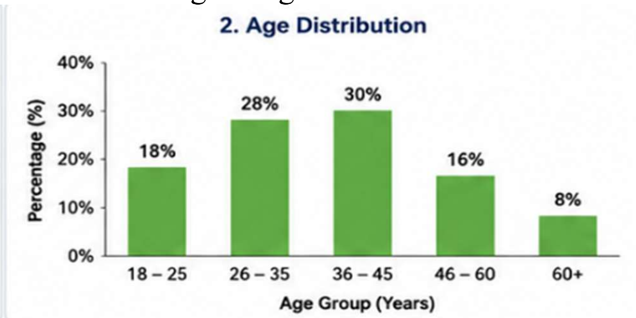
IV. Results and Analysis

Figure 1:- Gender Distribution



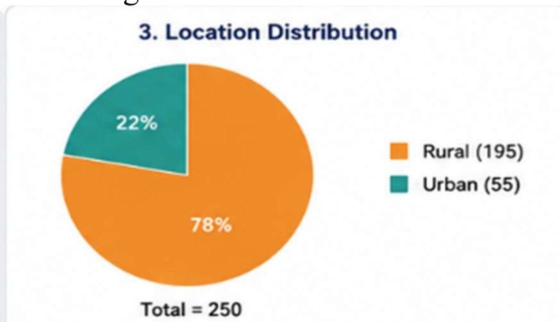
Analysis: Figure 1 (Pie Chart) provides the distribution of 250 total male and female study participants (weather-based crop prediction). • Male = 165 (66%) • Female = 85 (34%) Therefore, male subjects make up most (66%) of the total sample; females make up 34%; thus indicating most of these participants' perspectives and experiences will be from male farmers, as is typical for most agricultural communities where the majority of farm work is performed by men.

Fig. 2: Age Distribution



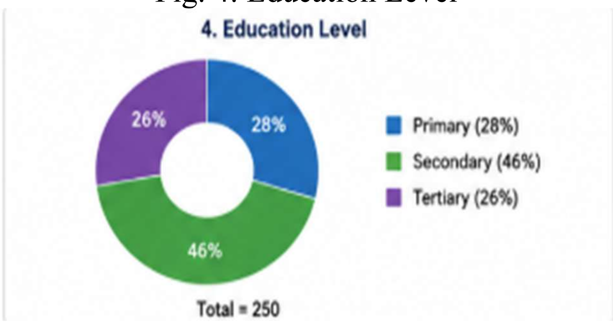
- **Analysis:** The greatest number (30%) of total subjects is in the 36–45 year age group.

Fig. 3: Location Distribution



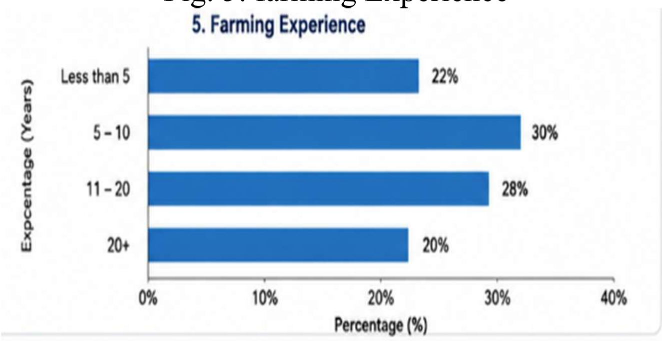
Analysis: Examination of the location distribution for the 250 participants in the Weather-Based Crop Prediction System study shows that 195 (78%) of the participants reside in rural areas, and 55 (22%) of them reside in urban areas. This indicates that a large number of participants are farming in rural areas, thereby establishing a primary relevance of the Weather-Based Crop Prediction System within agricultural-based communities located in these rural areas. The extent of the number of participants from rural areas reflects the need for weather-based crop recommendations to assist with farming practices and enhance productivity within the agricultural sector.

Fig. 4: Education Level



Analysis: The education level of the participants in the Weather-Based Crop Prediction System study is summarized in the table below (Figure 4). 28% - Primary Education 46% - Secondary Education 26% - Tertiary Education In sum, the majority of participants (46%) have completed their education through secondary education, thus demonstrating a moderate level of formal education for most participants. Another 28% of the participants completed their education through primary education, while 26% were educated to a level of tertiary education for this study.

Fig. 5: farming Experience



Analysis: The Response (the proportion of all respondents based on the number of years that they have been involved in some type of farming) responses were broken out as follows: • 22% of the respondents reported having less than 5 years' experience • 30% of respondents reported having 5-10 years' experience • 28% reported 11-20 years' experience • 20% reported having more than 20 years' experience The group of respondents with 5-10 years' of farming experience (30%) were the largest group of respondents followed closely by 11-20 years' experience (28%). This illustrates that a majority of the participants have a good deal of agricultural experience and knowledge about how to farm.

V. Discussion

Information collected during the analysis of the demographic information provided valuable insight about the characteristics of the Weather Based Crop Prediction System survey participants. One of the major findings is that male participants comprised 66% of all survey respondents suggesting that men dominate agricultural activity in the study area. Although 34% of respondents were female, their participation also demonstrates the importance of women in the agriculture-related decision-making process. The age distribution of respondents shows that the overwhelming majority of them are in the 26-45 age range, which comprises 58% of all respondents. This suggests that the majority of survey respondents are active and productive farmers who are more likely to adopt new agricultural technology and incorporate data into their farming practices. The lower participation rate of older farmers could show that older farmers are less likely to adopt technologies that are related to the use of data.

According to the location analysis, 78% of respondents reside in rural areas, indicating a very strong correlation between the Weather-Based Crop Prediction System and the way it will benefit the farming population that lives in rural areas. The majority of agriculture occurs in these areas; consequently, the system has great potential to assist a farmer's decision-making regarding crop selection with respect to the predicted weather conditions.

The analysis regarding education levels shows that close to half of the respondents possess secondary degrees, while approximately one-quarter have attained a degree from an institution of higher learning. This indicates that respondents possess enough of an educational background to use technology in agriculture and are therefore likely to use the proposed system. The proposed system must therefore be designed so as to provide a simple and user-friendly interface for all users regardless of their educational levels.

Most of the respondents possess between five and 20 years of real world experience in farming according to the distribution of years of farming experience among the participating farmers, which suggests that they have developed an extensive base of experience to evaluate the practicality of weather-based crop recommendations and, therefore, give the study results a higher degree of reliability and provide support for the proposed system's reasonable application.

VI. Conclusion

The demographic evaluation of the Weather-Based Crop Prediction and Decision Support System yielded essential information regarding the characteristics of the respondents and future users of this system. The assessments revealed that the respondents were predominantly male farmers, aged between 26 and 45, living in rural locations, possessing secondary education (i.e. a high school diploma or equivalency), and had extensive experience working in agriculture.

This suggests that the system will be of great benefit financially for those users actively involved in agricultural operations because weather/satellite technology can assist them with farming activities. The significant number of rural respondents also indicates the need for the development of a system that provides support to an agricultural population in which the weather significantly impacts crop yields. In addition, the respondents' education and experience working in farming imply that they can appropriately use weather/satellite-based recommendations to make informed decisions regarding their crop production plans.

In conclusion, this demographic analysis indicates the existing Weather-Based Crop Prediction and Decision Support System will be readily accepted by farmers and serve as an excellent tool to support their decision-making processes. Therefore, the Weather-Based Crop Prediction and Decision Support System will help to improve the overall productivity of agriculture, lower the risk of crops failing to produce, and promote sustainable farming practices.

VII. Limitations

- Dependence on Quality of Data Set (Accuracy of Prediction): The dependence that the quality of data set used to build a prediction model has a major effect on the accuracy of the model. If the data used to build the model is not complete or accurate, then the quality of the prediction would be reduced.
- Limited Weather variables: The model created by this system considers only variables that are directly related to weather (e.g., temperature, humidity and rainfall); however, other important variables such as soil nutrients supplied to crops, pest infestations in and around crops, and methods used by farmers to maintain their crops are not accounted for.
- Regional Variability: The model is trained using a specific data set which has its own characteristics, so it is not likely that this model will perform equally well in all geographic regions that may have distinctly different types of climatic and agricultural conditions.
- Uncertainty in Prediction, Due to Climate Change: Rapid and severe weather changes may occur due to climate change, so data from past weather events may not accurately account for rapid changes due to climate change; thus affecting the accuracy of historical data used in predictions.

VIII. Future Scope

1. Integrate soil parameters- To improve the prediction accuracy, soil parameters (e.g., pH, nitrogen, phosphorus, potassium and soil fertility) should be put equally with other weather parameters.
2. Implement more advanced machine learning algorithms/technologies- improve the performance of the predictions through the use of newer technologies such as Deep Learning, Artificial Neural Networks (ANN) and Ensemble Learning.Real-Time Weather Data Integration
3. Connect the system with live weather forecasting services to provide dynamic and up-to-date crop recommendations.
4. Integrate Real-time Weather Information- Connect to live weather forecasting services that provide continuous change in weather will provide up-to-date recommendations on what crops to plant.

Acknowledgment

I would like to thank the authors, researchers, and developers determined and provided useful data sources for this research. The various contributions helped me to understand machine learning, data analysis, and agricultural prediction systems. In addition, I want to give special thanks to my fellow friends and coworkers, who provided me with encouragement, assisting suggestions and help throughout the project. Their assistance has been significantly helpful to the completion of this project.

References

[1] Rajeswari, S, Suthendran, K. "C5.0: Advanced Decision Tree (ADT) Classification Model for Agricultural Data Analysis on Cloud" Computers and Electronics in Agriculture 156 (2019), p. 530-539, 2019.

[2] Sujatha, R. and Isakki, P. A Study on Crop Yield Forecasting Using Classification Techniques International Journal of Computer Applications 181 (11), 2018, pp. 1-5, 2018.

[3] Jeong, M, Kim, J, Choi, H. Crop Prediction Model Using Machine Learning Techniques for Smart Agriculture. IEEE Access 8 (2020), p. 195112-195122, 2020.

[4] K. R. Patel and M. P. Patel, "Machine Learning-Based Crop Advisory System," International Journal of Scientific Research in Computer Science, Engineering and Information Technology, Volume 6, Issue 3, Pages (450-456), 2020.

[5] A. Kumar, S. Kumar, and V. Reddy, "Prediction of Crop Growth Using Weather Data: A Random Forest Approach," Proceedings of the International Conference on Intelligent Computing and Smart Communication, 2021, Pages (215-220).

[6] P. Sharma, A. Srivastava, and S. Saini, " Crop Growth Prediction and Advisory System Based on Machine Learning," International Journal of Advanced Research in Computer Science, Volume 11, Issue 2, Pages (25-31), 2020.

[7] J. Han, MG. Kamber, and J. Pei, Data Mining: Principles and Applications, Third Edition, Burlington, Massachusetts: Morgan Kaufmann, 2012.

[8] T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition, New York: Springer, 2009.

[9] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning, Cambridge, Massachusetts: MIT Press, 2016.

[10] F. Pedregosa, et al., "Scikit-learn: Machine Learning in Python," Journal of Machine Learning Research, Volume 12, Pages (2825-2830), 2011.

[11] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," Proceedings of the International Conference on Learning Representations (ICLR), San Diego, California, 2015

[12] A. Géron, Tell me, what is Machine Learning? 3rd ed. Sebastopol, California, United States: O'Reilly Media, 2022.

[13] Food and Agriculture Organization of the United Nations (FAO), FAO's State of Food and Agriculture - 2024. Rome, Italy: FAO, 2024.

[14] United Nations Intergovernmental Panel on Climate Change (IPCC), Climate Change 2023: Synthesis Report. Geneva, Switzerland: IPCC, 2023.

[15] J. Brownlee, Python Programming Masters Course .Melbourne, Australia: Machine Learning Mastery, 2020.