

A Microclimate based Crop Recommender System for Precision Agriculture

AI Mustarik

American International University-
Bangladesh
Department of Computer Science
Dhaka, Bangladesh

Md. Tawhid Sultan

American International University-
Bangladesh
Department of Computer Science
Dhaka, Bangladesh

Md. Mahidul Islam

American International University-
Bangladesh
Department of Computer Science
Dhaka, Bangladesh

ABSTRACT

Agriculture is not only necessary for producing food but also provides the raw material for other industries, like the fashion industry, dairy industry, sugar industries, etc. Agricultural production is one of the most effective ways to support a country's economy. In this paper, it has been tried to figure out predicting crop production using microclimate data. While doing so, the IoT system has been used here to make this research more efficient and accurate. In this paper, Using IoT devices and machine learning techniques, a crop recommendation model has been proposed. A cloud server-based system has been built to store the machine learning model and a database of previous data readings. This crop recommendation model is beneficial for farmers as well as for researchers. Django Framework has been used to create a web application that shows the predicted results to the farmers based on the machine learning algorithm. And the environment, where the introduction of such a system will reduce farmer risk, save money and time, and reduce the waste of agricultural commodities.

Keywords

Microclimate & Precision agriculture, Regression Analysis, Crop recommendation.

1. INTRODUCTION

Most of the developing countries' economic growth depends on their agriculture sector. As per the UN SDG goals, a country can't achieve sustainable development without erasing hunger for all of its populations [30]. Having a strong agricultural system is necessary to achieve this goal. But every year a substantial portion of the world population faces a shortage of food. The scarcity of food is caused by several factors and one of the major reasons is crop failure. Crop fails for so many reasons. One of the major reasons is selecting the wrong crops for a specific location without considering its geographical features.

A specific climate is a subtle need for all of us. Most of the previous work was based on machine learning and data analysis. Using the previous data predicting crop production is one of the most efficient approaches. But most of the previous work was theoretical representation and they haven't introduced microclimate in their works. Also, not so much use of the IoT system. which can be so useful for this kind of prediction model.

So, a research gap has been found that very few researchers have looked at this approach. The main focus of this research work is on the relation between microclimate and crop production.

2. BACKGROUND STUDY

The microclimate is a local atmospheric region where the atmosphere varies from the surrounding area. In Paper [28], many Statistics show that in different evolutions, microclimates are often overlooked, despite increasing evidence that microclimates are important for different production dynamites and processes. The paper [29] describes the need for crop microclimate studies on the ability of agroecological management to reduce climate effects concerning quality parameters [ex. temperature, humidity, rainfall, etc.].

Eco physiological models are used mostly for simulating the impacts of ecological factors on agricultural and natural biological systems [20]. Rising atmospheric CO₂ levels significantly affected climate change. Which affected the global food supply [21]. Increasing carbon dioxide in the atmosphere results in warmer air which is accountable for accelerated plant phenology and also reduces crop yields by 10-40% [22]. A. Abdullah, et al. said in their research that agricultural productivity is affected by two factors of climate-induced effects, one of them is temperature and precipitation and another one is an indirect effect on soil moisture [23] [24].

Using data mining applications for crop prediction is another significant feature for data science by using the weather parameters and crop production of the previous year's data. S. Veenadhari et al. proposed a decision tree model for crop yielding prediction using the ID3 algorithm [25]. In 1983 Quinlan, et al. proposed an algorithm called Iterative Dichotomiser 3 (ID3). The major limitation of this algorithm is that, does not give the proper attributes for that's why this algorithm is not suitable for attribution [1]. Kushwaha, et al. uses an algorithm named the "Agro algorithm" implemented in the Hadoop platform to manage large volumes of data sets [2] [3]. Snehal, et al. have proposed a model, which uses meteorological data to predict crop yields based on the "Artificial Neural Networks" (ANNs) approach that would find the most successful crop yielding factor. But some factors like measurements are difficult and also cost-effective [4]. In paper [5], based on available data Multiple Linear Regression (MLR) used in the technique and Density-based clustering technique for crop yield prediction. Researchers have also introduced supervised machine learning techniques to predict crop production. Many of them used KNN Regression to predict the output. But they haven't discussed working with real-life data and also, they have a lack of proper data collection which they have worked with [6]. Veenadhari et al. tried to develop a website to assess the effect of climate parameters on crop production based on the C4.5 algorithm used to set various Decision rules for prediction [8]. Various

Supervised Learning algorithms are used to compare the productivity of the wheat yield based on soil and crop data in paper [9]. Linear regression has been conducted regarding the prediction of agricultural crop growth [12]. A Multilayer Perceptron Neural Network has been developed for validating the data by using the cross-validation process. This research describes the development of a predictive model of rice crop yields using ANNs [13].

Many technologies and techniques of IoT have been created to reduce labor strain and to quickly do the hard work. Kamilaris, et al. in [14] introduced an IoT-based framework for the real-time processing observation and justification of inputs in the agricultural environment, based on semantic web technologies. By analyzing the correlation Analysis between historical statistics and real-time data by the IoT application, the agricultural forecast was attempted to support the information system in [15]. Khattab et al. introduced a cloud-based IoT architecture in [16] which could be used in various applications for agricultural precision. The architecture proposed is composed of several layers. Besides that, Paper [17] proposed an IoT system-based concept for interconnected farms consisting of three main components, connected IoT devices with sensor and controller tracking, an IoT gateway called the &Cube, and an IoT service platform called the Mobius. Rodríguez et al developed a WSN within a greenhouse of roses which enables data collection, visualization in a web or mobile application, and generating a predictive model using machine learning techniques [18].

Scientists and researchers are trying to focus on these agriculture-related fields based on microclimate for better crop production. But most of the works were theoretical implementation and only very few of the works have real-life implementation and examined with real-time data. But the model that has been proposed in this paper is going to work with the real-time data and also going to work along with the machine learning techniques. An IoT based architecture has been designed and built where real-time data can be collected by using different kinds of sensors and a predefined machine learning model will be used where those data can be analyzed and give the prediction result based on the data and for visualization of all these, small web applications have been developed.

3. DATASET PREPARATION

3.1 Data Source and Description

For the proposed model two types of data set were needed. One is data of crops and the other is weather data of the Bangladesh region. The dataset used for this paper is adopted from the years 1988 to 2013. The crop production data has been collected from the Bangladesh Bureau of Statistics, a government organization [1]. And the Weather data has been collected from Kaggle [2].

3.1.1 Crops Data

In the crop's dataset, there is information about 6 major crops (Aus, Amon, Boro, Jute, Potato & Wheat) of Bangladesh. Crop production information of 64 districts is in the dataset. The data has been collected financial year-wise.

3.1.2 Weather Data

The weather data in the dataset has been collected from

different weather stations around the country. Monthly data has been taken in the dataset.

3.2 Dataset Integration

3.2.1 Data Processing

In the crops production dataset, there was not much information about 64 districts' crop production. And also, in the weather dataset but for this model, only three attributes have been used. they are temperature, rainfall, and humidity. So, 23 districts have been selected for the data analysis according to the availability of weather and crop production data.

3.2.2 Reducing the dataset

The dataset used for this paper hasn't many missing values. Some of the information was not much accurate in the crops dataset so one crop data has been removed from the main dataset.

3.2.3 Replace missing value with mean

Only a few numbers of missing values were there in the dataset that has been used for the paper. So, the row which contains missing values was replaced with the mean value.

3.3 Data Combined and Final Dataset

Districts and years that are the common attribute in both datasets. In the combined dataset there is information about 23 districts' weather and crop production. The attributes in the combined dataset are year, districts, production area, production per acre, temperature (minimum and maximum), humidity, and rainfall. To make the dataset more usable and efficient Min-max normalization has been applied on those datasets.

4. PROPOSED MODEL

To describe the proposed model this section has been divided into two parts. In the first part, the whole architecture model will be elaborated, and in the second part how machine learning works in this model will be described.

4.1 Architecture

Based on the proposed architecture (see Figure 1), some individual elements or methods will be applicable that work together to create the model and help to deploy it as a service. This proposed architecture is discussed below as several components. Firstly, to configure the prototype and to introduce the IoT implementations, certain equipment that would have to be discussed at first will be listed below:

Sensors: This proposed model requires multiple sensors for the data in real-time. Three sensors have been used for that. These are the sensors of (1) temperature, (2) humidity, and (3) rainfall. The sensor which can be used for monitoring temperature and humidity value is DHT22. Water sensors can be used to detect rainfall. Water sensor block is designed to measure water, and can also be frequently used to track rainfall, water level, and even liquid leak.

Microcontroller: Most IoT gadgets use microcontrollers of any kind. In this proposed architecture a microcontroller named "Arduino UNO" is used. The Arduino Uno is an open-source microcontroller board based on the Microchip ATmega328P microcontroller.

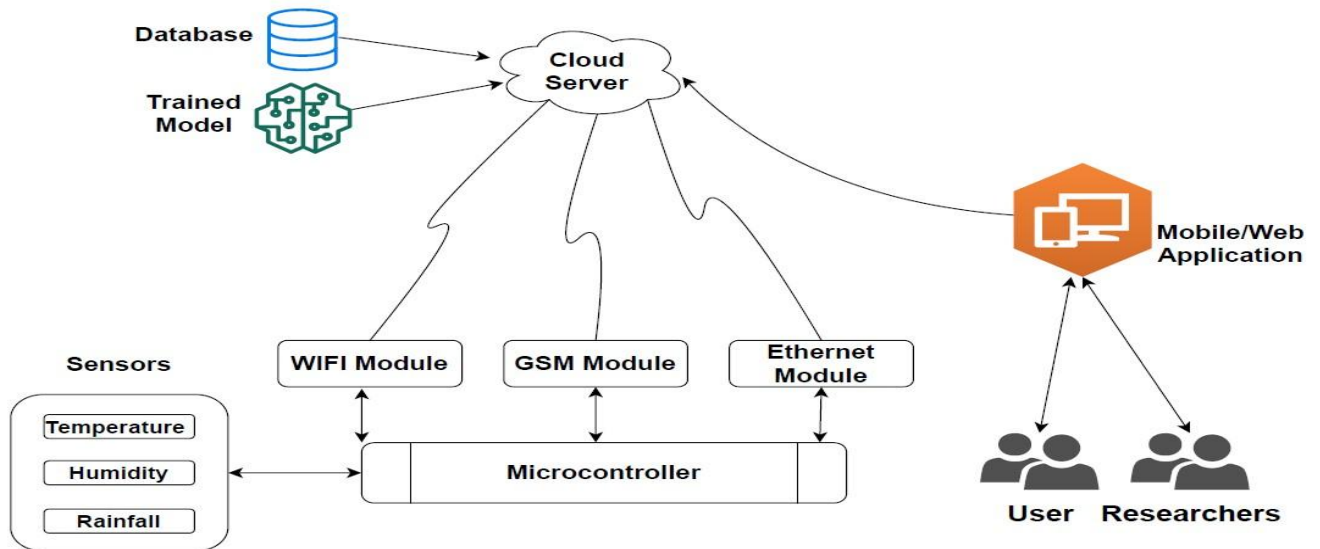


Fig 1: Overall model for crops prediction

Ethernet Shield: Ethernet shields can be used for delivering internet access from Arduino to the server with an ethernet module. Ethernet Shield provides internet connectivity to the Arduino board through its Ethernet library. Sensor data will be collected to Arduino and these values will then be stored on the data servers via ethernet shield.

Cloud Server: This cloud server comprises certain components such as train model for the proposed system and a database repository where numerous different information records are kept. The individual parts and their operations are listed below

Database: It has already been planned for the database management system with 2 forms of table configuration in it. One is about customer information where the user details specifics and also the productId from the defined approach that one will use. Another table is about the data which will come from reading the sensor against a particular product Id. To a single user, this product Id is the unique identifier. Such data are used in the small web application that has been proposed in this model and the prediction that was wanted to make for more implementations.

Trained Model: The model of machine learning requires the results to decide the features which best predict the results. This train model should be able to be set up through the dataset from the machine learning algorithm. How this model train is built will be defined in the next section.

4.2 Machine Learning Model

For the implementation of the proposed system, three algorithms have been used and the algorithms are a. Linear Regression, b. Polynomial Regression, c. Gradient Boosted Tree.

But before this algorithm is used there are several steps ahead of it. How the whole data set is integrated and how all these items are combined is listed in section 3. The entire set of data is then divided into two parts, the train, and the test dataset. Eighty percent of the entire data set is used for model training using those three data mining algorithms mentioned above. The trained model is then saved for further prediction in the cloud server and performance analysis is also performed using the test data set.

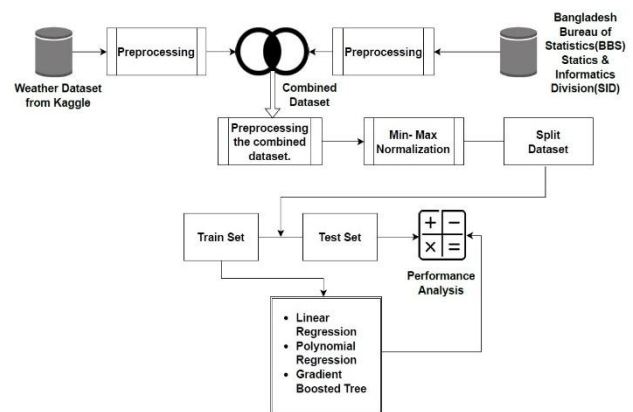


Fig 2: Dataset preparation for crops prediction

5. EXPERIMENT

For the proposed IoT model users are connected through a device. In the web application, users must register in the system. In the device, one will be identified as a system user by a unique ID that generates all the information. When a user is connected to the system, the device (see Figure 3), sends data to the cloud server.

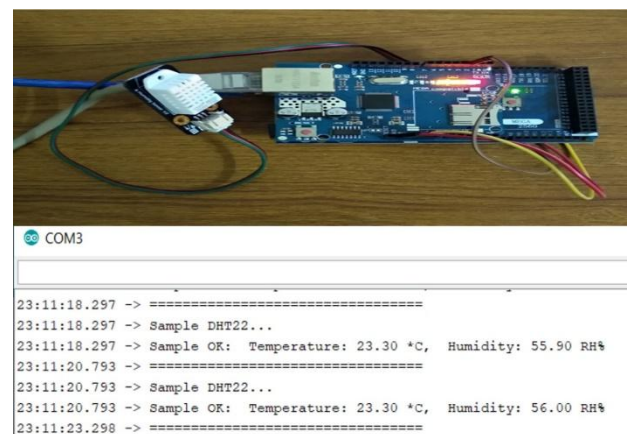


Fig 3: Collecting data using “DHT22” sensor

In the cloud server, the predefined train model and database

are stored. Here the user's data works as a test set and by analyzing this data the application can give the result to the user.

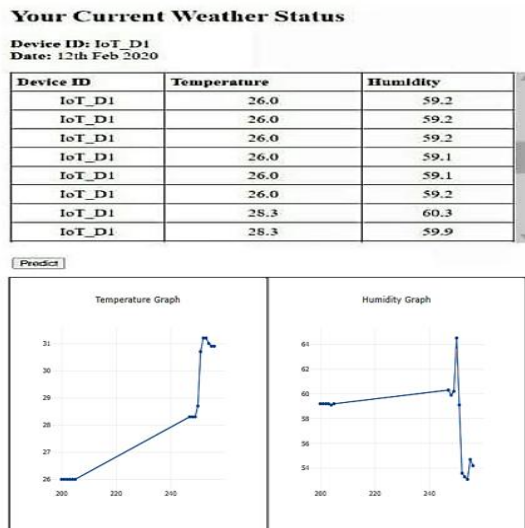


Fig 4: Web application page to show one user's data

To build this model, Arduino UNO, Arduino Ethernet Shield, and sensors are used. Arduino Ethernet Shield and sensors are connected through the Arduino UNO and send data to the cloud server. There's a small web application that is designed by using the Django framework for model prediction. Through that web application, a user can know the details of his device readings as shown in (see Figure 4).

If any user then wants to know the expected crop production rate based on their data then this can be obtained which is demonstrated in (see Figure 5).

Predicted Result for All Crops

Crops Name	Predicted PPA
Potato(M Ton)	2.862
Wheat(M Ton)	1.365
Amon Rice(M Ton)	3.332
Aus Rice(M Ton)	1.253
Boro Rice(M Ton)	0.783

Fig 5: Predicted result in the web portal

From this table, one can see the prediction result according to the real-time data collected from the IoT device. And here it can be seen that Potato's production rate is higher in this weather condition than in other crops. This is how the proposed model is going to be worked.

6. RESULT ANALYSIS

By using Regression analysis, it indicates the significant relationship between dependent and independent variables. It also shows the effect of intensity for a dependent variable of several independent variables. In this study, three regression algorithms have been used named Linear Regression, Polynomial Regression, Gradient Boosted Tree Regression. Based on weather and previous production data and tried to predict future production. The evaluated results are shown below (see Table 1).

Table 1. R² and Root Mean Squared Error Value in all Algorithms

LINEAR REGRESSION					
	Aus	Amon	Boro	Potato	Wheat
R ²	0.487	0.485	0.232	0.456	0.468
Root Mean Squared Error	0.13	0.078	0.028	0.144	0.04
POLYNOMIAL REGRESSION					
	Aus	Amon	Boro	Potato	Wheat
R ²	0.18	0.138	0.175	0.194	0.189
Root Mean Squared Error	0.164	0.1	0.029	0.175	0.05
GRADIENT BOOSTED TREE					
	Aus	Amon	Boro	Potato	Wheat
R ²	0.427	0.482	0.091	0.389	0.165
Root Mean Squared Error	0.137	0.078	0.031	0.153	0.05

7. CONCLUSION AND FUTURE WORK

Agriculture plays a vital role in a country's financial status. This paper represents a future agricultural improvement with the help of machine learning algorithms and also IoT devices. The main target of this paper is to present a prediction model using different kinds of data mining and machine learning algorithms. Where a model will work to complete the whole process based on regression analysis. A device has been built using different kinds of microcontrollers and also has built a website using the Django framework. This proposed model can help the farmer to choose the right crop for yielding according to the weather and also can be used in many other ways. The limitations are negligible compared to success. There are a few of the districts that have no weather station, so we took the nearest weather station's data of the districts.

By adding a GPS module, it can be possible to trace the device and in case any damage occurs then it'll be easy to take action in a short time. And all of these tasks would be handled by using an advanced mobile application and the application also gives the proper suggestion to the farmer for smart agricultural farming. Finally, in the future, it will help the meteorological department to collect accurate weather information from different locations.

8. ACKNOWLEDGMENTS

This paper is about a microclimate-based crop recommender system for precision agriculture. In this research work, the main target is to figure out predicting crop production using microclimate data. First, we thank Allah for giving us the chance, courage, and strength to work on such a major topic.

We also wish to express our sincere appreciation to all the professors and lecturers of our university who helped us a lot. They convincingly guided and encouraged us to be professional and do the right thing even when the road got tough. Without their persistent help, the goal of this thesis would not have been realized.

Additionally, we wish to acknowledge the support and great love of our family. They kept us going on and this work would not have been possible without their input. Most significantly, this research would not have been possible without the help and expertise of other researchers who have previously worked on this subject.

9. REFERENCES

- [1] Somvanshi, Madan, and Pranjali Chavan. "A review of machine learning techniques using decision tree and support vector machine." In 2016 International Conference on Computing Communication Control and automation (ICCUBE), pp. 1-7. IEEE, 2016.
- [2] Padhy, Neelamadhab, and Rasmita Panigrahi. "Multi relational data mining approaches: A data mining technique." arXiv preprint arXiv:1211.3871 (2012).
- [3] Kushwaha, Ashwani Kumar, and Sweta Bhattacharya. "Crop yield prediction using Agro Algorithm in Hadoop." International Journal of Computer Science and Information Technology & Security (IJCSITS) 5, no. 2 (2015): 271-274.
- [4] Dahikar, Snehal S., and Sandeep V. Rode. "Agricultural crop yield prediction using artificial neural network approach." International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering 2, no. 1 (2014): 683-686.
- [5] Ramesh, D., and B. Vishnu Vardhan. "Analysis of crop yield prediction using data mining techniques." International Journal of research in engineering and technology 4, no. 1 (2015): 47-473.
- [6] Shakoor, Md Tahmid, et al. "Agricultural production output prediction using supervised machine learning techniques." 2017 1st International Conference on Next Generation Computing Applications (NextComp). IEEE, 2017.
- [7] Ahamed, AT M. Shakil, Navid Tanzeem Mahmood, Nazmul Hossain, Mohammad Tanzir Kabir, Kallal Das, Faridur Rahman, and Rashedur M. Rahman. "Applying data mining techniques to predict annual yield of major crops and recommend planting different crops in different districts in Bangladesh." In 2015 IEEE/ACIS 16th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), pp. 1-6. IEEE, 2015.
- [8] Veenadhari, S., Bharat Misra, and C. D. Singh. "Machine learning approach for forecasting crop yield based on climatic parameters." In 2014 International Conference on Computer Communication and Informatics, pp. 1-5. IEEE, 2014.
- [9] Pantazi XE, Moshou D, Alexandridis T, Whetton RL, Mouazen AM. Wheat yield prediction using machine learning and advanced sensing techniques. Computers and Electronics in Agriculture. 2016 Feb 1; 121:57-65.
- [10] Kumar, Rakesh, M. P. Singh, Prabhat Kumar, and J. P. Singh. "Crop Selection Method to maximize crop yield rate using machine learning technique." In 2015 international conference on smart technologies and management for computing, communication, controls, energy and materials (ICSTM), pp. 138-145. IEEE, 2015.
- [11] Chlingaryan, Anna, Salah Sukkarieh, and Brett Whelan. "Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review." Computers and electronics in agriculture 151 (2018): 61-69.
- [12] Paswan, Raju Prasad, and Shahin Ara Begum. "Regression and neural networks models for prediction of crop production 1." (2013)
- [13] Gandhi, Niketa, Owaiz Petkar, and Leisa J. Armstrong. "Rice crop yield prediction using artificial neural networks." In 2016 IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR), pp. 105-110. IEEE, 2016.
- [14] Kamilaris, Andreas, et al. "Agri-IoT: A semantic framework for Internet of Things-enabled smart farming applications." 2016 IEEE 3rd World Forum on Internet of Things (WF-IoT). IEEE, 2016.
- [15] Lee, Meonghun, Jeonghwan Hwang, and Hyun Yoo. "Agricultural production system based on IoT." 2013 IEEE 16th international conference on computational science and engineering. IEEE, 2013.
- [16] Khattab, Ahmed, Ahmed Abdelgawad, and Kumar Yelmarthi. "Design and implementation of a cloud-based IoT scheme for precision agriculture." 2016 28th International Conference on Microelectronics (ICM). IEEE, 2016
- [17] Ryu, Minwoo, et al. "Design and implementation of a connected farm for smart farming system." 2015 IEEE SENSORS. IEEE, 2015.
- [18] Rodríguez, Schubert, Tatiana Gualotuña, and Carlos Grilo. "A system for the monitoring and predicting of data in precision agriculture in a rose greenhouse based on wireless sensor networks." *Procedia computer science* 121 (2017): 306-313.
- [19] Koshy, Santosh Sam, et al. "Application of the internet of things (IoT) for smart farming: a case study on groundnut and castor pest and disease forewarning." *CSI Transactions on ICT* 6.3-4 (2018): 311-318.
- [20] White, Jeffrey W., et al. "Methodologies for simulating impacts of climate change on crop production." *Field Crops Research* 124.3 (2011): 357-368.
- [21] Mearns, Linda O., Cynthia Rosenzweig, and Richard Goldberg. "Effect of changes in interannual climatic variability on CERES-Wheat yields: sensitivity and 2x CO2 general circulation model studies." *Agricultural and forest meteorology* 62.3-4 (1992): 159-189.
- [22] Tubiello, Francesco N., et al. "Effects of climate change and elevated CO2 on cropping systems: model predictions at two Italian locations." *European Journal of Agronomy* 13.2-3 (2000): 179-189.
- [23] Abdullah, Ahsan, et al. "The Case for an Agri Data Warehouse: Enabling Analytical Exploration of Integrated Agricultural Data." *Databases and Applications*. 2004.
- [24] Abdullah, Ahsan, et al. "Learning dynamics of pesticide abuse through data mining." *Proceedings of the second workshop on Australasian information security, Data Mining and Web Intelligence, and Software Internationalisation-Volume 32*. Australian Computer Society, Inc., 2004.

- [25] Veenadhari, S., Bharat Mishra, and C. D. Singh. "Soybean productivity modelling using decision tree algorithms." *International Journal of Computer Applications* 27.7 (2011): 11-15.
- [26] Mestre-Sanchís, Fernando, and María Luisa Feijóo-Bello. "Climate change and its marginalizing effect on agriculture." *Ecological economics* 68.3 (2009): 896-904.
- [27] Tanny, Josef. "Microclimate and evapotranspiration of crops covered by agricultural screens: A review." *Biosystems Engineering* 114.1 (2013): 26-43.
- [28] Zellweger, Florian, Pieter De Frenne, Jonathan Lenoir, Duccio Rocchini, and David Coomes. "Advances in microclimate ecology arising from remote sensing." *Trends in Ecology & Evolution* 34, no. 4 (2019): 327-341.
- [29] Ahmed, Selena, and John Richard Stepp. "Beyond yields: Climate effects on specialty crop quality and agroecological management." *Elementa: Science of the Anthropocene* 4 (2016): 92.
- [30] "Sustainable Development Goals | UNDP". 2021. *UNDP*. <https://www.undp.org/content/undp/en/home/sustainable-development-goals.html>.