Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN : 2581 - 611X)

Volume7-Issue1, January 2024
Paper: 25

AGRICULTURE YIELD PRODUCTION USING MACHINE LEARNING

Beschi Joshua Emmanuel[1]
Student,
Department of Computer Science and Engineering,
Sathyabama Institute of Science and Technology, Chennai, India
beschijoshua@gmail.com

Dr. Ramya. G. Franklin[2]
Assistant professor
Department of Computer Science and Engineering,
Sathyabama Institute of Science and Technology, Chennai, India
ramyagfranklin.cse@sathyabama.ac.in

Dr. L. Suji Helen[3]
Associate professor
Department of Computer Science and Engineering,
Sathyabama Institute of Science and Technology, Chennai, India
sujihelen.cse@sathyabama.ac.in

***Abstract*— The lack of awareness of the various IT applications is one of the reasons for the decline of agriculture sector. They are not able to predict the climate and so they are not able to choose the type of crop for that yield. The proposed system uses machine learning techniques like CNN, decision-tree, and random-forest to estimate the yield of crop by taking in account of major elements like seasons, rainfall, temperature, and locations, we can overcome this challenge.**

Keywords: machine learning techniques, CNN, decision-tree, random-forest, crop yield.

1. Introduction

Agriculture has been viewed as one of India's primary economic activities and considered to be one of the main practices in the country. Around 60% of the land in our country is utilized for agriculture and to cater to 1.2 billion people's needs. Therefore, agriculture is an essential occupation and is significant to the country's prosperity. As a result, crop yield prediction is considered to be an integral part of agriculture it is very important to predict the yield of crops such that the farmer can sow the right seed which is suitable for the land owned by them. Each crops yield is influenced by its different factors such as temperature, soil type and characteristic, humidity, temperature, moisture and so on. The agricultural sector, which is regarded as India's backbone, has grown stronger as a result of increased demands and advancements in technology. Since it is one of the primary sectors to be touched by numerous sources including climatic changes, soil qualities, seasonal variations, etc these innovations are critically needed to satisfy the needs of every individual considering the rapid population growth, therefore yield forecasting is essential since it allows farmers to grow the appropriate crop on their land.

The prediction of crop production is based on a variety of data types that have been gathered and processed using data mining techniques like machine learning techniques from numerous sources that are beneficial for the growth of the crop. It is an art to estimate crop and yield amounts in advance even before actual harvest occurs. Which helps the farmer to be aware of the seeds to be sown beforehand to reduce the inefficient crop production Basically, it is about being aware of what kind of crop can be grown in the area where the land is located and it's dependency factors. It is also useful to cut down on wasteful crop production. For farmers, being able to forecast crop yield could be extremely helpful.

They can plan the crop to be sown before harvest if they know roughly how much production they can expect, which frequently results in a more reasonable rate than if they waited until after harvest. However, due to abnormal climatic changes, food production and forecasting are currently declining which will have a negative impact on farmers' economics which will result in a low yield and also makes it less predictable. Farmers can boost productivity by early detection and managing issues related to crop production indicators to get more yield. But by utilising the proposed model and decision-making tools for crop production prediction, such problems can be resolved. Likewise, industries will get benefits from yield predictions by better planning the logistics of their business. It will be incredibly beneficial for the farmers and other people involved in agricultural industry to make smart decisions about storage and business activities. In this project we consider three parameters visualization of soil type, soil attributes, and the area where the land is located. Based on these parameters data is collected and suggestions are given to the farmers about suitable crops to be cultivated on that geo-location. By monitoring the agricultural area based on the soil's characteristics and advising farmers on the best crop to grow, the proposed project would help farmers significantly increase productivity and minimize loss. This would be done by

Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN : 2581 - 611X)
Volume7-Issue1, January 2024
Paper: 25

monitoring the agricultural area based on the soil's characteristics. Machine learning algorithms are used to predict using historical data of the dependent parameters in that specific region.

Crop Yield Prediction aids farmers in raising crops that produce higher yields in accordance with the algorithmic prediction and it helps to reduce inefficient production for the farmers and results in giving optimal production with available resources, and reduces the capital spent on crops which are not suitable to grow in that particular area. In our project, we have developed a recommendation engine that utilizes various ML algorithms to suggest best suited crop depending on the input characteristics of soil. Consequently, this approach reduces the financial losses that a farmer incurs as a result of cultivating the unsuitable crop. Additionally, helps the farmer to find newer varieties of crop that could be cultivated in that location. This approach provides solutions such as recommending a recommendation system through an ensemble model with majority voting methods using Random tree, Convolutional Neural Network, and Decision Tree for the learner to recommend a suitable crop based on parameters of soil with high efficiency and accuracy.

2. Literature Review

A study of crop yield prediction using ML was done by Thomas van Klompenburg, Ayalew Kassahun, Cagatay Catal. ML is a key tool for forecasting agricultural yields, enabling recommendations about which crop to cultivate and what actions to take while they are in the growing seasons. In this study, they have done a Systematic Literature Review [SLR] to synthesize and extract the algorithm and features implemented in crop yield prediction study. Temperature, rainfall, and soil type were found to be the most frequently utilized features and artificial neural networks to be the most widely employed method in these models, per their research.

E.Manjula, S.Djodiltachoumy in their research they have given They have employed various data mining techniques and tested them in agriculture to forecast the production of crops for the upcoming year. The prediction of crop yield using data mining is quickly examined based on association rules for the selected region. Which was focused on building recommendation engine it puts into practice a system to forecast agricultural yield based on past information about the land. They have achieved it by applying association rule mining on agriculture data.

Alexandros Oikonomidis, Cagatay Catal and Ayalew Kassahun, have done a research on crop yield prediction. Crop yield forecasting is difficult because it depends on numerous parameters even though multiple models have been created the performance of these models are not accurate. In this study they have forecasted the crops yield using various ML algorithms such as XGBoost, Convolution Neural Network, Deep Neural Network, CNN-Long Short-Term Memory and CNN-Recurrent Neural Network. On a public soybean dataset with 25,345 samples and 395 characteristics, such as weather and soil conditions, they have conducted experiments for case study and the results are more accurate than the previous models.

M.K.Dharani, R.Thamilselvan, P.Natesan, PCD.Kalaivaani and S.Santhoshkumar. They have conducted a study utilizing deep learning methods to forecast agricultural productivity. Each country's agriculture sector has a significant impact on its gross domestic product. How effectively the crop was managed or farmed depended totally on the final yield and the market price. On the time of delivery, the entire crop yield determinant depends, hence it needs monitoring and guidance of Artificial Intelligence. This offers a method to automatically predict the production and keep an eye on the crop. Artificial neural networks, deep neural networks, and recurrent neural networks are some of the hybrid deep learning techniques that have been researched. It helped to pinpoint the ways in which agricultural productivity is improved by artificial intelligence technologies. The research study explains the concept and necessity of recurrent neural networks.

Nidhi H Kulkarni, Dr. G N Srinivasan, Dr. B M Sagar, Dr.N. K Cauvery. They have conducted research about how to optimize Crop Productivity using a crop recommendation system built with Ensembling Techniques This proposed system is used to suggest the appropriate crop based on the unique characteristics of the soil and parameters, such as the annual average rainfall and surface temperature with high accuracy. This system's functionality supported a number of machine learning techniques, including Linear SVM, Random Forest, and Naive Bayes. The input soil dataset was classified by this crop recommendation algorithm into the recommended crop types, Kharif and Rabi. Applying the suggested approach produced a 99.91% accuracy rate.

Potnuru Sai Nishant, Pinapa Sai Venkat, Bollu Lakshmi Avinash, B.Jabber. The yield of basically all crops grown in India is predicted in this research. This study uses uncomplicated characteristics like State, district, area,

Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN : 2581 - 611X)
Volume7-Issue1, January 2024
Paper: 25

and season. By tweaking these parameters, a user can forecast the crop's production for the year they desire it to grow. For forecasting the yield they have utilized advanced regression techniques such as Kernel Ridge, Lasso, and E-Net algorithms. It makes use of the idea of stacking regression to improve the algorithms and provide more accurate predictions.

N. Saranya and A. Mythili. In this research they are predicting crop yield based on classification of soil type they use the features and characteristics of the soil for predicting the crop that is best suited for that soil. They have laid out a system for identifying soil types based on the macro- and micronutrient composition, that further forecasts the types of crops that can be grown in each type of soil. They have employed a variety of machine learning algorithms, that includes the likes of logistic regression, support vector machines, bagged trees, and K-Nearest Neighbor (k-NN). It offers crop yield forecast data that are optimal.

3. Proposed Solution

Our proposed system is a web application which has an user-friendly UI that allows the farmer to give their inputs data such as soil type, soil characteristics, geolocation of the land and the amount of rainfall the specified region gets, PH level of soil, potassium level, Nitrogen level and phosphorus level of the soil using which it predicts the crop to be sow in the specified soil type and characteristics and calculates the corresponding yield it will give for the land having that particular soil type. The crop that will be sown depends on a number of factors, which includes rainfall, temperature, wind speed, humidity, etc. These parameters play a vital role in forecasting the crop's yield and it also becomes difficult as these parameters vary place-to- place. To solve this a weather API is developed to gather weather reports the yield depends on the area and the production capacity of that crop. CNN, decision trees and random forests will be used to make predictions in this project. It will deliver the most accurate figures for the crop prediction than the previous models. The application will use the government database for analyses which will use previous harvests data of that particular region and it will give the result through a message that which crop is suitable to grow in that region and its yield.

This project also has fertilizer recommendation section which will get different metrics, such as the nitrogen, PH, and potassium values of the soil. and the crop that the farmer wants to sow based on the information fed the recommendation engine predicts the amount of fertilizer to be used for that particular soil and crop so that it can cure that disease and make the crop yield more optimistic way. It will also have a plant- disease prediction section where it will get the image from the farmer and processes that image using Convolution Neural Network and predicts the disease the plant has and displays the remedy for curing the disease through a message. These sections will help the farmer to make more profit and it will increase their productivity which in turn will attract more people to start farming. Since our nation caters 1.50 billion of people agriculture serves major role in the country's development. Through this project farmers can know more about their land and they will be aware of what types of crops can be sown for their region since these play an essential role in crop prediction. Farmers can choose which crop is best suited for cultivation in that field with the help of the provided methodology.

This study is being carried out for understanding the characteristics of the crops so that can be utilized to efficiently and effectively harvest crops. Farmers would benefit from the precise forecasting of several specified crops throughout various districts. The yield rate of crop production increases as a result, which benefits the Indian economy. The proposed methodology reduces the financial risks involved in framing and helps the framer to cultivated crops that are suited for their soil enabling it to yield more and helps to utilize the land of the farmer to optimum level.

4. System Architecture

The data is collected using datasets in Kaggle, then the raw data is processed into trained data using the help of ML algorithms (Decision Tree, CNN, Random-forest). When the user enters their data system it evaluates their data entered by the user with the trained data sets using which it predicts the crop that can be cultivated for that specified soil and area. The sampling frame for gathering the data for crop yield estimation will be the sample chosen for area estimation. A brief exploration of the sampling strategies utilized to pick the sample for crop area and yield estimation is made.

A double sampling regression estimator is used to integrate the subjective and objective techniques. Based on the domain estimation method, crop area and yield estimation utilizing a double sampling regression estimator are

Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN : 2581 - 611X)
Volume7-Issue1, January 2024
Paper: 25

also seen. The theory of domain estimation enables the distinct estimation of crop area and yield of various crops and blends them into a single sample. The criteria for choosing the sample size are also presented.

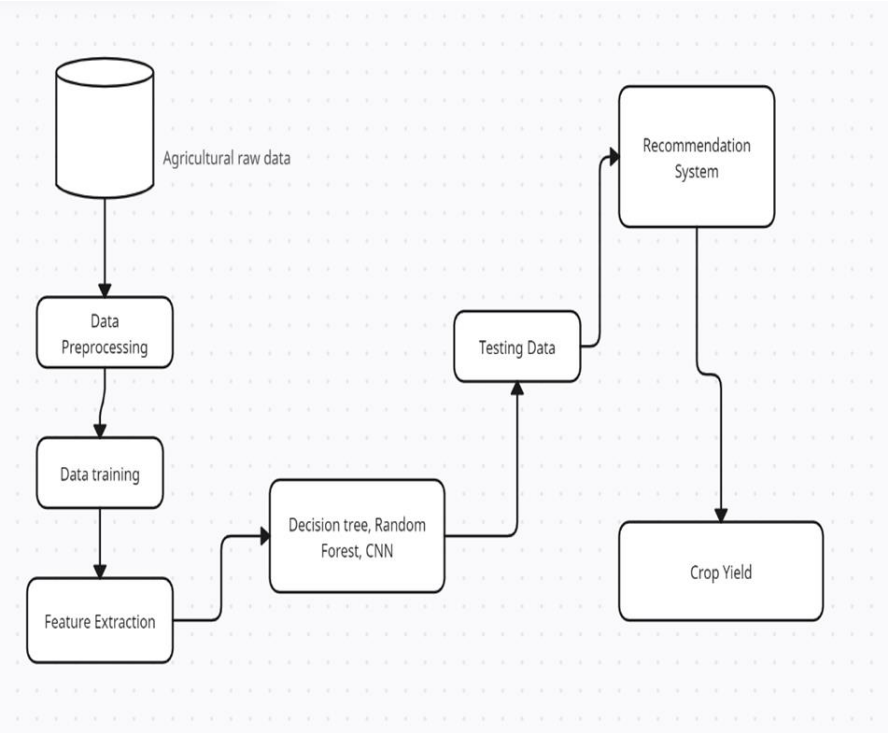

Fig. 1. System Architecture

5. Modules

5.1 Algorithms Used

Random forest
CNN
Decision tree

Are three main algorithms used for forecasting the crop yield that takes various parameters like soil type, characteristics, PH level, nitrogen level, potassium level, phosphorus level of the soil and in addition the amount of rainfall in the specified region.

5.2 Random Forest

Random forest is an ensemble method used for classification. In random forest we create multiple decision trees and compare it with one another and select a single tree in decision tree model. It is a predictive model which adds up multiple decision trees in the random which in turn gets rid of overfitting elements and produces a very strong predictive model. Each tree offers a classification for categorizing a new item based on qualities, and we can say the tree votes for that class. This method classifies based on most votes a tree has. Random Forest will be implemented in both the fertilizer section and crop recommendation section. Which helps the farmers to be aware of the amount of fertilizer to be used so that it will not hinder the crops yield and helps the crop to yield more optimistic way. Thus, helps to utilize the optimum of the land.

5.3 Convolution Neural Network (CNN)

Convolution neural network is one of the simplest ML algorithms based on Supervised Learning techniques. CNN is used for the disease prediction section of the web page. CNN is regarded as the best algorithm for image processing. A set of filters are applied to the input image, each of which highlights different aspects of the image. It completes the task and sends the output to the subsequent layer's filter. The technique is ultimately repeated for numerous iterations of layers as each layer learns to recognize the various fears. After processing the image data, the outcome is presented. CNN algorithm will be implemented in the plant disease section of the project, which

Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN : 2581 - 611X)
Volume7-Issue1, January 2024
Paper: 25

gets the image as input from user and predicts the particular plant disease it has and provides the cure for the disease which will be help the farmer to be aware of their crops remedy for the disease.

5.4 Decision Tree

Decision tree is well-known and effective tools for prediction and classification. Decision trees are used in knowledge systems like databases to describe rules that can be comprehended by humans. It is a hierarchical supervised learning model. Which identifies the solution through a series of iterative method of splitting in smaller stages. A decision contains a leaf terminal and several internal decision nodes. Discrete outcome is produced by each decision node as it implements a test function with branch. A tree is applied with an input at each node, and one of the trees starts at root level and this process is repeated iteratively till it reaches a leaf node, at that time the value which is entered in the leaf node acts as the output. Decision tree will be implemented in the crop prediction section, which gets the input parameters from the user and predicts their yield accordingly. which will help the farmers to be aware of crop which will be best suited for their land and will help to utilize the land in the most optimal level.

6. Conclusion

The proposed agricultural recommendation system is a promising solution that addresses various agricultural challenges by leveraging machine learning algorithms and historical data of dependent factors to predict the most suitable crops for farmers based on their input soil parameters. Through the use of advanced technologies such as Python, numpy, pandas, torch, flask, angular, and bootstrap, the system has been developed with a robust and user-friendly frontend that facilitates easy interaction and usage.

The successful completion of this project holds immense potential to significantly impact the agricultural sector in multiple ways. By helping farmers choose the right crops to cultivate based on accurate predictions, the system can contribute to increased productivity and reduced losses. This, in turn, can lead to improved economic outcomes for farmers and contribute to the growth of the national economy.

Moreover, the system's ability to recommend new types of crops that can be cultivated in a particular area can open up new possibilities for farmers and attract more people to start farming. This can foster innovation and diversification in agricultural practices, leading to a more sustainable and resilient agricultural sector.

Furthermore, by incorporating real-time data, such as weather conditions, soil moisture levels, and market prices, and by integrating remote sensing and IoT technologies, the system can continuously update its predictions and provide farmers with more accurate and up-to-date recommendations. This can further enhance the system's effectiveness in supporting farmers' decision-making process and optimizing their farming practices.

Overall, the successful completion of this agricultural recommendation system project has the potential to revolutionize the way farmers make crop selection decisions, reduce financial losses caused by planting the wrong crops, and enable the adoption of more sustainable and profitable agricultural practices. With its machine learning-based approach, intuitive frontend, and potential for future enhancements, this project can have a positive and lasting impact on the agricultural sector, benefiting farmers, communities, and the national economy as a whole.

References

[1]. Satish Babu (2015) ‘A Software Model for Precision Agriculture for Small and Marginal Farmers’.

[2]. Mrs. N. Saranya, Ms. A. Mythili (2020)’Classification of Soil and Crop Suggestion using Machine Learning Techniques’-International Journal of Engineering Research & Technology (IJERT).

[3]. Franklin, Ramya G. ; Muthukumar, B. A Prediction System for Diagnosis of Diabetes Mellitus; Journal of Computational and Theoretical Nanoscience, Volume 17, Number 1, January 2020.

[4]. Dr. Y. Jeevan Kumar, (2020)’ Supervised Learning Approach for Crop Production’,2020.

[5]. Aruvansh Nigam, Saksham Garg, Archit Agrawal (2019) ‘Crop Yield Prediction using ML Algorithms’.

[6]. Tanmay Banavlikar, Aqsa Mahir, Mayuresh Budukh, Soham Dhodapkar (2018) ‘Crop Recommendation System Using Neural Network’ - International Research Journal of Engineering and Technology (IRJET).

[7]. Mishra S., Mishra, D. Santra, G. H, (2020)' Applications of machine learning techniques in agricultural crop production'.

[8]. Ranjini B Guruprasad, Kumar Saurav, Sukanya Randhawa (2019) 'Machine Learning Methodologies for Paddy Yield Estimation in India: A CASE STUDY'.

[9]. Potnuru Sai Nishant, Pinapa Sai Venkat, Bollu Lakshmi Avinash, B.Jabber "Crop Yield Prediction Based On Indian Agriculture Using Machine Learning ",2020.

[10]. M.K.Dharani,R.Thamilselvan,P.Natesan,PCD.Kalaivaani and S.Santhoshkumar "Review on Crop Prediction Using Deep Learning Techniques " ,2021.

[11]. B. Rajesh, M. V. Sai Vardhan and L. Sujihelen, "Leaf Disease Detection and Classification by Decision Tree," 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184), Tirunelveli, India, 2020.