Agriculture Prediction Using ML

Ronak Chavda, Borad Zenish, Aniket Parmar and Nihal Patel

March 18, 2024

AGRICULTURE PREDICTION USING ML: A Survey-Based Study

1st CHAVDA RONAK
Department of Computer Science and Engineering
Parul University
Vadodara, India
200303105103@paruluniversity.ac.in

2nd BORAD ZENISH
Department of Computer Science and Engineering
Parul University
Vadodara, India
200303105101@paruluniversity.ac.in

3rd PARMAR ANIKET
Department of Computer Science and Engineering
Parul University
Vadodara, India
200303105041@paruluniversity.ac.in

4th PATEL NIHAL
Department of Computer Science and Engineering
Parul University
Vadodara, India
200303105049@paruluniversity.ac.in

Abstract—Agriculture is one of the most important and widely practised professions in India, and it has contributed significantly to the growth of our nation. Around 60used for agriculture. Growing crop output is seen as a crucial component of agriculture since it helps meet the demands of 1.2 billion people. Using machine learning to solve practical and real-world crop productivity problems can be challenging. Usually, if we own a piece of land, we need to have a basic knowledge of the kinds of crops that ought to be grown there. Many aspects of the soil must be present for agriculture to thrive. Crop production is a challenging undertaking because it requires taking into account a number of factors, including temperature, soil type, humidity, and others. Making educated decisions about storage and business will be significantly simpler for farmers and other stakeholders if it is simple to locate the crop to be grown before seeding it. By monitoring agricultural regions based on soil qualities and counselling farmers on the best crop, the proposed project will assist in the settlement of agricultural difficulties by advising them on how to significantly increase production and decrease loss. According to the description, this study is a recommendation system that uses a number of machine learning techniques to suggest suitable crops based on input soil factors. The purpose of this study is to forecast agricultural production using a variety of machine learning techniques. Data from Logistic Regression, Nave Bayes, and Random Forest are combined in the classifier models inferred here, with Random Forest displaying the highest accuracy. Forecasts based on machine learning algorithms that take into account all of these variables, including rainfall, temperature, and region, may one day assist farmers in deciding what plant to produce. So, this strategy lessens the f inancial losses that farmers experience when they decide to plant the improper crops. It also helps farmers in their search for new crop varieties that are suitable for cultivation in their region.

I. Introduction

Since its inception, agriculture has been the main activity in every society and civilization that has existed throughout human history. It is not only a huge part of the expanding economy, but it is also necessary for our survival. It is also a vital sector for the Indian economy and the future of humanity. Also, it makes up a sizable amount of employment. As time goes on, the demand for production has dramatically expanded. People use technology in an utterly incorrect manner in order to produce in large quantities. Every day, new hybrid kinds are created. These kinds, however, don't offer the same critical components as a crop grown naturally. These artificial methods degrade the soil. All of this causes more environmental deterioration. Most of these unconventional methods are used to prevent losses. Yet, the loss is reduced when agricultural growers have access to reliable crop production information. Machine learning is a rapidly expanding methodology that supports decision-making across all industries to provide the most useful of its applications. The majority of modern gadgets benefit from models being examined before deployment. The fundamental idea is to use machine learning models to boost the agricultural sector's throughput. The amount of knowledge imparted during the training period is another element that influences the prediction, as the number of parameters was higher in comparison. Precision agriculture, which prioritises quality over unfavourable environmental variables, would be the main focus. Many machine learning classifiers, including Logistic Regression, Naive Bayes, Random Forest, etc., are used to urge a pattern in order to conduct reliable prediction and stand on the erratic patterns in temperature and rainfall. Our analysis of the aforementioned machine learning classifiers led us to the conclusion that the Random Forest method offers the highest level of accuracy. The system forecasts crops based on the collection of historical data. The information is provided using historical data on the weather, temperature, and a number of other variables. Our application runs an algorithm and displays a list of crops that match the inputted data and their anticipated yield values.

II. Literature review

A. Historical Perspective of Agriculture Prediction

Agriculture price prediction has been historically significant for several reasons, primarily driven by the need to manage risks, make informed decisions, and ensure stability in the agricultural sector.hostorical purposes are risk management,supply chain planning,finanncial planning,market stability,policy formulation,food security,research and development.

B. Ensemble Learning Methods

Ensemble learning methods in agriculture price prediction involve combining predictions from multiple models to improve overall accuracy and robustness. These methods leverage the diversity among individual models to overcome the limitations of any single model.some methods are bagging,random forest,boosting ,stacking,voting classifier,weighted averaging.Ensemble learning methods are valuable in agriculture price prediction as they help mitigate overfitting, improve generalization, and enhance the overall predictive performance of the model. The choice of ensemble method depends on the characteristics of the data and the specific challenges of the agriculture price prediction task.

C. Evolution of Methodologies and Techniques

In the past, agriculture predictions relied heavily on traditional methods such as expert judgment, historical averages, and basic statistical models. As data collection and computational capabilities improved, statistical models like linear regression and time series analysis became more prevalent. The integration of remote sensing technologies and Geographic Information System (GIS) in agriculture brought a significant advancement. Satellite imagery, aerial surveys, and GIS mapping enabled the monitoring of crop health, land use changes, and environmental factors, contributing to more informed predictions.With the rise of machine learning techniques, agriculture prediction moved towards more data-driven and algorithmic approaches. Ensemble learning methods, such as bagging, boosting, and stacking, evolved to address the limitations of individual models. The advent of big data technologies and cloud computing provided the infrastructure to handle large-scale agricultural datasets. This allowed for more comprehensive analysis and modeling, accommodating the increasing volume of data generated by sensors, satellites, and other sources.IoT devices, such as sensors and actuators, became integral to precision agriculture. Deep learning, a subset of machine learning, gained prominence with the use of neural networks for agriculture prediction tasks.Blockchain technology has been explored to enhance transparency and traceability in the agricultural supply chain.As AI models become more complex, there is a growing emphasis on making predictions interpretable. Explainable AI techniques aim to provide insights into how models arrive at specific predictions, which is crucial for gaining trust in the agricultural community and facilitating informed decision-making.

D. Key Challenges in Agriculture Prediction

One of the key challenges in agriculture prediction is the inherent complexity of the agricultural system, which involves numerous interrelated factors. Here are some key challenges:Data variability and quality, Multifactorial Nature,Climate Change and Extreme Events, Limited Historical Data, Rapid Technological Advancements, Model Interpretability. .

III. Methodology

The Methodology represents a description about the framework that is undertaken. It comprises various milestones that should be accomplished in order to satisfy the objective. We have undertaken different data mining and machine learning concepts. The accompanying figure, Fig.1 represents step-wise undertakings that should be finished.

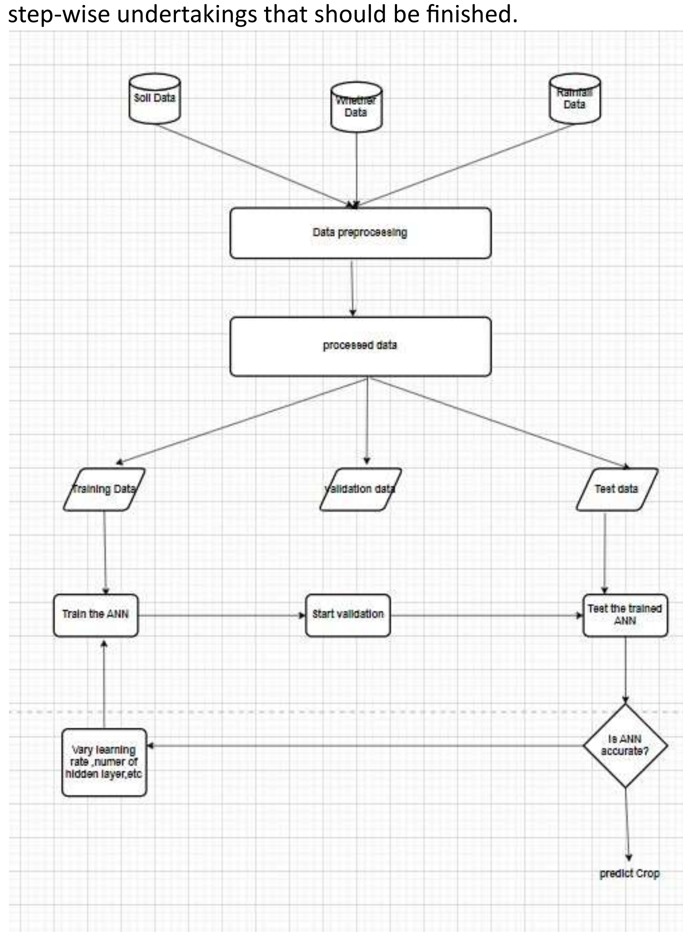

Fig. 1. Methodology

A. Data Pre-processing

A technique called data preprocessing is used to turn the raw data into a clean data set. The data are acquired from various sources, however because they are collected in raw form, analysis is not possible. We can convert data into a comprehensible format by using several strategies, such as substituting missing values and null values. The division of training and testing data is the last step in the data preprocessing process. Due to the fact that training the model typically requires as many data points as possible, the data typically tend to be distributed unevenly. The training dataset is the first dataset used to hone machine learning algorithms and make accurate predictions.

B. Factors affecting agriculture prediction

The yield and productivity of any crop are impacted by a wide range of variables. These are essentially the characteristics that aid in estimating a crop's annual yield. We take into account variables like temperature, rainfall, area, humidity, and wind speed in this essay.

C. Comparison and selection of machine learning algorithm

We must first assess and compare potential algorithms before selecting the one that best fits this particular dataset. The best method for solving the crop production problem practically is machine learning. Several machine learning methods are employed to forecast agricultural yield. The following machine learning techniques for selection and accuracy comparison are included in this paper: A. Logistic regression: The probability of a target variable is predicted using the supervised learning classification algorithm known as logistic regression. Because the dependent variable's nature is dichotomous, there are only two viable classes. The accuracy of the logistic regression approach when applied to our dataset is 87.8B. Na¨ıve Bayes: The Naive Bayes classifier makes the assumption that a certain feature's

presence in a class has no bearing on the presence of any other feature. Simple to construct and especially helpful for very big data sets is the naive Bayes model. Even with being straightforward, Naive Bayes is known to perform better than even the most complex classification techniques. It offers a 91.50C. Random forest: Crop development in relation to current climatic conditions and biophysical change can be examined using Random Forest. The random forest algorithm builds decision trees using several data samples, predicts the data from each subset, and then determines which answer is best for the system through user voting. The bagging approach is used by Random Forest to train the data, increasing the accuracy of the outcome. RF offers a 92.81accuracy for our data. It is clear that among all three algorithm's ,Random forest is the best algorithm because of greater accuracy

D. Random forest model for crop prediction

In order for each tree to depend on the values of a random subset sampled independently and with the same distribution for all the trees in the forest, random forests are the collection of tree 13 predictors. The bagging method was utilised by Random Forest to train the data, increasing the accuracy of the outcome. The Random Forest approach, which provides accuracy for model-based predictions and the actual results of predictions in the dataset, was utilised to achieve high accuracy. Analyzed is the model's anticipated accuracy, which is 91.34crop yield is shown above.

E. System Architecture

: The weather API is the key component of the system architecture shown in Fig. 3, from which we retrieve data on temperature, humidity, rainfall, and other variables. The server module receives the data that was retrieved from the API. The server's database is where the data is kept. The user can provide information such as location, area, etc. through the mobile application. The user can register once for an account on the mobile app and the server receives all of these entered data. The trained Random forest model installed on the server analyses all the input and fetched data to estimate crop yields and locates the yield of the named crop in the specified location

IV. Challenges and Limitations

A. Data Quality and Availability

- Limited access to high-quality and diverse datasets can hinder the development of accurate models.
- Inconsistent or incomplete data can lead to biased predictions and reduced model performance.

B. Data Variability

- Agricultural data is highly variable due to factors like weather conditions, soil composition, and crop types. This variability can make it challenging to generalize models across different regions and time periods.

C. Complexity of Agricultural Systems

- Agriculture involves complex, interconnected systems with numerous variables influencing crop growth, such as pests, diseases, and environmental factors. Capturing and modeling these complexities accurately can be challenging.

D. Dynamic Nature of Agricultural Processes

- Agricultural systems are dynamic and subject to constant changes. Sudden weather events, disease outbreaks, or changes in farming practices can impact predictions, requiring models to adapt in real-time.

E. Model Interpretability

- Some machine learning models, especially deep learning models, can be complex and lack interpretability.

Understanding how a model makes predictions is crucial for gaining trust and making informed decisions.

F. Adoption and Acceptance

- Farmers may be hesitant to adopt new technologies, particularly if they perceive them as a risk or if the technology does not align with their traditional practices. Convincing users to embrace machine learning-based solutions is a significant challenge.

V. Application and Case Studies

A. World Application of Agriculture Prediction Models

- Agriculture prediction using machine learning has numerous real-world applications that can significantly benefit farmers, increase crop yields, optimize resource use, and contribute to sustainable farming practices. Some notable applications include: Crop yield prediction,Disease detection, fertilizer prediction,etc.
- By harnessing the power of predictive analytics, farmers can make more informed decisions, adapt to changing conditions, and enhance overall farm productivity.

B. Case Studies Highlighting Successful Implementations

- Examining specific case studies provides valuable insights into the practical implications and successes of real estate price prediction models.
- This subsection will delve into noteworthy case studies where predictive models have played a pivotal role.
- Examples may include instances where machine learning algorithms accurately predicted Crop name, leading to profitable investments.

VI. Conclusion and Future Scope

Conclusion

The use of machine learning algorithms to anticipate crops and determine their yield is the main topic of this study. The calculation of accuracy uses a variety of machine learning techniques. The crop prediction for the selected district made use of the Random Forest classifier. created a technique to anticipate crops using data gathered in the past. The suggested method aids farmers in choosing which crop to plant in the field. This work is done to learn more about the crops that can be used to harvest things in an effective and helpful way. Farmers in Kerala will benefit from the precise forecasting of many specific crops across various districts. As a result, the yield rate of crop production is maximised, which benefits our Indian economy.

Future Scope

The use of machine learning algorithms to anticipate crops and determine their yield is the main topic of this study. The calculation of accuracy uses a variety of machine learning techniques. The crop prediction for the selected district made use of the Random Forest classifier. created a technique to anticipate crops using data gathered in the past. The suggested method aids farmers in choosing which crop to plant in the field. This work is done to learn more about the crops that can be used to harvest things in an effective and helpful way. Farmers in Kerala will benefit from the precise forecasting of many specific crops across various districts. As a result, the yield rate of crop production is maximised, which benefits our Indian economy.

References

[1] Thomas van Klompenburga , Ayalew Kassahuna , Cagatay Catalb .Crop yield prediction using machine learning: A systematic literature review,accessed in 2020.

[2] Dr. Pathan Ahmed Khan, Meesum S Hussain, B Mir Mustafa Ali, MD Zakir Ali Khan, CROP YIELD PREDICTION USINGMACHINELEARNINGALGORITHMS,accessedon5th may,2022

[3] A.P.S Manideep , Dr. Seema K , A Comparative Analysis of Machine Learning Prediction Techniques for Crop Yield Prediction in, accessed in 2022.

[4])Nischitha K, DhanushVishwakarma, MahendraN,Ashwini, ManjurajuM.R.CropPrediction using Machine Learning Approaches, accessed on 8th august, 2020.

[5] Dr. Poorna Shankar , Dr. Prashant Pareek , Ms. Urvi Patel , and Mr. Canny Sen, Crops Prediction Based on Environmental Factors Using Machine Learning Algorithm, accessed in 2022.

[6] Anakha Venugopal, Aparna S, Jinsu Mani, Rima Mathew, Prof. Vinu Williams, Crop Yield Prediction using Machine Learning Algorithms, accessed in 2021.

[7] . Ms. Priyanka Bakare Ms. Shamika Ghodke Under the guidance of Prof. Rakesh Borase, Project Report on Data analysis and prediction for agricultural production, accessed on April,2021.

[8] X.E. Pantazi , D. Moshou , T. Alexandridis , R.L. Whetton , A.M. Mouazen, Wheat yield prediction using machine learning and advanced sensing techniques, accessed in 2015.

[9] Marizel B. Villanueva, Ma. Louella M. Salenga, Bitter Melon Crop Yield Prediction using Machine Learning Algorithm , accessed in 2018.

[10] Taman Kumar , Kiran Jyoti , and Sandeep K.Singla,, Design and Development of Machine Learning Model for Crop Yield Prediction.

[11] Prof. Swati Dhabarde, Swapnil Bisane, Arti Yadav , Devyani Pote , Akshay Gupta, Smart Farming Using Machine Learning, accessed on 1st March, 2022

[12] Spyridon Mourtzinis, Paul D. Esker, James E. Specht and Shawn P. Conley, Advancing agricultural research using machine learning algorithms.

[13] by Ishwarya R, Nagapooja , Raghavi R, Soundarya K, Prof. Chitra C, CROP YIELD PREDICTION USING MACHINE LEARNING ALGORITHM, accessed on 7th July, 2022.

[14] Madhuri Shripathi Rao, Crop prediction using machine learning, accesed in 2022.

[15] Mayank Champaneri , Darpan Chachpara , Chaitanya Chandvidkar , Mansing Rathod, Crop Yield Prediction Using Machine Learning, accessed in 2018.

[16] Mrs. R. Usha Devi 1 , Dr. N.A. Sheela Selvakumari, Crop Prediction And MappingUsing Soil Features With Different Machine Learning Techniques.