Seasonal Crop Recommendations for High Productivity in India

K Jamunadevi[1], Dr. K Banuroopa[2]

[1]Department of Information Technology, Dr. N.G.P Arts and Science College, Coimbatore, India

[2]Assistant Professor, Department of Information Technology, Dr. N.G.P Arts and Science College, Coimbatore, India

ARTICLE INFO

ABSTRACT

Tamil Nadu, being a sea-facing state, faces agricultural uncertainty, which diminishes productivity. One can expect higher productivity with a greater number of people and acres of land but cannot do the same. Growers utilize word-of-mouth but are unable to do so with climate conditions. Agricultural factors and terms are utilized. Give information that can be utilized to know more about Agri-facts. Some of the key points in agriculture are motivated by the expansion of the IT world. Sciences to assist farmers with accurate agricultural information. The intelligence of using advanced technological approaches in the agricultural area is optimal in this present situation. Machine Learning techniques create a clear model with the data and enable us to achieve predictions. Agricultural problems such as crop forecasting, rotation, water need, fertilizer need, and protection may be solved. Because of the climatic factors of the environment that vary, a proper method to support crop production, and help farmers improve their production, and management is needed. This would help potential agriculturalists. Better their farming. Using data mining, a farmer can get a system of suggestions to guide them in crop cultivation. Crops are suggested based on climatic factors and how much to use, such as method. Data analytics provides a way to develop useful extraction from agriculture databases. Crop Data set has have been tested, and crop suggestions based on productivity and season have been provided.

Keywords : Machine Learning, Light GBM, Crop and Yield Prediction, Knowledge Discovery in Databases.

I. INTRODUCTION

Tamil Nadu is India's seventh-largest state and has the sixth-highest population. It is the world's largest producer of agricultural products. Tamil Nadu's main source of income is agriculture. In this ruthless world, agriculture has a strong voice. Cauvery is the main water source. The Cauvery Delta is Tamil Nadu's rice bowl. In Tamil Nadu, rice is the most important crop.

Peanuts, cotton, sugarcane, coconut, and paddy are among the other crops that are farmed. Biofertilizers are efficiently produced.

The agriculture sector of a country has a significant impact on its economy. These days, agricultural farming is deteriorating due to changes in natural causes. Agriculture is directly affected by the climate, the highest and lowest temperatures, rainfall, soil type, sunlight, humidity, fertilizers, pesticides, and other elements.

II.LITERATURE REVIEW

For the agriculture industry to be able to forecast crop yields and assist farmers in making informed choices to improve farming quality, there should be a methodical and evident way. Patrizia Busato and Konstantinos G. Liakos, "Machine Learning in Agriculture: A Review," 2019.In Agri technology, which is an interdisciplinary science, machine learning has opened new horizons for data-intensive science owing to advancements in big data technologies and supercomputers. In this paper, we present a comprehensive review of the literature on machine learning applications in agricultural production systems. Palepu, Ramesh Babu (2020)An analysis of agricultural soils by applying data mining techniques. Applying data mining techniques in agriculture, particularly regarding soils, can enhance crop yields and change the pledge- making process. Soil analysis is necessary to solve various problems.A proper and systematic approach is necessary in the agricultural sector for the estimation of crop production and helping farmers make better-informed decisions to enhance quality farming. Konstantinos G. Liakos, Patrizia Busato(2019)-Machine Learning in Agriculture: A Review With advancements in big data technologies and high-performance computing, machine learning has provided new opportunities for data-intensive science in the multi-disciplinary agri-technology field. In this paper, We provide a detailed analysis of research on

machine learning applications in agricultural production systems in this research. Ramesh Babu Palepu(2020)-An Analysis of Agricultural Soils using Data Mining Techniques. The application of data mining methods to agriculture, especially on soils, can increase the yields of cultivation and alter the nature of pledge-making. Soil analysis is instrumental in solving many issues in the agricultural sector. This paper addresses the

application of data mining in the context of soil analysis in agriculture, as well as numerous data mining techniques and their respective work by various authors in the context of soil analysis. A. Swarupa Rani(2021)- The Impact of Data Analytics in Crop Management based on Weather Conditions. Data mining aims to derive knowledge from a given set of data and convert it into a new, distinct human-readable form for use in the future. Crop management in a specific agricultural area is determined by local climatic conditions since climate plays an important role in crop yield. Live weather information can help in proper crop management. The application of information and communications technology enables the automation of extracting meaningful data to gain knowledge and trends.

III.MODULS DISCRIPTION DATA COLLECTION

The process of gathering data is the first step in creating a machine-learning model. This crucial stage will have a cascading effect on the model's performance the more and better data we collect, the better the model will function. Data can be gathered using a variety of methods, including physical interventions, web scraping and more. The information utilized in this Indian crop suggestion was obtained from another source.

DATASET
The dataset contains 821 unique data points. The 14 columns that make up the dataset are described below.

1. States: all of the Indian states combined.

2. Rainfall: precipitation in millimeters.

3. Ground Water: The total groundwater level.

4. Degrees Celsius are used to express the temperature.

5. Soil types: A range of soil types.

6. Season: What is the ideal time of year for crops?

7. Crops: Types of crops.

8. Necessary fertilizers: The kinds of fertilizers that are required.

9. Cultivation cost: The entire expense of cultivation.

10. Expected earnings: The total amount of expected earnings.

11. The quantity of seeds per hectare is known as seeds per hectare.

12. Cultivation duration: the total number of days that the plant is grown.

13. Demand for crops: high, low, or neither.

14. Which crops are appropriate

DATA SETTING

Prepare the data for training by sorting it. Eliminate duplicates, correct errors, handle missing values, normalize, convert data types, and tidy up everything that could require it. To remove the influence of the particular order in which the data were collected and otherwise prepared, they should be randomized. To detect relevant associations between variables, spot class imbalances (watch out for bias), or do more exploratory analysis, employ data visualization. Separate the sets into evaluation and training.

MODEL SELECTION

Decision trees are flowchart-like tree structures in which an internal node represents a feature or attribute, a branch represents a decision rule, and each leaf node represents the conclusion. The node at the top of a decision tree is called the root node. It acquires the capacity to divide based on the attribute's value. The technique for dividing the tree recursively is called recursive partitioning. This flowchart-like framework facilitates decision-making. Similar to a flowchart diagram, this kind of depiction easily mimics human-level cognitive processes. For this reason, decision trees are easy to understand and evaluate.

ANALYZE AND PREDICTION

Only seven features were selected from the actual dataset they are below:

States: The entire count of Indian states. Precipitation: Millimeters.

Groundwater: the total amount of groundwater. Temperature: temperature in degrees Celsius. Soil type: The variety of soils.

Seasons: which ones?

IV.SYSTEM STUDY EXISTING SYSTEM

Tripathy et al. created a system for controlling pesticides used in crop production using data mining. Pritam Bose developed an SNN model to have a spatiotemporal analysis with crop estimation. Shreya S. Bhamose's agricultural and yield prediction model predicts agricultural production and water demand using a modified k-means clustering algorithm.

DISADVANTAGES OF EXISTING SYSTEM

1. Instead of considering all the states and other aspects, the present system just considered one state.

2. Built more slowly in comparison.

3. Interpretation is challenging.

4. Computationally pricey.

PROPOSED SYSTEM

1. Numerous agricultural factors influence crop productivity. The crops that can be suggested to farmers are based on the production of crops in prior years.

2. Farmers might use these suggestions to ascertain whether particular region has historically produced successfully. Water problems, agricultural illnesses, and numerous other factors can lower agricultural output.

3. When farmers are considering production, they can find out which crop is selling well that year. Based on

this, farmers can make decisions regarding recent crop trends.

4. When giving farmers advice, the agricultural output season will be considered.

5. The project's stated issue is recommending crops to farmers.

ADVANTAGES OF PROPOSED SYSTEM

1. While only one state was taken into consideration in the current system, we have used a vast dataset in our proposed method that takes into account all of India's states.

2. These suggestions can be taken and used to educate farmers. A more thorough understanding of the crops to be grown is demonstrated to the farmer by pictorial representation.

3. It doesn't need to be scaled or normalized, is simple to construct, and comprehend, and is computationally less expensive

V METHODOLOGY – ALGORITHMS RANDOM FOREST

Random Forest is a flexible, easy-to-use machine learning method that usually produces good results, even without hyper-parameter adjustment. It is also one of the most often used algorithms due to its adaptability and simplicity of usage (it can be used for both classification and regression problems). In this post, we'll learn how the random forest method works, how it differs from other algorithms, and how to use it.

ENSEMBLE LEARNING

By combining the predictions from several machine learning algorithms, an ensemble method can produce predictions that are more accurate than those from any one model alone. An ensemble model is a model made up of numerous models.
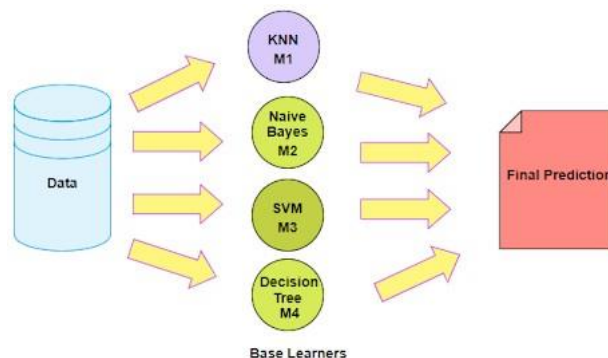

Base Learners

TYPES OF ENSEMBLE LEARNING:

I.BOOSTING

A class of algorithms known as "boosting" uses weighted averages to turn weak learners into stronger ones. The key to boosting is "teamwork." The features that the subsequent model will concentrate on are determined by each model that runs. As the name implies, boosting involves learning from others, which further enhances learning.

BOOTSTRAP AGGREGATION

Random sampling with replacement is referred to as bootstrap. We can have a better understanding of the dataset's bias and variance thanks to Bootstrap. For algorithms with large variance, such as decision trees, this generic process can be used to lower the variance. Each model runs independently thanks to bagging, and the final results are combined without favoring any one model over another.

SYSTEM ARCHITECTURE
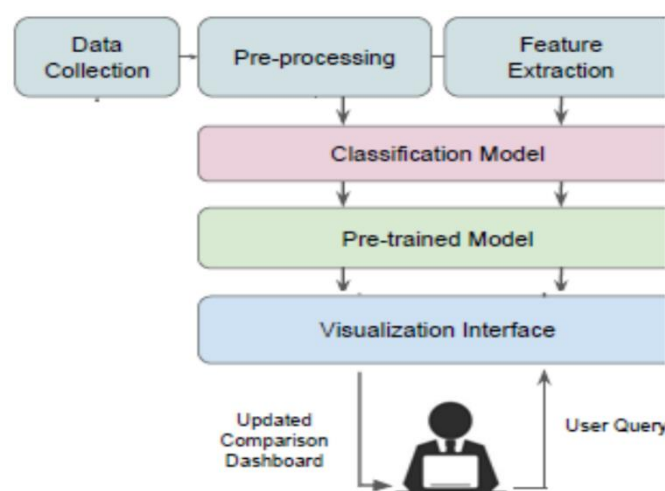
VI. CONCLUSION

A lot of research was done on the importance of crop management. Farmers need help with cutting-edge technology to grow their crops. Farmers can be informed about appropriate harvest forecasts on a timely basis. Any machine learning techniques have been used to examine the agricultural parameters. A literature review looks at a few farming methods used in different fields. Blooming neural networks and other soft computing methods are important for making recommendations. By considering variables like production and season, farmers can get more tailored and relevant advice that motivates them to produce a large amount of food. When comparing different algorithms, the Support Vector Regression, Linear Regression, and Adaptive Neuro Fuzzy Inference System offer the maximum degree of accuracy.
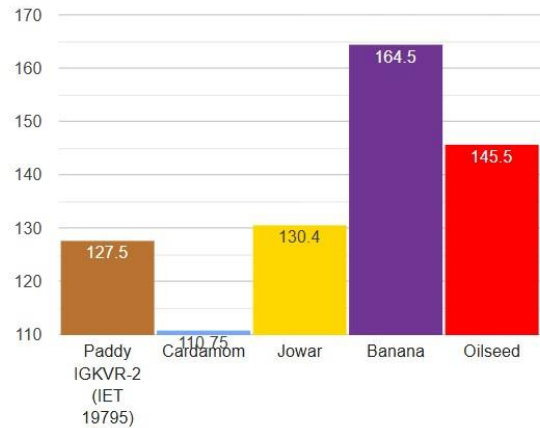
VII. OUTPUT

Figure 1.1

Figure 1.2

References

[1]. Patel, K., & Patel, H. B. (2023). "Multi- criteria Agriculture Recommendation System using Machine Learning for Crop and Fertilizers Prediction."

[2]. Apat, S. K., Mishra, J., Raju, K. S., & Padhy, N. (2023). "An Artificial Intelligence-based Crop Recommendation System using Machine Learning."

[3]. Yerra, S. Enhancing Inventory Management through Real-Time Power BI Dashboards and KPI Tracking. (2025). International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 11(2), 944-951. https://doi.org/10.32628/CSEIT25112458

[4]. Prakash, V., & Tajuddin, A. (2020). "Agricultural Data Mining for Crop Recommendation and Yield Prediction."

[5]. Mohapatra, B. N., & Kale, V. (2024). "Crop Recommendation System using Machine Learning."

[6]. Rakesh Kumar, M. P. Singh, Prabhat Kumar, & J. P. Singh. (2015). "Crop Selection Method

[7]. Geetha, V., Punitha, A., Abarna, M., Akshaya, M., Illakiya, S., & Janani, A. P. (2020). "An Effective Crop Prediction using Random Forest Algorithm."

[8]. Ray, R. K., Das, S. K., & Chakravarty, S. (2022). "Smart Crop Recommender System— A Machine Learning Approach."

[9]. Vigneswaran, E. E., & Selvaganesh, M. (2020). "Decision Support System for Crop Rotation using Machine Learning."

[10]. Modi, D., Sutagundar, A. V., Yalavigi, V., & Aravatagimath, A. (2021). "Crop Recommendation using Machine Learning Algorithm."

[11]. Kushal, B. J., Sp, N. J., Raaju, N. S., Gv, K. G., Kp, A. R., & Gowrishankar, S. (2022). "Real-Time Crop Prediction based on Soil Analysis using Internet of Things and Machine Learning."

[12]. Yerra, S. Optimizing Supply Chain Efficiency Using AI-Driven Predictive Analytics in Logistics. (2025). International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 11(2), 1212-1220. https://doi.org/10.32628/CSEIT25112475