

E-Auction: Recommendation System for Farmers using Machine Learning

Avinash Bhat¹, Ahmed Morve², Mahesh Dhokate³, Ayush Dhar⁴, Prof. Wasudeo Rahane⁵

Students, Information Technology, NBN Sinhgad School of Engineering, Ambegaon BK., Pune^{1,2,3,4}

Guide, Information Technology, NBN Sinhgad School of Engineering, Ambegaon BK., Pune⁵

Abstract: *A major contribution to our country's GDP is agriculture. However, farmers have not yet received the required number of crops. This is largely due to improper irrigation or crop selection or sometimes because yields are lower than expected. The best crops that will have the highest yields and crop residues can be predicted by analyzing the soil and atmosphere in a particular area. This forecast will help growers to choose the right crop based on soil type, temperature, and humidity, water level, depth, soil pH, season, fertilizer and months. And our contribution applies to fertilizer recommendations, sale of goods and agricultural equipment to be rented. And to remove the third person between the farmers and the consumers in order to get a direct profit from the farmers. Program E: Auctions and Recommendations aims to eliminate the problems that appear in the manual auction system. The results obtained in subsequent tests were very impressive in terms of time, as well as benefits for farmers compared to the manual system. Such a system with all of these skills will go a long way in fixing the problems mentioned above with the existing manual voting system at auction.*

Keywords: Agriculture, Crop Soil, E-auction, Prediction, Machine Learning, etc.

I. INTRODUCTION

Data analysis is the process of cleaning and modeling for the purpose of obtaining useful information and conclusions. It is a process of analyzing, extracting and predicting meaningful data from big data to pattern extraction. This process is used by companies to get useful information about the raw data of their customers. In the agricultural sector this analysis can also be used. Many farmers rely on their long experience in the field for certain crops in anticipation of high yields in the next harvest. Mainly due to improper watering or crop selection, or crop yields are sometimes lower than expected.

Farmers emphasize the need for an effective method of predicting and improving crop growth and most agricultural research focuses on biological methods of identifying and improving plant growth. The effect of crop yields depends largely on factors such as crop diversity, seed type and environmental conditions such as sunlight (temperature), soil (ph), water (ph), crop abundance and moisture. The best crops that will have the highest yields and crop residues can be predicted by analyzing the soil and atmosphere in a particular area. Farmers will benefit from that forecast.

To find the right soil type, temperature, water level, depth, soil pH, season, fertilizer and months, choose the right crops for your farm.

Agriculture is India's main activity. About 70 percent of primary and secondary agricultural companies. So many farmers have started using new technologies and techniques to improve farming. However, people do not know the right time and place to plant crops. The concept of determining crop suitability and yield based on the various factors that affect production in this regard can improve crop quality and yield, thereby increasing economic growth and gaining employment.

II. RELATED WORK

The C4,5 algorithm uses a logbook based on data entropy and the results are illustrated. a greenhouse data analysis tool for vegetables using data mining technology [1].

To predict agricultural production, use hybrid supportive vector regression. To predict [2], hybrid SVR models are used. Data mining method: estimates of crop yield survey in a few climate zones in Karnataka. The multi-level Regression and Random Forestry algorithm for predicting crop yields is categorized using drainage mind [3].

Agricultural production forecasting using the Machine Learning Techniques “Decision Tree Learning-ID3 (Iterative Dichotomiser 3) and K- Nearest Neighbors Regression algorithms used for prediction [4].

Crop Planning, Agricultural Intelligence with Weather Data” Data management, location visualization, chart production using PostgreSQL, PostGIS and mapping libraries like Leaflet [5].

We are looking at the development of a Specific Agriculture Based Information System including Geo-Spatial Data. Geo-spatial data is represented by interactive maps and Restful APIs such as the Sparklyr7 R interface of the Apache Spark [6]. Landscape Analysis and Climate Data Predicting Crops and Group of Bangladesh Agricultural Distribution Techniques such as K- means, PAM, CLARA and DBSCAN integration and four consecutive retrospective methods are used [7].

Improving productivity with big data tools in the agricultural sector. This article describes the yield forecast based on big data tools such as Hive, Scoop and Apache Hadoop [8].

Effective crop yield analysis based on the Random Forestry Approach using the Hadoop Framework. This method is used to predict crop yields using the Random Forest approach [9].

An algorithm for a strong and innovative time series to predict rice yields based on solid and fresh regression. Fuzzy Logical Relations of Different Degrees and Model Analysis Backlash Used [10] next in terms of frequency-based classification.

Existing Methods:

1) COLLABORATIVE FILTERING METHOD:

Incorrect set-based filters are presented in the study [2] to predict the average crop category of non-consumer crop categories and to improve the quality of new plant material. Provide custom cropping suggestion to customers using end-to-end system prototype comparing user site data with crop point sites, and repeat those topics according to youth and CPCC similarity. There are many advantages: Variety of plant material can be found automatically. An effective way to automatically find a missing user rating. Acquisition of a dynamic community does not work with this tool.

The access logs of the online plant publisher's web server are analyzed on paper [4] to determine trends in online learning. The model that can be used to predict which articles will be read by a particular user first develops, then evaluates the most important features of the articles and explains the model that has been read. There are many benefits: Predictability accuracy can be greatly affected by selecting the time window. The following is the downside: The addition of a user feature, which is the most in need of a computer.

[6] introduces PENETRATE, an integrated hierarchical based clustering-based Personalized Crop Recommendation Framework that provides attractive recommendation results. Unlike content-based strategies and collaborative filtering, we present a framework that takes into account both individual user activity and team behavior at the same time. In terms of benefits, they include more accuracy and efficiency. The disadvantages include: long-term investments.

According to this paper [5] it is possible to model users' reading preferences using long-term and short-term user profiles seamlessly. The long-term profile should be used to identify plant species and subspecies, and the short-term profile should be used to identify specific plant material to individual users. As a benefit, the short-term model can capture user preferences for modified or newly introduced crop titles. Increasing user interest in learning by introducing a wide range of topics and styles.

2) CONTENT BASED RECOMMENDATION:

According to the study [8], there are a number of measuring methods for making plant recommendations, which include multi- dimensional similarity calculations, rapid integration using Kmeans integration, and Top-N recommendations.

Common shared filtering suffers from a small amount of data, but a multi-dimensional analytical method solves this problem by calculating how two identical users are based on the rich plant properties and their behavior and time. The recommended recommendation method has the following advantages. It effectively solves the problem of stiffness. With less data, the quality of the plant recommendations is improved. Only content-based recommendations are presented in this project, not collaborative filtering.

Web content recommendation is a paper policy [9]. Users' long-term interests and popular models are built using their roaming history and connected to a recommendation system. The importance of the web is given to users based on whether the content is similar to their mental models. Benefits include: a roaming model that frees users from annoying and repetitive web searches, automatically segmenting web pages, and increasing productivity.

There are two ways in the paper [10] that consider the meaning of words. In processes, concepts and semantic similarities are used to clarify similarities between film elements. Second Method, Semantic Similarity (SSV), First Method, Sync Frequency - Inverse Document Frequency (SF-IDF) (SS). Ceryx, the Hermes Crop Portal's personal botanical service addon, uses the recommended methods for recommending recommended items. There are many benefits: It exceeds the TF-IDF at the mathematical level.

3) HYBRID METHOD

Personal crop recommendations with a clear semantic analysis (LP-ESA), a proposed mixed method that [1,] takes into account both user preferences and interests while making crop recommendations an alternative method, known as LP-DSA (in-depth analysis) semantic, proposed to address issues of LSA-large LP size, minimum, and duplication using in-depth semantic analysis. Based on the recommendations of the deep neural network, LP-DSA quotes include presentations from locals, pedestrians, and other organizations. For example, the LP-DSA method uses deep neural networks to address the magnitude of LP-ESA's dimensionality, sparseness, and unwanted problems. Improve the efficiency and effectiveness of plant recommendations. Calculation time reduced. The following is the downside: Improve LP-DSA's performance by becoming an advanced space master of the invisible feature.

It proposes a personal crop recommendation in the article [3], a combination of content-based filtering algorithms and filtering algorithms that use specific domain information embedded in plants to produce recommendations (i.e., CB and CF, for short). By using both specific and standard domain features, this is the first recommendation program of its kind, to our knowledge. To increase the accuracy of plant recommendations, this paper proposes a novel algorithm for CB that makes the advantage of a trending feature (a feature-based feature from plants), which shows that different plant categories have different life cycles and play different roles in collecting users. profiles. There are many benefits: Benefits of using plant-specific features to improve crop recommendation systems. Both individual approaches and balanced mix strategies work better than diversification-based strategy in terms of accuracy and stability. FereBSP and FereRBML are better than their comparable benchmarks in terms of performance. The following is the downside: Those publishers who are able to track their readers for a long time are the only ones who can benefit from the recommended program. Based on both user content and integrated filtering of Wesomender architecture, a mixed recommendation engine.

Collaborative filtering and content-based filtering are two important aspects of Wesomender [7]. Each section evaluates and makes independent crop recommendations that the user has not yet considered or evaluated. There are many benefits to using sensitive content in journalism. Flexible complimentary search engines that recognize the context can assist journalists with their day-to-day activities by obtaining relevant information in real time. By rejecting older plants or those that occur far too far, heuristics can be used to reduce effort in one area.

4) MULTIDIMENSIONAL APPROACH

Additionally, it allows for Maximum size, user profile details, and sequential integration of recommendation programs. It is possible to estimate critical measurements using this method, using a machine-readable retrospective model trained to simulate integrand targeting. In order to ensure that the final end result has a

statistically accurate measure of uncertainty due to the prediction error introduced by machine learning, any bias in the scale is corrected using a biased correction term.

5) TF-IDF VECTORIZATION

Data retrieval and data extraction are key functions of the TF-IDF sub-project. An effort is made to highlight the importance of the word in the text that forms part of the document chorus. When using the document term-frequency frequency, the text is converted into a usable vector. Terms Frequency (TF) is combined with Document Frequency (DF) (DF). The frequency of a term in a document is measured using the term frequency. How frequently a word is used can indicate how essential it is inside an article or other piece of writing. Each document is represented in the data as a matrix with rows and columns that represent the number of distinct phrases in each document.

6) RANDOM FOREST

Random forests are a machine learning algorithm. The basic premise of the algorithm is that building a small resolution tree with fewer features is a computer process. The algorithm works like this: for each tree in the forest, we select a bootstrap sample from S when $S(i)$ means i th bootstrap. Then we read the decision tree using the modified decision tree algorithm. The algorithm is configured as follows: in each tree area, instead of exploring all the possible fragments, we randomly select a specific set of features.

Where, F is the set of features. The node is then split into F . By making f greater, less than F . Determining which aspect to distinguish is often the most expensive computer-aided decision-making process. By reducing the set of features, we greatly accelerate the learning of the tree

III. PROPOSED SYSTEM

The project predicts crop and crop yields, taking into account information such as soil type, temperature, humidity, time of year, fertilizer and months. The system provides easy and quick access to all the basic information about location, rainfall, irrigation, crops, annual yields, fertilizers that users can use to analyze crops and choose to predict farming parameters in order to get the right yield for their farm. This program provides easy visibility, allowing users to easily understand and analyze objects.

An online auction program is a web-based system, where the seller can sell goods by staying in his or her home, so the main advantage of this app is that it is no longer a problem with the requirement to comply with the system. The great advantage of an online auction system is that the user can have the best and most time-saving options, and with this program the user can invest in his chosen company. This project covers the shortcomings encountered during the auction process which may pose a significant safety risk. It uses a web application that will come in handy when you do. The applicant will be able to file his or her bid without tracking down malicious people. The user will be able to set their own information which will be stored and can only be viewed by the administrator. This will ensure safety from exposure to robbers and other malicious people.

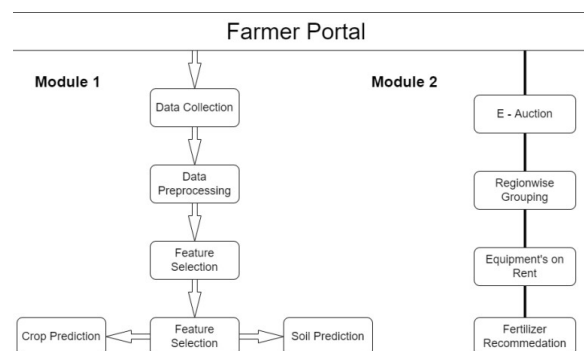


Figure: System Architecture

IV. ADVANTAGES:

- This system provides access to fundamental information quicker and easier.
- The User can adapt the crop parameters with this system.
- Results and prediction easy to understand.
- The system will be authenticated.
- Improved understanding of trends in agriculture in various areas.

V. LITERATURE REVIEW

- 1 Jason Rhuggenaath, Alp Akcay, Yingqian Zhang and Uzay Kaymak, "Fuzzy Logic based Pricing combined with Adaptive Search for Reserve Price Optimization in Online Ad Auctions". 2019 IEEE, in this paper we review an online publisher that sells advertising space and suggest how to read the right reservation prices at second pricing auctions. We are reading a limited information setting where bids are not disclosed and there is no historical information about bids available. Our proposed method combines a flexible search process with pricing action to set fixed prices and is suitable for non-stop areas.
At a reasonable price point, we take the gap between the winning bidder and the second highest bid and show that this leads to better pricing decisions. Examination using real-life ad auction data shows that the proposed method exceeds the popular bandit algorithms.
- 2 Ermatita, Ika Nurlaili Isnainiyah, Yulnelly Yulnelly, Amalia Nurul Balqis, "Usability Analysis using Principal Component Analysis (PCA) Method for Online Fish Auction Application". 2019 IEEE, the information systems or applications used by users need to go through a high level of analysis, depending on the needs of the information systems or applications that can be used. Performance monitoring and analysis and user engagement rate needs to be done to measure the level of application of the auction in the online fish auction. This analysis uses the Kaiser-Meyer-Olkin Measure of Sampling Adequacy and the Bartlett's Test of Sphericity methods to assess eligibility. Also, use the Main Component Analysis method to find size features that represent all sizes. The test results of the index show some of the most important variables, namely PUS5 and FI3 with the highest values of 0.926 and 0.792, which are indicators representing the variable involvement (Visual Performance and Sensible Involvement).
- 3 Dou An, Qingyu Yang, Wei Yu, Xinyu Yang, Xinwen Fu, and Wei Zhao, "SODA: Double Auction Scheme Online Strategy-Proof Online for Multi-micro-grids Bidding". 2017 IEEE, in this paper, introduces the theory and design of a dual online auction of power trading within a smart grid with microgrid (MGs). Double internet auction has the potential to cause the distribution of the remaining electricity to MGs that require the most profitable electricity in real-time markets. However, two important issues remain a challenge when designing a dual online auction system in such a system.
- 4 First, since agents are allowed to come and go at any time, the seller needs to make decisions without the knowledge of additional bids and inquiries. Second, the economic aspects of strategic evidence, individual flexibility, and (weak) budget balance must be satisfied.
To address these issues and to facilitate multi-unit power transactions between local MGs, in this paper, we propose a dual online auction system (SODA) with proof of strategy, where the remaining MGs and MGs in the system are considered as retailers and consumers, respectively, and the MG center controller is able to enhance the MG social well-being by equating buyers and sellers appropriately. Through theoretical analysis, we prove that SODA can achieve aspects of individual considerations, (weak) budget balance, strategic validation, and computer efficiency. Testing also shows that SODA is capable to reduce the cost of purchasing energy for MGs and to change volume, while achieving efficiency in terms of social well-being, consumer satisfaction/consumer satisfaction, social efficiency, and accounting.

VI. FUTURE WORK

The future work on this project can be adding more options such as:

1. Crop disease detection and prevention, crop price prediction.
2. Current Market status and analysis for getting the information about market rates of crop, production cost, fertilizers.
3. Government schemes option to know new government schemes related with loans, fertilizers, and crop.

REFERENCES

- [1] Ponce-Guevara, K. L., Palacios-Echeverria, J. A., Maya-Olalla, E., Dominguez Limaico, H. M., Suarez-Zambrano, L. E., Rosero-Montalvo, P.D., Alvarado-Perez, J. C. (2017). Green Farm-DM: A tool for analyzing vegetable crops data from a greenhouse using data mining techniques (First trial). 2017 IEEE Second Ecuador Technical Chapters Meeting (ETCM).
- [2] Jheng, T.-Z., Li, T.-H., Lee, C.-P. (2018). Using hybrid support vector regression to predict agricultural output. 2018 27th Wireless and Optical Communication Conference (WOCC).
- [3] Manjunatha, M., Parkavi, A. (2018). Estimation of Arecanut Yield in Various Climatic Zones of Karnataka using Data Mining Technique: A Survey. 2018 International Conference on Current Trends Towards Converging Technologies (ICCTCT).
- [4] Shakoor, M. T., Rahman, K., Rayta, S. N., Chakrabarty, A. (2017). Agricultural production output prediction using Supervised Machine Learning techniques. 2017 1st International Conference on Next Generation Computing Applications (NextComp).
- [5] Grajales, D. F. P., Mejia, F., Mosquera, G. J. A., Piedrahita, L. C., Basurto, C. (2015). Crop-planning, making smarter agriculture with climate data. 2015 Fourth International Conference on Agro-Geoinformatics (Agro-Geoinformatics).
- [6] Shah, P., Hiremath, D., Chaudhary, S. (2017). Towards development of spark based agricultural information system including geo-spatial data. 2017 IEEE International Conference on Big Data (Big Data).
- [7] Afrin, S., Khan, A. T., Mahia, M., Ahsan, R., Mishal, M. R., Ahmed, W., Rahman, R. M. (2018). Analysis of Soil Properties and Climatic Data to Predict Crop Yields and Cluster Deferent Agricultural Regions of Bangladesh. 2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS).
- [8] Sekhar, C. C., Sekhar, C. (2017). Productivity improvement in agriculture sector using big data tools. 2017 International Conference on Big Data Analytics and Computational Intelligence (ICBDAC).
- [9] Sahu, S., Chawla, M., Khare, N. (2017). An efficient analysis of crop yield prediction using Hadoop framework based on random forest approach. 2017 International Conference on Computing, Communication and Automation (ICCCA).
- [10] Garg, A., Garg, B. (2017). A robust and novel regression based fuzzy time series algorithm for prediction of rice yield. 2017 International Conference on Intelligent Communication and Computational Techniques (ICCT).