



Real time Data Analytics in Crop Management based on Weather Conditions using Machine Learning

Dr.V.Suganthi¹, R.Hariprakash²

¹Associate Professor, PG & Research Department of Computer Science, Sri Ramakrishna College of Arts & Science

²PG Student, PG & Research Department of Computer Science, Sri Ramakrishna College of Arts & Science

Abstract: As a coastal state, Tamil Nadu faces uncertainty in agriculture which decreases its production. With more population and area, more productivity should be achieved but it cannot be reached. Farmers have words-of-mouth in past decades but now it cannot be used due to climatic factors. Agricultural factors and parameters make the data to get insights about the Agri-facts. Growth of IT world drives some highlights in Agriculture Sciences to help farmers with good agricultural information. Intelligence of applying modern technological methods in the field of agriculture is desirable in this current scenario. Machine Learning Techniques develops a well-defined model with the data and helps us to attain predictions. Agricultural issues like crop prediction, rotation, water requirement, fertilizer requirement and protection can be solved. Due to the variable climatic factors of the environment, there is a necessity to have a efficient technique to facilitate the crop cultivation and to lend a hand to the farmers in their production and management. This may help upcoming agriculturalists to have a better agriculture. A system of recommendations can be provided to a farmer to help them in crop cultivation with the help of data mining. To implement such an approach, crops are recommended based on its climatic factors and quantity. Data Analytics paves a way to evolve useful extraction from agricultural database. Crop Dataset has been analyzed and recommendation of crops is done based on productivity and season

Keywords: weather condition, Machine learning algorithms, Types of crops, fertilizer.

I. INTRODUCTION

Agriculture is the backbone of Indian Economy. In India, majority of the farmers are not getting the expected crop yield due to several reasons. The agricultural yield is primarily depends on weather conditions. Rainfall conditions also influences the rice cultivation. In this context, the farmers necessarily requires a timely advice to predict the future crop productivity and an analysis is to be made in order to help the farmers to maximize the crop production in their crops. People of India are practicing Agriculture for years but the results are never satisfying due to various factors that affect the crop yield. To fulfill the needs of around 1.2 billion people, it is very important to have a good yield of crops. Due to factors like soil type, precipitation, seed quality, lack of technical facilities etc the crop yield is directly influenced. Hence, new technologies are necessary for satisfying the growing need and farmers must work smartly by opting new technologies rather than going for trivial methods.

Machine learning is the process of extract helpful and significant information from huge sets of data. Data Mining in agriculture field is a comparatively novel research field. Yield prediction is a very important agricultural problem. Any farmer is interested in knowing how much yield he is concerning to be expecting. In the earlier period, yield prediction was performing by considering farmer's experience on particular field and crop. In any of Data Mining actions the training data is to be collected from past data and the gathered data is used in terms of training which has to be exploited to study how to categorize future yield predictions.

From ancient period, agriculture is considered as the main and the foremost culture practiced in India. Ancient people cultivate the crops in their own land and so they have been accommodated to their needs. Therefore, the natural crops are cultivated and have been used by many creatures such as human beings, animals and birds. The greenish



goods produced in the land which have been taken by the creature leads to a healthy and welfare life. Since the invention of new innovative technologies and techniques the agriculture field is slowly degrading. Due to these, abundant invention people are been concentrated on cultivating artificial products that is hybrid products where there leads to an unhealthy life. Nowadays, modern people don't have awareness about the cultivation of the crops in a right time and at a right place. Because of these cultivating techniques the seasonal climatic conditions are also being changed against the fundamental assets like soil, water and air which lead to insecurity of food. By analyzing all these issues and problems like weather, temperature and several factors, there is no proper solution and technologies to overcome the situation faced by us. In India there are several ways to increase the economical growth in the field of agriculture.

Crop yield prediction is an important agricultural problem. Each and Every farmer is always tries to know, how much yield will get from his expectation. In the past, yield prediction was calculated by analyzing farmer's previous experience on a particular crop. The Agricultural yield is primarily depends on weather conditions, pests and planning of harvest operation. Accurate information about history of crop yield is an important thing for making decisions related to agricultural risk management. This research focuses on evolution of a prediction model which may be used to predict crop yield production. The proposed method use data mining technique to predict the crop yield production based on the association rules. The main objective of these problems is collecting agricultural dataset which can be used to analysed for useful crop yield forecasting. To predict the crop yield with the help of data mining technique, advanced methods can be introduced to predict crop yield and it is also helps the farmer to choose the most suitable crop, thereby improving the value and gain of the farming area.

II. RELATED WORKS

An agricultural sector necessitate for well defined and systematic approach for predicting the crops with its yield and supporting farmers to take correct decisions to enhance quality of farming. The complexity of predicting the best crops is high duet unavailability of crop knowledge-base. Crop prediction is an efficient approach for better quality farming and increase revenue. Use of data clustering algorithm is an efficient approach in field of data mining to extract useful information and give prediction. Various approaches have been implemented so far are worked either for crop prediction. Crop prediction model aiding farmers to take correct decision. This indeed helps in improving quality of farming and generates better revenue for farmers. Traditional clustering algorithms such as k-Means, improved rough k-Means and-means++ makes the tasks complicated due to random selection of initial cluster center and decision of number of clusters. Modified K-Means algorithm is thereby used to improve the accuracy of a system as it achieves the high quality clusters duet initial cluster centric selection [1].

Data driven precision agriculture aspects, particularly the pest/disease management, require a dynamic crop-weather data. An experiment was conducted in a semi-arid region to understand the crop-weather-pest/disease relations using wireless sensory and field-level surveillance data on closely related and interdependent pest (Thrips) - disease (Bud Necrosis) dynamics of groundnut crop. Data mining techniques were used to turn the data into useful information/knowledge/relations/trends and correlation of crop-weatherpest/disease continuum. These dynamics obtained from the data mining techniques and trained through mathematical models were validated with corresponding surveillance data. Results obtained from 2009 & 2010 kharif seasons (monsoon) and 2009-10 & 2010-11 rabi seasons (post monsoon) data could be used to develop a real to near real-time decision support system for pest/disease predictions.

Agriculture is the most basic function to accomplish food demand all over the globe; it is a backbone particularly in the developing countries like India. The application of Data mining techniques in agriculture especially on soils can revise the situation of pledge making and improve cultivation yields in a better way. The analysis of soils plays an indispensable role for resolution making on several issues related to agriculture field. This paper presents about the role of data mining in perspective of soil analysis in the field of agriculture and also confers about several data mining techniques and their related work by several authors in context to soil analysis domain. The data mining techniques are of very up-to-the-minute in the area of soil analysis.



Soil is an essential key factor of agriculture. The objective of the work is to predict soil type using data mining classification techniques. Methods/Analysis: Soil type is predicted using data mining classification techniques such as JRip, J48 and Naive Bayes. These classifier algorithms are applied to extract the knowledge from soil data and two types of soil are considered such as Red and Black. Findings: In this paper, Data Mining and agricultural Data Mining are summarized. The JRip model can produce more reliable results of this data and the Kappa Statistics in the forecast were increased. Application/Improvement: For solving the issues in Big Data, efficient methods can be created that utilize Data Mining to enhance the exactness of classification of huge soil data sets.

Agriculture is the most significant application area particularly in the developing countries like India. Data mining plays a crucial role for decision making on several issues related to agriculture field. The goal of the data mining process is to extract knowledge from an existing data set and transform it into a unique human understandable format for some advance use. Crop management of certain agriculture region is depends on the climatic conditions of that region because climate can make huge impact on crop productivity. Real time weather data can helps to attain the good crop management. Utilization of information and communications technology enables automation of extracting significant data in an effort to obtain knowledge and trends, which enables the elimination of manual tasks and easier data extraction directly from electronic sources, transfer to secure electronic system of documentation which will enable production cost reduction, higher yield and higher market price. Also identified that how the data mining helps to analyze and predict the useful pattern from huge and dynamically changed climatic data. In agricultural and biological engineering, researchers and engineers have developed methods of fuzzy logic, artificial neural networks, genetic algorithms, decision trees, and support vector machines to study soil, climate conditions and water regimes related to crop growth and pest management in agriculture. In this paper summarizes the application of data mining techniques, Neural Networks, Support Vector Machine, Big Data analysis and soft computing in the agriculture field base on weather conditions.

III. METHODOLOGY

Crop production depends on many agricultural parameters. Proposed work is based on the production of crops in previous years, crops can be recommended to the farmers. This kind of suggestions will make farmer to know that whether that particular is yielding a good production in recent years. Production of crops may become less due to any crop disease, water problem and many other factors. While considering about the production, farmers may get knowledge about which crop is in high volume in the market in that year. Based on this farmer can take decision of trend on crops in recent years. Farmers will be given recommendation by considering the season of crop production. The problem statement of the paper is to recommend crops to the farmers using Decision Tree Classifier. The basic process of this paper is that we will pre-process the data provided to us, then it is used to prepare the model for the backed and using flask to connect it to the UI interface to show the full and final output.

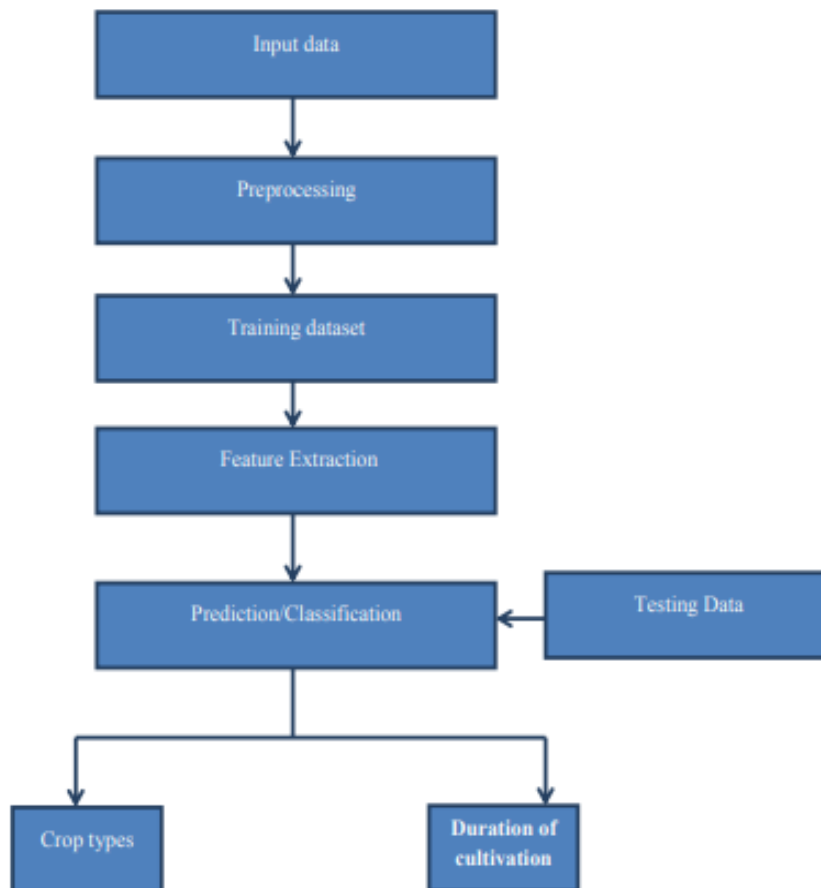


Figure-1 System Architecture

Proposed Algorithm

Three different machine learning algorithms has taken for simple analysis. the three algorithms are then compared with some quality metrics such as Classification Ratio, Detection Ratio and Malicious Ratio. The experimental results from weka tools are displayed for analyzing different results.

3.1 K-Nearest Neighbors Algorithm

In design acknowledgment, the k-Nearest Neighbors algorithm (or k-NN for short) is a nonparametric technique utilized for classification and regression.[1] In the two cases, the info comprises of the k nearest preparing models in the component space. The yield relies upon whether k-NN is utilized for classification or regression: In k-NN classification, the yield is a class participation. An article is grouped by a dominant part vote of its neighbors, with the item being appointed to the class generally basic among its k nearest neighbors (k is a positive whole number, ordinarily little). In the event that $k = 1$, at that point the article is basically relegated to the class of that solitary nearest neighbor. In k-NN regression, the yield is the property estimation for the article. This worth is the normal of the estimations of its k nearest neighbors. k-NN is a kind of occasion based learning, or languid learning, where the capacity is just approximated locally and all calculation is conceded until classification. The k-NN algorithm is among the most straightforward of all machine learning algorithms.

3.2 Naive Bayes:

It is a likelihood based classification procedure. It considers all highlights autonomous of one another. It computes likelihood of each element freely for a specific class mark. Naïve Bayes is utilized in this paper for malware prediction utilizing web traffic data. These are the means behind the Naïve Bayes algorithm:

1. Preparing data set is taken as information.
2. Highlights are extricated from that preparation data. In this paper web traffic data comprises of 43 highlights.



3. At that point from the preparation data for each component Naïve bayes figures likelihood that in the event that element has specific worth, at that point the dataset class be will malicious or not.
4. In the event that each component has constrained potential qualities, at that point above probabilities can be determined. Be that as it may, if the huge number of qualities is there for each element, scope of qualities can likewise be taken.
5. At that point for each line of test data set after the preparation stage. Based all things considered probabilities determined from preparing data decision is taken.

3.3 Decision tree:

This kind of classifier models data with the assistance of a tree. Tree is having highlights as the inward hubs and edges show the estimations of highlights. And edges isolated hubs dependent on the qualities. All the leaf hubs of the decision tree speaks to a class which is relied upon to be acquired on the off chance that we have every one of the highlights having particular qualities which are in the way from the root to that class having middle of the road include hubs. Probably the most mainstream decision tree algorithms are ID3, C4.5, CART. ID3 is one of the most straightforward decision tree approaches it utilizes idea of data gain as the parting criteria. C4.5 is the development of ID3. It takes a shot at the guideline of addition proportion.

IV. RESULTS AND DISCUSSION

Three different algorithms are used for machine learning comparison.

Classification Ratio

| K-Nearest Neighbor Algorithm | Decision Tree Algorithm | Naïve Bayes Algorithm |
|------------------------------|-------------------------|-----------------------|
| 0.02 | 0.09 | 0.04 |
| 0.05 | 0.14 | 0.08 |
| 0.09 | 0.19 | 0.13 |
| 0.14 | 0.25 | 0.19 |
| 0.19 | 0.3 | 0.22 |

Table 1: Comparison table of Classification Ratio

The comparison table of classification ratio explains the different values of K-Nearest Neighbor algorithm, Decision Tree Algorithm and Naïve Bayes Algorithm. While comparing these algorithms the decision tree algorithm is better than the other. The K-Nearest Neighbor Algorithm values are starts from 0.02 to 0.19, The decision tree algorithm values are starts from 0.09 to 0.3, naïve bayes algorithm values are starts from 0.04 to 0.22.

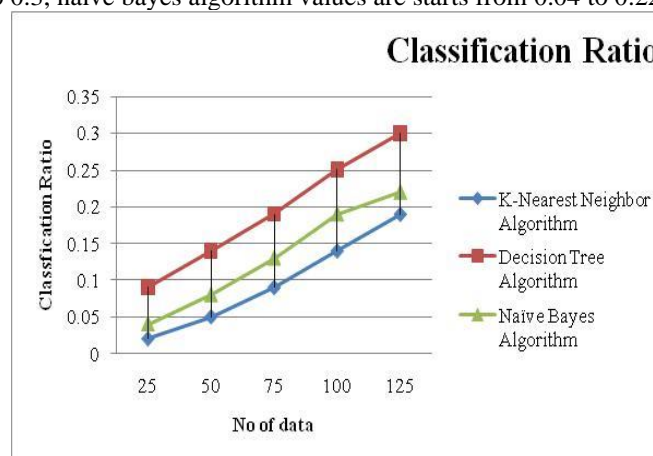


Figure 2: Comparison chart of Classification Ratio

The comparison chart of classification ratio shows the values of K-Nearest Neighbor algorithm, Decision Tree Algorithm and Naïve Bayes Algorithm. No of data in X axis and classification ratio in Y axis. The K-Nearest Neighbor Algorithm values are starts from 0.02 to 0.19, The decision tree algorithm values are starts from 0.09 to 0.3, naïve bayes algorithm values are starts from 0.04 to 0.22. the decision tree algorithm values are higher than the other algorithm.



Figure 3- Sample Classification results

V. CONCLUSION

In this paper, significance of management of crops was studied vastly. Farmers need assistance with recent technology to grow their crops. Proper prediction of crops can be informed to agriculturists in time basis. Many Machine Learning techniques have been used to analyse the agriculture parameters. Some of the techniques in different aspects of agriculture are studied by a literature study. Blooming Neural networks, Soft computing techniques plays significant part in providing recommendations. Considering the parameter like production and season, more personalized and relevant recommendations can be given to farmers which makes them to yield good volume of production.

REFERENCES

- [1] Shreya S. Bhanose, Kalyani A. Bogawar (2016) "Crop And Yield Prediction Model", International Journal of Advance Scientific Research and Engineering Trends, Volume 1, Issue 1, April 2016
- [2] Tripathy, A. K., et al.(2011) "Data mining and wireless sensor network for agriculture pest/disease predictions." Information and Communication Technologies (WICT), 2011
- [3] Ramesh Babu Palepu (2017) "An Analysis of Agricultural Soils by using Data Mining Techniques", International Journal of Engineering Science and Computing, Volume 7 Issue No. 10 October.
- [4] Rajeswari and K. Arunesh (2016) "Analysing Soil Data using Data Mining Classification Techniques", Indian Journal of Science and Technology, Volume 9, May.
- [5] A.Swarupa Rani (2017), "The Impact of Data Analytics in Crop Management based on Weather Conditions", International Journal of Engineering Technology Science and Research, Volume 4, Issue 5, May.
- [6] Pritam Bose, Nikola K. Kasabov (2016), "Spiking Neural Networks for Crop Yield Estimation Based on Spatiotemporal Analysis of Image Time Series", IEEE Transactions On Geoscience And Remote Sensing.
- [7] Priyanka P.Chandak (2017), "Smart Farming System Using Data Mining", International Journal of Applied Engineering Research, Volume 12, Number 11.



- [8] Vikas Kumar, Vishal Dave (2013), "KrishiMantra: Agricultural Recommendation System", Proceedings of the 3rd ACM Symposium on Computing for Development, January.
- [9] Savae Latu (2009), "Sustainable Development : The Role Of Gis And Visualisation", The Electronic Journal on Information Systems in Developing Countries, EJISDC 38, 5, 1-17.
- [10] Nasrin Fathima.G (2014), "Agriculture Crop Pattern Using Data Mining Techniques", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 4, May.
- [11] Ramesh A.Medar (2014), "A Survey on Data Mining Techniques for Crop Yield Prediction", International Journal of Advance Research in Computer Science and Management Studies, Volume 2, Issue 9, September.
- [12] Shakil Ahamed.A.T.M, Navid Tanzeem Mahmood (2015)," Applying data mining techniques to predict annual yield of major crops and recommend planting different crops in different districts in Bangladesh", ACIS 16th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD),IEEE,June.
- [13] Napoleon D. and Praneesh M. "Detection of Brain Tumor using Kernel Induced Possiblistic C-Means Clustering", volume no.3, issue no.9, pp 436-438, 2013
- [14] Shreya S.Bhanose (2016),"Crop and Yield Prediction Model", International Journal of Advence Scientific Research and Engineering Trends, Volume 1,Issue 1,ISSN(online) 2456- 0774, April.
- [15] Agaj i Iorshase, Onyeke Idoko Charles,"A Well-Built Hybrid Recommender System for Agricultural Products in Benue State of Nigeria", Journal of Software Engineering and Applications,2015,8,581-589
- [16] G. Adomavicius and A. Tuzhilin(2005), "Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions," IEEE Trans. Knowledge and Data Eng., vol. 17, no. 6, pp. 734-749, June.
- [17] Avinash Jain, Kiran Kumar (2016),"Application of Recommendation Engines in Agriculture", International Journal of Recent Trends in Engineering & Research, ISSN: 2455-1457.
- [18] Kiran Shinde (2015),"Web Based Recommendation System for farmers", International Journal on Recent and Innovation Trends in Computing and Communication, Volume 3,Issue 3, ISSN:2321- 8169,March.
- [19] Konstantinos G. Liakos, " Machine Learning in Agriculture: A Review", Sensors 2018, 18, 2674; doi:10.3390/s18082674
- [20] S.Vaishnavi, M.Shobana, N Geethanjali, Dr.S.Karthik, "Data Mining: Solving the Thirst of Recommendations to Users", IOSR Journal of Computer Engineering IOSR-JCE), Vol.16, no.6, 2014.