Rainfall Based, Crop Production Identification by using different Algorithm technique

[1]Rohan Pawar, [2]Isha Bhalchandra, [3]Yashraj Singh, [4]Prashobh Praveen, [5]Prof .Asmita Kalamkar

***Abstract*: Rainfall depends on various factors such as temperature, humidity, cloudiness, wind speed, etc. Rainfall prediction is a major concern for many fields like the meteorological department, agriculture department, etc which is directly associated with the economy and sustenance of human life. In the proposed system we are going to predict the rainfall and after this prediction, estimate the yield, solutions for fertilizing or growing alternative crops. A further step in the process also consists of finding out if the crop has certain kind of disease or not. Agriculture study is growing rapidly, due to innovation in technologies and upcoming challenges. It has been proven to play a crucial role in improving the overall growth rate of any country. Agriculture matters a lot in economic growth. Very wide work is going on from several years to improve productivity by using data mining & evolutionary techniques. The losses incurred in crop production are a large issue for agriculture as they come in the way of sufficient produce. There is a very limited analysis performed on the loss of crops, due to the changing environmental conditions. The correct prediction of loss in crops helps the farmer to apply suitable actions to overcome these issues and maintain the expected production. As agriculture matters a lot in economic growth, to overcome this issue, we have used the data mining approach i.e classification. The objective of the system is to calculate the produce of a particular crop, based on the rainfall patterns. Data mining techniques help in agriculture for the exclusion of manual jobs and for decision making which enables to decrease the production cost and improve productivity.**

***Keywords*: Crop, Classification, Naïve Bayes, KNN, RF**

Introduction:

The agricultural study has strengthened the optimal economical income, globally, and is a vast and important area to gain more benefits. However, it can be enhanced by the use of different technological resources, tools, and procedures. Today, the term data mining is an interdisciplinary process of analyzing, processing and evaluating the real-world data-sets and prediction based on the patterns found. Data mining is the collection of exponentially growing techniques that are used to find some useful information, patterns, and knowledge from already given data. This useful information helps to advance the existing research and enhance productivity. The applications of data mining are countless; it can be applied in every aspect of life. Some major applications of data mining are Health-care, Market Analysis, Finance, Education, Manufacture Engineering, Corporation Surveillance, and Agriculture [6]. Lots of technology and research have been developed in the field of agriculture. To improve the rate of agriculture, the researchers have used data mining techniques to solve problems arising in agriculture. In this paper data mining technique is used to find patterns in the rainfall [3] and provide further solutions and suggestions. It supports the farmer while taking the right decisions about agriculture. This paper focuses on the prediction of the crop produce and rainfall using data-sets. For the classification and prediction purposes, the Naïve Bayes, KNN, RF algorithms are used in the proposed system.

Literature Survey:

Agriculture's input in GDP is important for many countries, especially for Asian countries. Researchers are doing comprehensive work from the last few years to enhance agricultural productivity by incorporating Data Mining techniques in their research and development.

In [13] researchers have proved with output that agriculture is a backbone of economic growth for many countries. The prediction of diseases and loss in crop help farmers to improve it.

There are different classifiers used to build a model for predicting the spray decision on crops [5,6,9].

In [2,12] authors have used Decision Tree, Support Vector Machine and Neural Network classifiers. They concluded that theses classifier have good prediction accuracy as compared to k^{th} Neural Network and GNB.

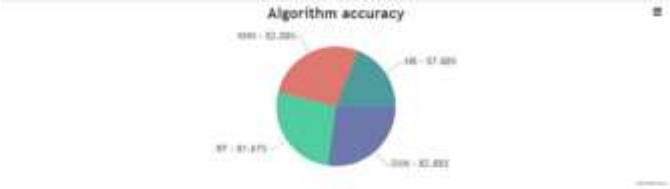
In one of the papers, proposed different matrices for evaluation such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and accuracy for both true & false classes in all mentioned classifiers. Author also describes that the use of ensemble model is a future research direction.

In [3] author articulated that different classifiers can be used on different features in ensemble model for predicting the true class. Each model for the ensemble classifier is trained on different set of features to achieve the best results. And high vote getting the class will be allocated to test sample.

In [5] authors provide an extensive survey on classifiers and explain the ways for better prediction. Authors explained that we cannot solve all the problems with same classifier. Different classifiers work in different way so different classifiers can be used in different scenarios.

Dhawal Hirani, Nitin Mishra [13] "A Survey on Rainfall Prediction Techniques". This paper mainly discusses the various machine learning techniques used for early prediction of rainfall and some cognitive approaches. Research mainly explained about approaches to empirical and dynamical Methods.

Pinky Saikia Dutta, Hitesh Tahbilder [14] "Prediction of Rainfall Using Data Mining Technique over Assam". In this paper, they have described data mining technique in forecasting monthly Rainfall and traditional statistical technique Multiple Linear Regression.



The classification of soil is also another problem that is stated in the Base Paper. The authors have evaluated certain regions in their country, based on which their system recommends the crops that are best suited. They have also focused on the importance of trained data-sets, which has also been highlighted in the approach involved in this very paper.

Also, given the environmental conditions experienced today, it can be extremely difficult to target the prediction patterns over a longer period. This is why the approach used in [2] as a monthly prediction is taken into consideration. This can help in accurate predictions over shorter, immediate seasonal periods.

Motivation:

In agriculture, if the end-user cannot recognize crop production, it would lead to a loss in crop production. So, early detection in crop production and rainfall prediction can help farmers to take the right decision for their crops and minimize the loss in crop production.

Gap Analysis:

The weather conditions today are highly unpredictable. It makes it very challenging to perform a long term analysis. This is why the proposed system is dealing with shorter cycles of weather, helping us produce results in the immediate 3 month periods. This way, farmers can use this methodology to enhance the crop produce and be aware of the entire scenario, which was not possible with the earlier proposed methods. The method employed here lets the farmers to not get swayed away and confused by the technicalities. The algorithms used in this system for classification and prediction are very efficient. The results of the prediction are capable to guide the farmers in taking agricultural decisions. In a lot of earlier systems, the factor of usability was not well focused on. Also, the efficiency of the algorithms employed for prediction was subject to fluctuations. Another gap that this system has managed to fill is the ability of farmers to interpret these complex numbers and predictions. The report generated is able to paint a clear picture about what needs to be done and the entire dynamics of crop production. This gives way for a clear base for the farmers to base their further steps and to experience a greater crop produce.

System Architecture:

The proposed system is based on the rainfall prediction and production prediction of the crop. The system will predict the rainfall in the area and also find out the impact of given conditions on the crop. The result will show if these conditions are optimal to grow the required crop. Also, the estimated crop produce will also be projected. Further enhancement solutions like fertilizers or cultivation of alternative crops would also be included. So, for prediction and classification, the Naïve Bayes algorithm is used.

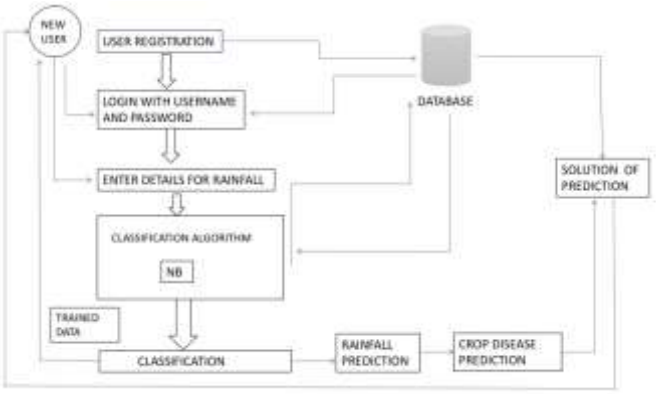


Fig. Proposed System

1. New user
This is the starting point where farmers are reached out to and persuaded in trying the proposed system.

2. User Registration
The farmer then gets registered to the system where his profile is created and the required parameters are entered.

3. Login Interface
Once registered, the farmer can access the interface and start the process to get a glimpse into the prediction process.

4. Detail Entry Point
This is where the input for all parameters, including rainfall in mm, temperature, humidity and other related parameters have to be entered, depending on the region where the farmer has to cultivate the crops.

5. Classification Algorithm
By taking references from the pool of Trained Data-set, the system then decides on what classification process would be the most ideal for the give parameters, to ensure maximum accuracy and efficiency in the prediction process.

6. Trained Data-set
For reference purposes, there is a Trained Data-set that is maintained in the system. This allows for a faster comparison to the kind of patterns input by the farmer and to decide the best methods and classification algorithms for prediction.

7. Classification
This is where the classification process will take place. The most efficient methods and algorithms for a given input will be employed for classifying the parameters that will then project the rainfall statistics.

8. Rainfall Prediction
This is the final output that the farmer will get. The statistics of the predicted rainfall on which the farmers' agricultural decisions will be based. This prediction would be based on certain periods for farmers to initiate immediate actions.

9. Solution
The solution will consist of the rainfall prediction pattern that is generated after processing. Also, there are suggestions related to the kind of fertilizers that can be used or alternate crops that are most suitable for the given climatic conditions.

10. Database
This database maintains the details of a farmer, the parameters input by them and the solutions which are generated at the end of the prediction process.

Mathematical Model

Mathematical model set theory S= {s, e, X, Y, Φ}

s= Start of the program

1. Register/Login into the system

2. Provide Attribute.

e= End of the program

Identify the Disease related to the Attribute

X= input of the program= {P, R, Q}

P = Type of disease

R= Rainfall Prediction Result.

Q = Using Machine learning algorithm predicting the disease

Y = Output of program = R

First, users provide the Attribute and predict the particular rainfall phase.

Let R be the set of Attribute

R = {R1, R2, R3... Rn} //Number of Attribute.

Let A be the set of disease type

Therefore,

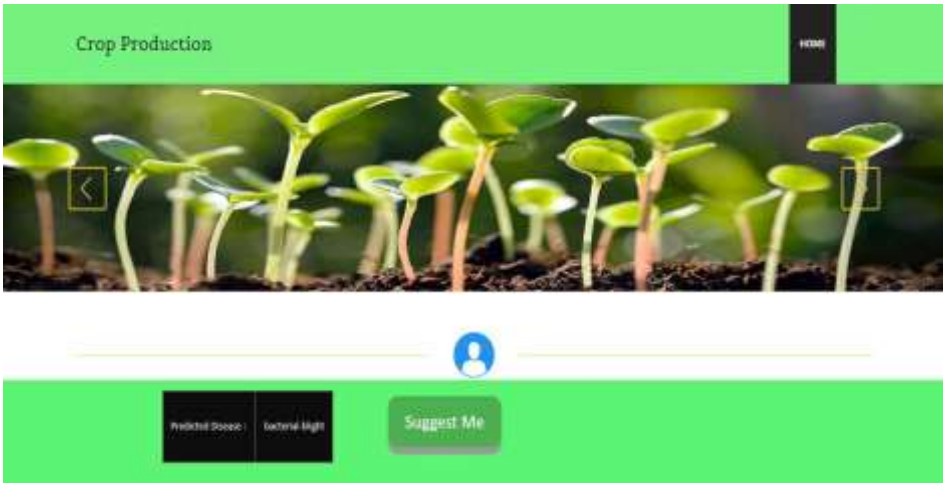
A = {A1, A2, A3..., Am} // Number of Disease

E = {E1, E2, E3,..., Em}

Overall Disease is evaluated with the help of this Attribute which basically represents rainfall prediction for the disease.

Y= E1+E2+... +Em / m

Where m is number of overall Disease



Conclusion:
In this paper, the Identification of classification algorithms with the best accuracy is done to predict rainfall. Furthermore, Identification of the perfect crop for those particular conditions and the estimation of yield is performed. The solution is then provided to the user according to the prediction. Based on these results, the farmers can decide what measures to take for their crops to survive in the predicted conditions. Also, if the existing crop is found to be unfavorable for the projected rainfall and climate conditions, suggestions for alternative crops are given. This will ensure that the land does not go uncultivated in any seasonal cycle.

Future Work:
The project involves the prediction of crop production based on weather and rainfall prediction. It is aiming to support and aid the farmers in understanding the dynamics of their crop cultivation. The system is capable of assisting them in making existing changes that can be done to use their soil for its complete, optimal use, its components, fertility and the geophysical characteristics that are in play. Also, it can get them acquainted with the alternate choices they have for cultivation on the same land.

Furthermore, the project aims on bringing marketplaces closer to their farms so that the fresh produce reaches to the markets. This can ensure an entirely utilized farming and market ecosystem under any conditions.

The work is also striving toward a better interface, a better interpretation of all this data and facts to the farmers for a clearer decision making process. Maximization of crop produce under and climatic conditions is of top priority. This is possible by working constantly towards improving the efficiency of the algorithms used and prediction of crop diseases as well as a further development stage.

References:

[1] C. C. Aggarwal, Data mining: the textbook. Switzerland: Springer International Publishing Switzerland, 2015.

[2] Monthly Rainfall Forecasting Using One-Dimensional Deep Convolutional Neural Network: A Haidar, B Verma - IEEE Access, 2018

[3] Rainfall prediction methodology with binary multilayer perceptron neural networks: JT Esteves, G de Souza Rolim, AS Ferraudo - Climate Dynamics, 2019 - Springer

[4] R. Chattamvelli, Data mining methods. Oxford, United Kingdom: Alpha Science International Ltd, 2009.

[5] C. C. Aggarwal, Data classification: algorithms and applications. Florida: CRC Press, 2014.

[6] Patel, Hetal and D. Patel, "A brief survey of data mining techniques applied to agricultural data," International Journal of Computer Applications, vol. 9, pp. 6-8, 2014.

[7] M. Stoces, J. Vanek, J. Masner and J. Pavlek, "Internet of things (iot) in agriculture-selected aspects," Agris on-line Papers in Economics and Informatics, vol. 8, pp. 83-88, 2016.

[8] B. Milovic and V. Radojevic, "Application of data mining in agriculture," Bulgarian Journal of Agricultural Science, vol. 5, pp. 18- 21, 2017.

[9] D. Kwetishe and O. Adenike, "Evaluation of predictive data mining algorithm in soil data classification for optimized crop recommendations," Computer and Electronics in Agriculture, 2015.

[10] Wai Yan Nyein Naing and ZawHtike. Forecasting of Monthly Temperature Variations Using Random Forests. ARPN Journal of Engineering and Applied Sciences, 2015.

[11] Ramsundram N, Sathya S and Karthikeyan S. Comparison of Decision Tree Based Rainfall Prediction Model with Data Driven Model Considering Climatic Variables. Irrigation and Drainage Systems Engineering, 2016.

[12] Nasimul Hasan, Nayan Chandra Nath and Risul Islam Rasel. A Support Vector Regression Model for Forecasting Rainfall. EICT, 2015.

[13] Dhawal Hirani, Nitin Mishra - "A Survey on Rainfall Prediction Techniques", International Journal of Computer Application, 2016.

[14] Pinky Saikia Dutta, Hitesh Tahbilder "Prediction of Rainfall Using Data Mining Technique over Assam"- IJCSE, 2014.

[15] Ansif Arooj, Mohsin Riaz, Malik Naeem Akram : “Evaluation of predictive data mining algorithms in soil data classification for optimized crop recommendation”, 2018 International Conference on Advancements in Computational Sciences (ICACS), IEEE, 2018.