e-ISSN:2582-7219

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

Volume 5, Issue 12, December 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.54

International Journal Of Multidisciplinary Research In Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.54|

| Volume 5, Issue 12, December 2022 |

| DOI:10.15680/IJMRSET.2022.0512007 |

Integrated Cotton yield Forecaster

Mrs.P.Neela Sundari [1], B.Yadavi [2], G.Gnapika [3], Ch. Madhavi [4], N.Krishna Priya [5]

[1] Assistant Professor, Dept. of CSE, KKR & KSR Institute of Technology and Sciences, Guntur, A.P, India

[2, 3, 4, 5] B. Tech Final Year, Dept. of CSE, KKR & KSR Institute of Technology and Sciences, Guntur, A.P, India

ABSTRACT: Farming is the economic pillar of every country. Wherein India more than 65% population depends on agriculture. Unpredictable variations in climatic circumstances shows great impact on crop cultivation. Since the farmers are not encouraged to embrace automation in their activities, individuals are losing interest in the topic. It is mainly due to the absence of Information Technology participation in the agricultural sector. The above issue can be conquered by using a variety of programming techniques to assist farm owners in estimating the turnover of their production and motivating the farmers to stay in their field area. The prediction of cotton yield entails predicting the yield of the cotton from the present chronological information like climate, fertilizers, and soil variables. Machine learning algorithms helps to forecast yields based on locality, using land area, and based on fertilizers used. The proposed model aims to design, develop and implement the training model by using different input data for cotton yield prediction with good accuracy. And at the same time the crop loss because of diseases is approximately 10 to 40%. Farmers determine the diseases by their experience in farming but, it is not accurate and appropriate way. Sometimes farmers take opinion from experts to detect the diseases but this is also a time-consuming way. At the time of examination of crop damage, the inspection committee faces many issues about the identification of disease and actual percentage loss of crop due to diseases. So, our mission is to make a web portal where farmers can upload diseases of the plant which uses Image processing Technique and by identifying the disease present in their plant leaf we provide the best possible solution for that disease and also those who don't have basic knowledge of which pesticides to use or where to buy them. They could have been benefited by providing E-commerce web interface that connects pesticide manufacture directly to farmers thereby achieving complete expulsion of 'Middle-Man' concept. Therefore, this project can help farmers to determine and prevent crop yield loss.

KEYWORDS: Yielding, Prediction, Agriculture, Cotton, Fertilizers, Gradient Boosting Algorithm, Image Processing, Multi-Variate Regression, Machine Learning.

I. INTRODUCTION

1. ABOUT COTTON:

Cotton known as "White Gold" is one of the most commercially important and natural textile fibre crops and a notable contributor of oil seeds. Cotton is an immensely important crop for the sustainable economy of India and livelihood of the Indian cotton farming community. It is cultivating of about 333 lakh hectares across the world and in around 120.69 lakh hectares in the country. Thus, India accounts for around 45.3% of the global cotton area and contributes to 26% of the global cotton production. Cotton recommences to enjoy a leading and the most favoured fiber status among all the Indian textile mills, as the major raw material for the textile industry. Presently, nearly 63 million people depend on cotton agronomy, marketing, processing and exports for their livelihood. India is also the only country in the world that grows not only the four cultivated species of cotton

But, also their specific hybrids on a commercial scale. The textile industry, which consumes the cotton, as its principal raw material, contributes about 6% to the GDP and is the major exchange payee for the country. Hence, growth and development of cotton cultivation and cotton yield has a vital bearing on the overall development of the Indian economy.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.54|

| Volume 5, Issue 12, December 2022 |

| DOI:10.15680/IJMRSET.2022.0512007 |

2. PARAMETERS EFFECT YIELD:

The majority of every nation's economy depends on agriculture. However, unpredictable changes in climatic conditions have affected the cultivation of crops. There is a declining interest for individuals to remain in this field since they are not encouraged to use technology for processes. This happens primarily because there is a lack of involvement of information technology in the farming sector. This obstacle can be overcome by applying several techniques of programming to help farmers estimate the yield of their produce and motivate them to remain in this field. Crop yield prediction comprises fruitful yield of the crop from available historical information like climatic, average parameters, fertilizers and soil parameters. Machine learning is one of the technique that can be used to predict the yield of crop. Using parameters that are associated with the environment, agriculture and farming, machine learning models can predict the crop yield with good precision and accuracy. The research previously done in this sector has proven to be instrumental in forming the basis of this model. The findings have shown that parameters such as nutrient levels, acreage and historical data have a greater impact on the overall yield in addition to the climatic conditions of the place being observed. In this work, a model has been proposed which helps cultivators to predict the yield of the crop even before cultivating directly onto the agricultural lands. The yield prediction based on location, acreage and fertilizer data for several regions in the United States has been experimented using various algorithms. The parameters tested and the features used introduce a unique insight into the relationship between nitrogen, phosphorous and potash (N-P-K) used (acreage covered and quantity) and the lint yield of cotton

3. EFFECT DUE TO DISEASES:

Nearly 60 million people in India are directly or indirectly connected to production of cotton crop out of which 20% to 30% is lost due to various diseases. In Cotton crop cultivation we face many problems due to diseases which affects crop a lot and it is not possible to figure out it by naked eyes. Mostly, the leaf will get effected of a major portion. About 60 to 75% of diseases on the plant is on its leaves. So, our study of interest is focused mostly on leaf of the crop rather than whole cotton crop. The primary focus is to detect disease and estimate its stage for a cotton plant using images by cascaded classifiers. Additionally, the yield prediction results have shown that the gradient boosting regressor proved to be the best performer, returning a standalone accuracy score of 90.1% for test set and 98% for train set on being validated using the fivefold cross validation method. After validation, 82.5% is accuracy. The proposed model aims to design, develop and implement the training model by using different input data. The machine will be able to learn the features and extract the crop yield from the data by using data mining techniques.

II. LITERATURE SURVEY

The research papers help us to find the existing models and routed us to develop a new thesis by overcoming the problems which have been found out.

1. In 2018, Chlingaryan, A., Sukkarieh, S., & Whelan has done an audit which is predominantly focuses and talked about on AI methods, yield estimation and, accuracy nitrogen the board. The survey exhibits the technique of back proliferation significance and its precision of harvest yield expectation for various vegetation lists. They described that gaussian procedure are valuable for foreseeing and finding various qualities of plant leaves which will be utilized for crop yield expectation model
2. Shailesh Shetty S et al. By this it suggests farmers evaluating which crop to grow in a specific area at a specific time and predicting whether it will be profitable or not. It gives the specifics by specifying whether the crop is profitable. As a result, this aids farmers in their decision-making process, allowing them to save time.
3. In 2012, P. Revathi, et al. and Hemalatha had proposed an algorithm named Homogenous Pixel Counting Technique for Cotton Disease Detection (HPCCDD) for detecting spot diseases in cotton leaves using Edge detection methods. By this the image enhancement was carried out.
4. In 2019, Bhanumathi, S., Vineeth, M., & Rohit, N. International Conference on Communication and Signal Processing (ICCSP) (pp. 0769-0773). IEEE. says how to predict yield and how to efficiently use Fertilizers.
5. Suvidha Jambekar et al. proposed a predictive modelling tool to predict crop production. According to the results, Random Forest Regression may be used to accurately estimate wheat, and rice, and maize production.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.54|

| Volume 5, Issue 12, December 2022 |

| DOI:10.15680/IJMRSET.2022.0512007 |

6. In 2013, Huang et al., Y. Huang, R. Sui, S.J. Thomson, D.K. Fisher proposed a system that estimates cotton yield with varied irrigation and nitrogen treatments using aerial multispectral imagery.
7. In 2015, Rakesh Kumar, M. P. Singh, Prabhat Kumar, J.P. Singh proposed a crop selection method using some Machine Learning techniques which gives yield rate, but it doesn't gives accurate results.
8. In the year 2020, European Journal of Molecular & Clinical Medicine ISSN 2515- 8260 Volume 7, Issue 07, detected an issue which is followed by segmentation process, and for obtaining beneficial segments the texture statistics are computed. At last, classifier is used for the features that are extracted to classify the disease.
9. In the year 2018, "Rose Plant Disease Detection Using Image Processing" was published by Varsha Sawarkar, S. Kawathekar. In this paper the identification of the rose plant diseases is observed. In this they observed the key for preventing the losses in the yield and quantity of the agricultural product.
10. International Conference for Convergence in Technology [IEEE] [2018] presents Image processing techniques for plant disease detection. Here, images are captured and then they are realized to match the size of the image which is stored in the databases. This approach can peculiarly support disease detection. However, diseases are detected by measuring leaf area only. Which leads us to get less accurate results because of the lack of factors used for disease detection. This system also recommended the use of harmful pesticides which could harm the soil in the long run.

III. REVIEW FINDINGS

After analysing the existing approaches, we have identified some drawbacks in them. They are described as follows.

1. Starting from the early 2000's. Many new techniques are being used by farmers in cultivation of cotton crop. But not much changes have been observed in the yield rate.
2. The farmers from various countries have different methodologies and practices which are followed by them in farming, cultivation. Depending upon the climatic conditions of that particular country the methodologies may change.

The farmers from various countries have different methodologies and practices which are followed by them in farming, cultivation. Depending upon the climatic conditions of that particular country the methodologies may change.

3. Nowadays, automation in agriculture sector has increased more rapidly in terms of predicting yield, detecting of diseases etc., but they produce results merely.
4. A system was proposed that which determines cotton yield using some machine learning techniques like Decision Tress, Naïve Bayes Classifier, Multi-Layer Perceptron which gives a less accurate values.
5. In maximum of the projects or research papers, they are discussing about systems that predict and interpret cotton yield and its determinants under long-term conservation management practices using machine learning techniques. Even though their predictions are less accurate.
6. In research relating to identify and diagnose diseases, it says that Indian Farmers may have to travel 100 of miles for the expert's advice, which very time-consuming and expensive.
7. In Random Forest, Support Vector Machine (SVM) –

- Random Forest contains less accuracy compared to Support Vector Machine in crop yield prediction.
- It doesn't defect any crop diseases.

8. Traditionally expertise of the biotechnology identifies the diseases by observing the plants with bare eyes. In villages, farmers identify the symptoms of diseases by their experiences. But sometimes due to lack of knowledge, difference between practical experiences, the same disease categorized in different group and inaccurate diagnosis leads to erroneous control measurements.

IV. PROPOSED WORK WITH METHODOLOGY

A. Existing System:

At the beginning, the decision trees and support vector machines methods are used which are costly and long process and also time consuming. Hence there is a need for new technologies and methods to enhance the existing

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.54|

| Volume 5, Issue 12, December 2022 |

| DOI:10.15680/IJMRSET.2022.0512007 |

system in order to get better accuracy. The existing model is proposed for predicting the cotton yield based on some features of soil, temperature and rainfall. The model does not consider attributes like land and fertilizers in predicting the crop. In the same way in existing systems diseases are predicted manually which may or may not be accurate.

B. Proposed Method:

Using various input data, this system intends to deploy, design, and develop a training model. Using the Data mining and Data Science approaches, the machine will be able to acquire the knowledge of the features and

bring out the cotton production from the details of data. Input the parameters like crop, yield, and area by using Data Mining technique. Pre-processing of data with noise removal are obtained. Collection of data can be made by setup and arranging in usual form. Data collection classifies into NPK used, crop data and area count. Yield of the particular seed or crop is obtained. Based on the amount of yield the partition of data can be made with 80% of the train set and 20% of test set. In the next step classification to the data can be performed by using Gradient Boosting Algorithm and Multi Variate Regression. Validating the coefficient determinate with Multiple R squared and Adjusted R squared parameters. Finally, the result can be obtained. And also, disease is also detected by scanning the image of disease. And checks whether the image uploaded by farmer is present in the database or not. If found detects the disease.

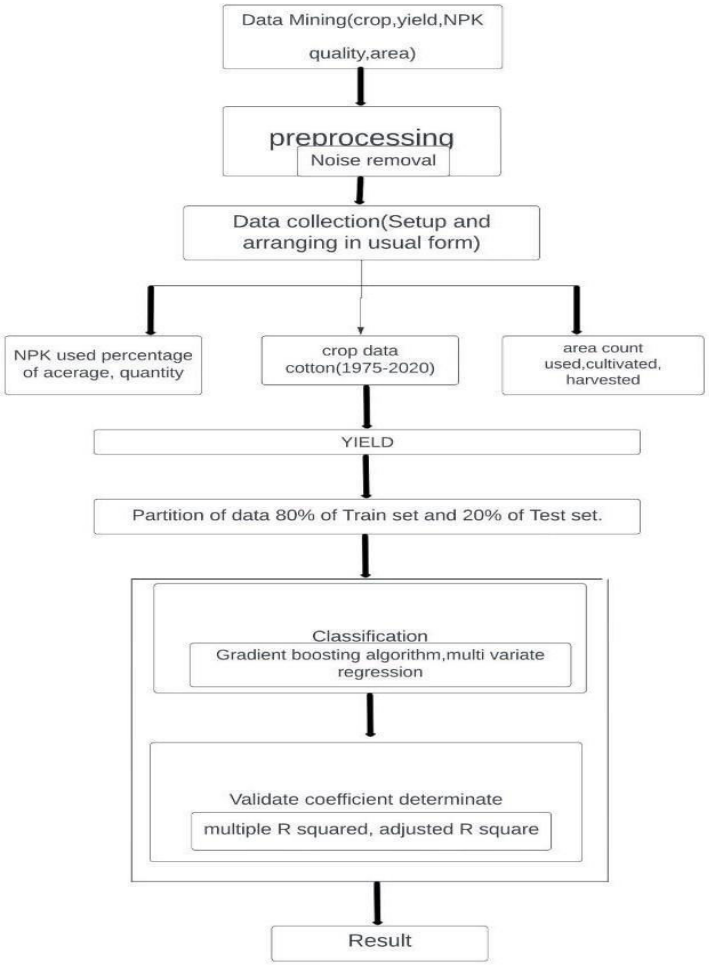

fig: SYSTEM ARCHITECTURE

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.54|

| Volume 5, Issue 12, December 2022 |

| DOI:10.15680/IJMRSET.2022.0512007 |

V.COMPARISON WITH EXISTING SYSTEM

1. One of the existing models is to automatically grade the disease on plant leaves. According to this system, plant experts and pathologists mainly gave observations on basis of naked eye prediction of disease from symptoms that are visible and gave certain scale grade to the disease. But this manual grading is not only time consuming but also not feasible.
2. Another existing model is usage of the decision trees and support vector machines methods which are costly and long process and also time consuming. Hence there is a need for new technologies and methods to enhance the existing system in order to get better accuracy. The existing model is for predicting the cotton yield based on some features of soil, temperature and rainfall. But this model does not consider attributes like land and fertilizers in predicting the crop. So, it won't give accurate results.

The existing systems tried to predict yield based on external features using some traditional methods, and by using some machine learning algorithms. They also produce output but some are time consuming and some are not producing accurate results.

1. The existing system do not provide accurate results, which are not useful for yield prediction.
2. The disease detection through naked eye does not give proper results. It needs continuous monitoring. Its time consuming and worthless.

VI. CONCLUSION

Agriculture plays a significant role in the growth of the national economy. India is a nation in which agriculture plays a primary role. Cotton is one of the major domains in agriculture which decides economy of India. So, our system maximizes the productivity of crop and it improves the crop yield. Thus, our work would help farmers. We aim to develop an integrated cotton yield system where we can predict the yield of cotton by entering the past data of several external factors like land, pH, NPK values, fertility etc., of their plot. For the prediction of cotton yield we are using Machine Learning techniques like Gradient Boosting Algorithm and Multi-variate Regression in which Gradient Boosting Algorithm has the highest accuracy value. The existing model exhibits Naive Bayes and Decision tree techniques which show a lower accuracy rate. Hence our system provides the result with high accuracy rate. It is automation for cotton yield prediction and disease detection in an efficient and economically faster way. Sometimes the farmer's does not have appropriate knowledge, then misidentification of disease can be possible and incorrect control measure like non-affecting pesticides can be used which leads to wasting of work, money and serious problem to crops. By using the proposed system, farmers can easily identify the disease on cotton plant and they can apply the correct control measure immediately. Also, this small-scale E-commerce system could help decrease in pricing of pesticides and helps farmers getting more profit which surely lead to decrease in the suicide rates between them. Hence the system maximizes the productivity of crop and it improves the crop yield. I've concluded that crop prediction is accurate and reliable. The study could be expanded to include the implementation of a smart irrigation system for farms in order to increase yield.

REFERENCES

1. Jamuna KS, Karpagavalli S, Vijaya MS, Revathi P, Gokilavani S, Madhiya E, "Classification of Seed Cotton Yield based on the Growth Stages of Cotton Crop Using Machine Learning Techniques," 2010.
2. "Crop Selection Method to Maximize Crop Yield Rate Using Machine Learning Technique," in 2015. International Conference on Smart Technologies and Management for Computing, Rakesh Kumar, M.P. Singh, Prabhat Kumar, J.P. Singh.
3. Mr. Santosh Mahagaonkar, Devdatta A "Machine Learning Algorithms for Crop Yield Prediction and Fertilizer Recommendation" -From the International Journal of Engineering Applied Sciences and Technology, published in 2019.
4. Mayank, Champaneri, Chaitanya Chandvidkar, Darpan Chachpara, Mansing Rathod "Crop Yield Prediction Using Machine Learning" -April 2020.
5. Thomas van Klompenburga, Ayalew Kassahuna, Cagatay Catalb," Crop yield prediction using machine learning: A systematic literature review"-August 2020.
6. D. Jayanarayana Reddy, Dr. M. Rudra Kumar "Crop Yield Prediction using Machine Learning Algorithm"- Proceedings of the Fifth International Conference on Intelligent Computing and Control Systems in August 2021

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.54|

| Volume 5, Issue 12, December 2022 |

| DOI:10.15680/IJMRSET.2022.0512007 |

7. N.R. Prasad, Patel N.R, Abhishek Danodia "Crop Yield Prediction in Cotton for Regional Level Using Random Forest Approach" - Spatial Information Research in July 2020.
8. Mamunur Rashid, Bifta sama Bari," A comprehensive review of crop yield prediction using Machine Learning Approaches with special Emphasis on palm oil yield prediction" - Digital Object Identifier in 2021.
9. Weicheng Xu, Pengchao Chen "Cotton yield estimation model based on machine learning using time series UAV remote sensoring data" - August 2021.

Impact Factor
7.54

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com