Literature Review Crop Yield Estimation Based On Soil Fertility Using Machine Learning Algorithm.

Smt. Pallavi C V Assistant Professor Department of Computer Science BNM Institute of technology Bengaluru, India pallavicv@bnmit.in Anil Kumar P N Department of Computer Science BNM Institute of technology Bengaluru, India anilkumarpn50@gmail.com Gamya T R Department of Computer Science BNM Institute of technology Bengaluru, India gamyathalanki21@gmail.com

Roshini R Department of Computer Science BNM Institute of technology Bengaluru, India roshinimargi7@gmail.com

Abstract— Agriculture is the backbone of India, where it plays a predominant role in the economic growth and it also contributes significantly to a country's GDP (Gross Domestic Production) of the country. As time passes, there are number of revelations that take place in order to improve agriculture field. The major drawback or the issues farmers facing in crop productivity is that the famers fail to select appropriate soil type for the selected crop. To improve or overcome the problems or issues faced by famers AI (Artifical Intelligence) system are helping to improve the overall harvest quantity, quality known as precision agriculture. In order to improve the productivity of the agriculture AI systems has provided many machine learning techniques. Though right crop selection is the main key for yield by analysing the soil type and considering some factors which include humidity, temperature and rainfall. Sometimes the lack of knowledge about current climatic changes and soil fertility and crop selection also will be reason for failure of crop. The framers having primitive knowledge about conventional farming are also facing some serious issues in decision of crop selection. The selection of same crop in every seasonal farming makes the soil to lose its fertility. This study is amid at providing an efficient and accurate system using machine learning models and sensors that accurately provide or help farmers to select crop and get maximum yield. Correct selection of crop and fertility of soil is the important factor in the field of agricultural arena. We propose a machine learning model which is based on data of soil fertility which includesNitrogen(N), Phosphours(P) and Potassium(K) (NPK) values ,rainfall and humidity. All these data will be analyzed and will be trained with various suitable machine learning algorithm. The system comes with a model which will be precise and accurate in predicting crop and soil pH value and deliver it to the end user with proper recommendations about required fertilizer for the crop and land which

Sahithya G K Department of Computer Science BNM Institute of technology Bengaluru, India sahithyagk9@gmail.com

enhance framers to grow crop and get good yield. For the efficient use of this model we develop a user friendly application which helps framers to fetch the real time data of his particular land. The real time data such as rainfall, humidity of the soil, pH value and nutrient content. These data will be stored in database and then gather in cloud storage and with the appropriate operations, it enables framers to optimize framing strategies an help the create a trend analysis. The proposed system works on ensembling technique where bagging and boosting is implemented to predict the crop.

KEYWORDS: Agriculture, NPK, pH value, ensemble technique

I. Introduction

Agriculture is the science, art, and practice of cultivating plants and livestock. Agriculture was the key development in the rise of sedentary human civilization, where the farming of domesticated species created food surpluses that enabled people to live in cities. The history of agriculture began thousands of years ago. The cultivation of plants and food grains began years ago in order to provide food to the city's population. Agriculture is the main source of livelihood; it provides a source for the people to earn. Most of the population in rural areas is dependent on agriculture as their main source of income. Most of the population in rural areas is dependent on agriculture as their main source of income. Agriculture contributes significantly to a country's GDP, which is its gross domestic product. With the passing of time, there have been a number of revolutions that have taken place in order to improve agriculture throughout the world or a country.

Machine learning (ML) approaches are used in many fields, ranging from supermarkets to evaluate the behavior of customers. Machine learning has also been used in agriculture for several years. Crop Yield prediction is one of the most challenging problems in agriculture, and Many models have been proposed and validated so far. This problem requires the use of multiple datasets because crop yield depends on many different factors, such as climate, weather, soil, use of fertilizer, and seed variety. This indicates that crop yields. Prediction is not a trivial task; instead, it consists of several complicated steps. Nowadays, crop yield prediction models can reasonably estimate the actual yield, but a better performance in yield prediction is still desirable. Machine learning, which is a branch of artificial intelligence (AI) Focusing on learning is a practical approach which can provide a more precise yield prediction based on several features. Machine learning (ML) can determine patterns, correlations and discover knowledge from other datasets.

The models need to be trained using datasets where the outcomes are represented based on past experience. The A predictive model is built using several features, and as such, parameters of the Models are determined using historical data during the training phase in order to test phase, some of the historical data that has. It is used for purposes other than training for performance evaluation purposes.

An ML model can be descriptive depending on the research based on any problem and research questions. While Descriptive models are used to gain knowledge from the collected data, make predictions, and explain what has happened. Models are used to make predictions for the future. ML studies consist of different challenges when aiming to build a highperformance predictive model. It is crucial to select the right algorithms to solve the problem at hand, and in addition, the algorithms and the underlying platforms need to be capable of handling the data. Agriculture is considered the science of plant cultivation, which was crucial for the well-being of humans According to the insights of 2016, more than 2 million. As a result, we must assist cultivators in understanding th significance of earlier yield forecasts so that they can thrive on the fundamentals information about soil quality, interpreting the insightful climate requirements of particular area in order to accomplish high harvest yield through our innovation arrangement. In automated soil testing, Human efforts will be reduced by monitoring the quality of the soil using NPK sensors. Since ancient times, agriculture has been the backbone of our country.

Nowadays, climatic conditions very often. So, it is hard to grow crops by understanding weather conditions. We need to use some technology to find or understand he crop details and guide the farmers to grow crops accordingly and Moreover, fertilizer is also one of the major factors in growing crops.

If fertilizer is used more or less in the field, the soil may lose fertility, and the crop may fail to produce the expected yield. So fertilizer is also used becomes a major factor in it. Mostly, understanding the temperature conditions is critical in India because we can't improve the Indian economy with the help of crop prediction because it plays a major role in the Indian economy. Generally, Machine learning algorithms will predict the most efficient output of the yield. Previously, the yield was predicted based on the farmer's prior experience, but this is no longer the case. The current weather conditions may change drastically, so they cannot guess the yield. Nowadays, modern people do not have awareness about the cultivation of the crops at the appropriate time place Because of these cultivating techniques, the right seasonal condition is also being changed, against the fundamental assists like soil, water, and air, which leads insecurity.

II. Literature Survey

A literature survey in a project report is that section which shows the various analyses and research made in the field of your interest and the results already published taking into account the various parameters of the project and the extent of the project. A Literature survey refers to getting the content from the books which are related to the topicor a given project. It should be referred from some research paper that is related to the topic.

Most of the existing system and articles for precision a smart agriculture are discussed about the crop maximization using ML algorithm for crop yield with suitable parameters available in the current location and the type of the soil. Some of the articles are projected in this section to show the advantages of the existing system and where the enhancement are needed to increase the crop yield using soil fertility.

Shakib Mahmud Dipto, Asif Ifekher [5] proposed a model Which plays a vital role in our agriculture sectors. In this model they demonstrate crop suggesting system based on N.P.K values by using machine learning which will determine the best cropin a particular soil type. Algorithm used in this model are SVM, random forest and logistic regression and archived a maximum accuracy using SVM. Random forests for global and regional crop yield prediction this journal generated outputs show that RF is an effective and adaptable machine-learning method for crop yield predictions at regional and global scales for its high accuracy and precision, ease of use, and utility in data analysis. Random Forest is the most efficient strategy and it outperforms multiple linear regression (MLR).

Nidhi H kulkarni,Dr.G.N srinivasan Dr. B M Sagar and Dr.N K Cauvery[1] proposed a ensembling technique to build a model which combines the predictions of multiple machine learning models together to recommend the right crop based on soil specific of type with high accuracy. The independent base learners used in theensemble model are random forest,linear SVM,Naive Bayes. Each classifier provides its own set of class labels with acceptable accuracy. The class labels of individual base learners are combined using the majority voting technique. The recommendable crop type in this model is karif and rabi crop.The collected data is initially subjected to prepossessing. The datasets is divided into training set and test samples. The proposed work provides a helping Hand to the framer in the accurate selection of the crop fro cultivation.

Crop yield prediction and efficient use of fertilizers as in a research paper carried out by S.Bhanumathi, M.Ninetieth and N.Rohit[2] analyze the various related attributes like location, pH value from which alkalinity of the soil is determined. Along with it, percentage of nutrients like Nitrogen, Phosphorous and Potassium locations isused along with the use of third party applications like API s for weather and temperature, type of soil, nutrient value of the soil can be determined. increase framer revenue. Algorithm used are Random forest, Back Propagation. In this proposed model random forest has got less error rate. The system comes with the model to be precise and accurate in predicating about required fertilizer ratio based on atmospheric and soil parameters of the land which enhance to increase the crop yield.

Meeradevi and Monica R Mundane [7] proposed a model uses only few sensors for experimental use, To overcome the flaws of traditional agriculture wireless sensor network based precision agriculture and prediction crop yield is done these nodes sense the environmental parameters like temperature, humidity, pH, NPK values. The system is simple and reliable. Techniques used in this model is Linear regression and sensors.

Mr.saiRamesh, T.Sindhu, K.Sekvakumar, V.Sathiyavat hi,R.Reshma [3] proposed a system. In which system soil monitoring is combined with IOT technology aids in the enhancement of agriculture by increasing yield through gauging the exact soil characteristics such as moisture, temperature, humidity, PH and nutrition content. IOT system consists of pH-sensors, humidity and temperature sensors, soil moisture sensors, NPK probes, micro controller equipped with WiFi and cloud server. For recommending system, the SVM and decision tree algorithm is proposed to get the crop for given data helps to enhance the growth. Classification of soil and crop suggestion using machine learning technique as in a research carried out by Mrs N. Saranya[4] a comparative analysis of machine learning techniques has been carried out. In this paper they have proposed a method for classifying the soil according to macro nutrients and micro nutrients and predicting the type of crop that can be cultivated in the particular soil type. Several type of machine learning algorithms are used such as KNN, bagged tree, SVM and logistic regression.

In this proposed model SVM has obtained the maximum accuracy. Amna Ikram ,Wagar Aslam[6] published a journal based on crop maximization. They proposed a ML model. based smart crop data selection(SCS) which is based on of meteorological and soil factors. These factors include nitrogen, phosphorous, potassium, CO2, PH, EC, temperature, humidity of soil and rainfall. Existing IOT models are not efficient compared to this model. In this model real time sensory data is sent to cloud for analysis. Its results are visualized in android app. SCS ensembles the following five ML algorithm to increase performance and accuracy. Decision tree, SVM, random forest, and Gaussian Navie Bayes techniques are used.

III. Summary of literature survey

To design an application where we compare the different machine learning to predict the crop yield. We build a new decision system using machine learning algorithm.



Fig 1 Architecture for soil behavior analysis and crop recommendation

The architecture for soil behavior analysis and crop recommended. Figure 1 shows the analysis of the soil dataset which integrates all modules. The preprocessing stage consists of three major processes namely data cleaning and missing data computation. System architecture give a detailed flow of each stages which defines the structure , behavior, and other which helps the developers to understand and construct the model. The algorithm implementation will be done when preprocessing of data is completed. The algorithm used in this proposed system will be used for calculating or used to find the accuracy of the specified dataset in the cloud. Then the model creation will be processed using from the given output of the algorithm. These data will be stored in cloud storage and user can access these data whenever he requires. These data will be displayed to the user with the help of web application. Finally, a report can be generated for the crop productivity along with the accuracy of each ML model. The accuracy which is near to 1 is considered as an ideal model and the model which has accuracy near to 0 are considered as unideal model.

There are a lot of machine learning algorithms used for predicting the crop yield. This experiment can be carried out in five phases. In the first phase, XG boost technique has been implemented to train the model for crop yield prediction. The second, third, fourth, and fifth phase deals with implementations of decision tree, ada boost and Random Forest (RF) respectively for training crop yield classifier. In each phase, data (with selected features) has been divided into two sets: training set and test set in the ratio of 70:30; thus training set has 3500 samples and test set 2500 samples. Each trained model has been validated with test data in order to assess its performance. The performance of all the trained models has been measured using performance measures namely accuracy, recall, precision, specify and F-score

IV.Conclusion

Crop yield prediction system enables framers to grow crop and get good yield of it. This technology ensures crop yield but major contribution is to grow a particular or a suitable crop by measuring soil pH values. The paper presented a survey on crop yield prediction and other existing methodology of the system. The paper also discussed the contribution of the proposed method and its effects on solving the crop failures. Finally, this paper proposes a crop yield prediction using ensembling technique for solving existing problem in agriculture field. This study can further be extended by implementing a hybrid model along with NPK sensors and benchmarking it with other models.

V. References

[1] Nidhi H kulkarni, Dr.G.N Srinivasan, Dr. B M Sagar and Dr.N K Cauvery, "Improving Crop Productivity through a Crop Recommendation System using Ensembling Technique", IEEE International Conference System and Information Technology for Sustainable Solutions, 2018. [2] S.Bhanumathi, M.Ninetieth and N.Rohit, "Crop Yield Prediction and Efficient use of Fertilizer", International Conference on Communication and signal Processing, April 4-6, 2019, India.

[3] Sai Ramesh, T.Sindhu, K.Sekvakumar, V.Sathiyavathi, R.Reshma, "IOT based Classification Techniques for Soil Content Analysis and Crop Yield Predication" Proceedings of the Fourth International Conference on I-SMAC 2020.

[4] N. Saranya "Classification of Soil and Crop Suggestion using Machine Learning Techniques", International Journal of Engineering Research and Technology(IJERT), issue 02, Feburary 2020.

[5] Shakib Mahmud Dipto ,Asif Ifekher, "Suitable Crop Suggesting System Based on N.P.K values Using Machine Learning Models" International Conference Paper. 2021.

[6] Amna Ikram ,Waqar Aslam, "Crop Yield maximization Using an IOT-Based Smart Decision", Hindawi Journal of Sensors, May 17,2022.

[7] Meeradevi and Monica, "krashigyan: A framer support sytem", Asian Journal for Convergence in Technology(AJCT), 2021.