Prediction of Crop Yield Based on Weather Using Machine Learning Techniques

[1]Kavya Shree S, [2]Dr.Prabha R

[1]Student, MTech,[2]Professor,

[1,2] Department of Computer Science,

[1,2]Dr. Ambedkar Institute of Technology, Bangalore, Karnataka, India

Abstract -The impact of climate change in India, most of the agricultural crops are being badly affected in terms of their performance over a period of the last two decades. Predicting the crop yield in advance of its harvest would help the policy makers and farmers for taking appropriate measures for marketing and storage. This project will help the farmers to know the yield of their crop before cultivating onto the agricultural field and thus help them to make the appropriate decisions. It attempts to solve the issue by building a prototype of an interactive prediction system. Random forest algorithm is used. By analysing all these issues and problems like weather, temperature, humidity, rainfall, moisture, there is no proper solution and technologies to overcome the situation faced by us. In India, there are many ways to increase the economic growth in the field of agriculture. Data mining is also useful for predicting crop yield production. Generally, data mining is the process of analysing data from various viewpoint and summarizing it into important information. Random forest is the most popular and powerful supervised machine learning algorithm capable of performing both classification and regression tasks, that operate by constructing a multitude of decision trees during training time and generating output of the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees.

Index Terms - Agriculture, crop-prediction, Machine Learning, Crop yield.

I. INTRODUCTION

Aanalysis of data is the method of creating data to find important information and results. It is process of analysis, extracting and predicting important information from many data to determine a unique design. Farmers use this method to collect raw data about their customers to gain useful information. This analysis can also be used in agriculture. Farmers rely on years of experience in planting, hoping to make more money the next time they harvest, but still fail to get the harvest. This is often due to poor water quality or poor crop selection, and crop yields are sometimes lower than expected. Agronomists say nearby is a need for better ways to predict and improve crop growth. Agricultural research often focuses on identifying biological processes that grow crops and increase yields. Planting mainly depends on limitations such as crop, seed type and environmental limitations such as sunlight (temperature), soil (ph), water (ph), precipitation and humidity. Through analysing the soil and climate of a region, yield can be maximized and yield can be estimated. This estimate will help farmers. Choose the right product for your farm based on soil type, temperature, humidity, water level, row spacing, soil pH, season, fertilizer and month. Predicting crop production stays a tough task as it is pretentious by countless factors, including crop resources, soil, climate, agricultural practices (planting date, irrigation and fertilization rate, etc.,) and biotic stress. Many types of product prediction have been developed, such as statistical, agrometeorological, empirical, biophysical and mechanical methods. A newly developed weather instrument and various techniques are used as input. To get a reasonable estimate of our actual results(called targets) and forecasting models, use the regression plus the mean is used for error rate, dedication coefficient were evaluated. Considering the validity of crop prices and output estimation, this research identifies and uses various estimation methods to calculate crop yield in different fields. Data analysis is the process of analyzing, cleaning and modelling data to find useful information and results. The process of analyzing, extracting and predicting significant data from big information to infer certain patterns of data is analysed. as soon as the drill data is not labeled or categorized. Unsupervised learning examines how computers might infer that functions describes hidden structures. The With labelled examples, you can apply the past lessons to new data. Afterwards necessary training, the classification canprovide targets for new inputs. Smart farming is the way information is sent from traditional farmers to knowledgeable farmers. An estimate of the general system act for several crop yields in a given situation is obtained by considering various methods and new air production methods. The regression and coefficient of determination analysis and mean error made a comparison between the farmer-friendly interfaces describing the crops built on the available data and the actual results, called the target and forecast model. Use dissimilar information mining techniques to predict crop yields to maximize yield, and planning for work because of the importance of predicting the crops in the field, the determination of using things in the yield

II. LITERATURE SURVEY

Feilong wang. et al [1] Timely and precise estimation of crop information is relevant to the national economy and human health. Today's research also provides a reason for major revisions of shipping estimates, but is often limited by locations and facts. UAV perform invisible reconnaissance with ultra-sensitive sensors that can capture high-resolution and extremely dangerous images when requested. Generally, time series vegetable data (VI) stands to measure crop yield. Therefore, comparison of display names can have external properties that can distress the accuracy of waste disposal data. Then, in this perspective, a relative vegetation and relative abdication list was prepared, then used in the direction of measure grain abandonment on pixel scale. Also, the best development phase will be determined by crop production estimates. The different growth stages are analyzed at this time for yield forecasting Finally, the differentiation of developmental phases was attempted to get the best combination of estimation agreement, has been validated on a pixel scale with the measurement results. The results show that relative vegetation information is suitable for predicting abdication on pixel scale.

Anil Suat Terliksiz. et al [2] Because of the expanding global population, more attractive infrastructure is required. Deep learning has been crucial in the usage of analysis, classification correction, and pruning hope as deep convolutional neural networks (CNNs) for image categorization have advanced. Always take into account access parameter-based abandonment estimation techniques in adding to traditional machine learning methods like inverse vector machines and selection trees. Two afresh proposed deep neural network for predictive modelling are CNN and long short-term memory. The primary goal of this work was to monitor soybean crops in Lauderdale Country, Alabama, using 3D CNNs to identify spatiotemporal features.

Mr . A. Suresh. et al [3] Prediction of major crop yields of Tamil-Nadu using K-means and modified KNN, More than 40% of the people in this state make their living mostly from agriculture. Researchers from the food and agriculture organization predict that the global population will rise by a third between 2010 and 2050. Compared to current crop production, the request for agricultural production will rise by 60%. Therefore, projection is crucial in determing the need for crop production in order to maximize yield. In order to do so, we provide in this study a forecast technique based on k-means and modified k nearest neighbour for the key crops of tamil-nadu. The tool for clustering and classification is matlabWEKA, respectively. The numerical outcome demonstrates that our strategy is superior than the conventional data mining approach.

Niketa Gandhi. et al [4] The popular of India's food production is reliant on on grains like rice, wheat, and other crops. The supporting capacity and efficiency of the planted area depend on suitable conditions. Uncertainty in the traditional climate can have a negative impact, and as result, the dry season reduces energy production. Making better agronomic and pruning decisions can benefit farmers and other stakeholders. This can be done by establishing the best programs to anticipate pruning efficiency in various climates. In various weather scenarios, pruning cancellation estimates can stay improved using machine learning approaches. This article provides a review of the application of this type of machine learning to the editorial variety of Indian rice. This article discusses the search outcomes obtained by applying the SMO classifier using WEKA source on the data of 27 districts in Maharashtra, India. Data used to predict rice abdication from the official database of the Indian government. The parameters considered for the study were precipitation, minimum temperature, average temperature, maximum temperature and reference crop evapotranspiration, area production and yield for the Kharif season (June to November). For the present study, the mean absolute error (MAE), root mean squared error (RMSE), relative absolute error (RAE) and root relative squared error (RRSE) were calculated. The experimental results showed that the performance of other techniques on the same dataset was much better compared to SMO.

Xinran Gao. et al [5] To circumvent the effects of an unfavorable climate and to forecast the degradation of trim abdication or quality owing to an unsightly collect plan, optimizing collecting plans demands a strategy for development date forecasting. However, as the majority of expectation models are statistically based and therefore non appropriate for territorial application, consistency is required for representations based on remote sensing. We demonstrated a system that integrates leaf region record data from the direct determination imaging spectral radiometer (MODIS) into the world Nutrient studies (WOFOST) edit development show. Figure atmospheric data from the THORPEX intelligently amazing worldwide ensemble (TIGGE) was used as climate information contribution for long-term periods.

Carl Menges. et al [6] Presents the output of an experiment carried out to relate the yield from various crops to TerraSAR-X dual polarimetric imagery. X-band wavelength has higher sensitivity to smaller crop structures, especially stem and head density making it suitable for relating yield to backscatter. The coherent dual-polarimetric mode of TerraSAR-X was also used to emphasize the volume scattering through dual-polarimetric entropy/alpha decomposition. Good correlations to yield data as gathered by harvester telemetry were obtained.

Zongnan Li. et al [7] Monitoring crop growth is important for forecasting, examine various indicators of harvest growth via remote sensing at different locations. The test was conducted at an investigated location in Hebei area within the North China plain. The aim edit of this thinking is winter wheat, one of the staple nourishments of India. The canopy-scale considered is based on a field tried with distinctive treatments of winter wheat to attain diverse yields. Canopy unearthly and LAI information for diverse crops was together and evaluated. Calculate the relationship coefficient between the vegetation record and the LAI. Visual development observing pointers for each phonological organize of the index were decided by the relationship coefficient. The most excellent pointer for the early prolongation stage at the joint level, the finest pointer. During head and breast broadening, the leading marker is NDVI. Utilize HJ-IA lackey multispectral farther detecting information to look at editing development checking within the shown region. The adequacy of vegetation record at distinctive spatial scales in observing edit development at distinctive phonological stages, by comparing the coefficient of vegetation list and LAI information. It that the vegetation file with the impression of making strides soil ripeness, is appropriate for regions where diverse crops and uncommon crops co-exist at the introductory sharing level, but has confinement at the boot arrange and in zones with thick crops.

Challinor.A. et al [8] According to the intergovernmental panel on climate change (IPCC), that climate change should reduce agricultural energy production. In sub-Saharan Africa and (to a lesser extent) South Asia, limited access to information and collective participation have limited agricultural exploration and progress. Here we look at the importance of the link between horticulture and climate and examine the importance of climate data for agricultural assessment in three steps. First, we review climate and the use of climate data in rural studies. Using the metadata from the survey, we found that rural workers preferred weather information (50.4% of data collected), how much data. It used in agriculture by wind turbines. We used a correlation scheme to assess the impact of weather stations on missing data and found that due to step terrain (the effect of missing data on precipitation and snow is measured at +1300mm/year+-3^{0}C when talking about temperature. Finally, we compare the climate indices and annual instability indices from the IPCC Fourth Assessment Report (4AR)climate models with specific observational data. Climate models are seen as inadequate for agricultural research in west Africa and south asia because they do not seem to reflect climate change, a view shared by more than 50% of national demonstration projects.

Martin. et al [9] Crop Syst, a crop simulation model, shows crop production and water/nitrogen budgets, In agriculture, nitrogen and water are two essential factor for controlling growth. However, the climate, soil, hydrology, edit characteristics, trim revolution, and administration cannot be used to characterize their administration in isolation. This page provides preliminary approval for these components and illustrates water, nitrogen, and trim development characteristics of coordinated edited renactment show. The crop's water use is accurately represented by the water budget. The expected nitrogen levels across the soil outline don't match the estimates from the filtering researches, but then again they are trustworthy when used with in general information designs. Through a total of 77 information focuses, the various indicators seemed to have a clear comprehension of the replicated and observed values for biomass and the abdication from maize, winter wheat, and spring crops. Trim system is promising as a tool for water analysis and nitrogen administration based on this initial approval. Long-term efforts should be made to exhaust advance approval of show components across various crops and situations

III. METHODOLOGY

(1)Dataset Collection

We currently gather data since many bases and create a database. Additionally, the offered dataset employs analysis (diagnostic and descriptive). Many online abstract sources are available, including data.gov.in and indiastat.og. the annual overview of a harvest determination be used for at-least ten years. These data sets typically permit non-governmental time series to behave as they do. This combines the essential abstract with the fundamental abstract. Forecasting harvest yields on a universal and regional ruler using random forests.

(2)Data Partitioning

The complete set of data is split into two sections: 25% is utilized for model testing and 75% is castoff for model training.

(3)To predict Machine Learning

Only Algorithms for supervised machine learning: using tagged examples, you can apply previous knowledge to new data. The system can give new inputs targets after enough training. Learning algorithms can also discriminate results from the planned and proper output and spot faults, allowing the model to be adjusted accordingly. Contrastingly, unsupervised machine learning methods are utilized as soon as the drill data is not labelled or categorized. Unsupervised learning examines how computers might infer from the unlabelled of data.

(3)Random Forest Classifier

Random forest is a prevalent and influential machine learning supervised algorithm that achieves sorting and deterioration responsibilities by constructing multiple decision trees the time of training, generating class outputs, and producing the class method(classification) or average prediction (regression) of a single tree. The supplementary plants or trees in the forest the stronger the prediction.

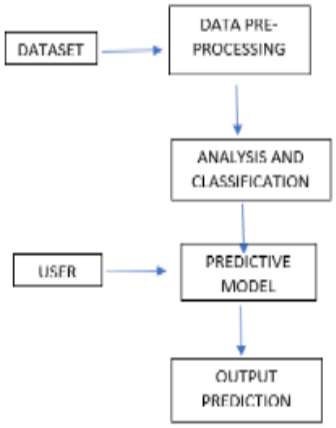

Fig.1 Methodology of Harvest

IV. SYSTEM ARCHITECTURE

Architectural diagrams can help system designers and developers visualize the global high- level structure of their systems or applications in order towards certify that the organization meets the needs of their users. They can also be used to describe the patterns castoff in the entire design. It's somewhat like a blueprint that can be used as a guide to facilitate discussions, improvementsand cooperation among teams. A system architecture or an engineering framework can be a conceptual framework that describes the assembly and behavior of the framework. They have principles and relations that describe how they work together to completes the whole process in the system.

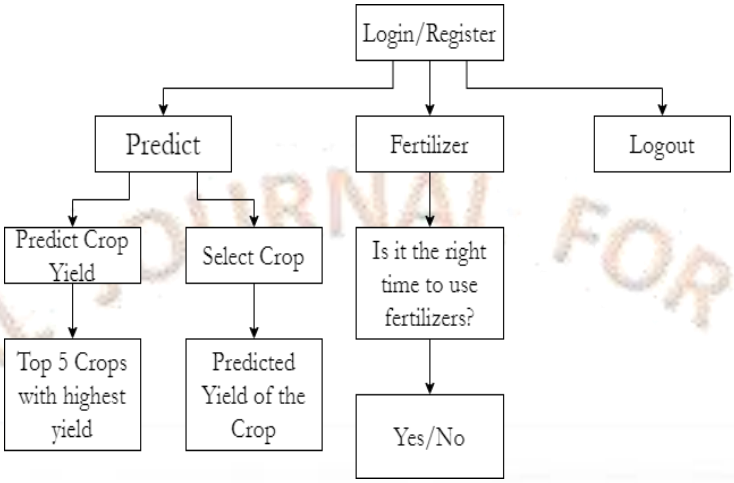

Fig.2 System Architecture

Above fig shows the system architecture: First phase is to register for the application or login. In the next step, you have three options: Predict, Fertilizer and Logout. The operator can select one of the three options and continue. In Predict, the system provides two options, depending on whether the user knows what to plant and whether he has not yet decided on thecrops. In both cases, the input is taken since the user besides the predicted value is given to the user. When selecting a fertilizer module, the user receives a pop-up message indicating whetheror not he can use the fertilizer and whether it will rain for the next 15 days. The last is the logout that deletes the user and brings him back to the login/registration page.

V. RESULT AND DISCUSSION

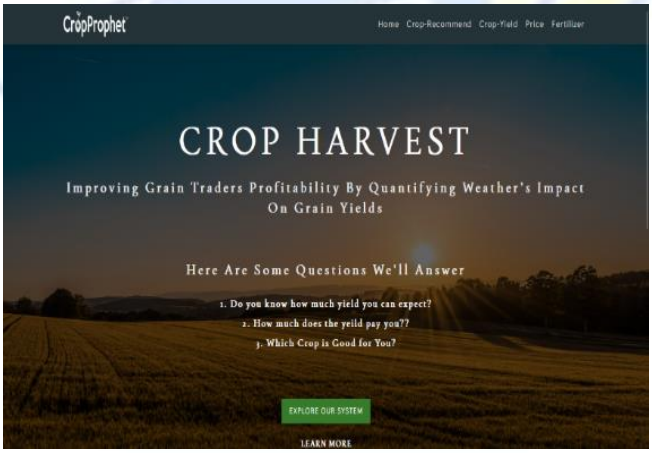

Fig.3 Home Page

The above fig3 shows the homepage of the website where the person accessing the website, enters the details such as the crop-recommend, crop-yield, price, fertilizer.

Fig.4 Crop Yield Prediction Page

The above fig4 shows where the person enters the required details such as Year, season, crop, temperature, humidity, area, state , district of the soil to get a suitable recommended crop to grow in a farm.

Fig.5 Crop Fertilizer Prediction

The above fig5 shows where the person enters the required details such as temperature, humidity in %, moisture, soil type, crop type, nitrogen, potassium, phosphorous and by clicking on submit the result is printed of the crop fertilizer of the soil to get a suitable crop to grow in a farm.

VI. CONCLUSIONS

We have proposed an approach of predicting the crop, crop-yield and crop-price using RFA random forest and back propagation algorithm. To improve agribusiness, this approach reduces farmers losses and improves yields. In this phase of the project, we reviewed the literature on price prediction and crop-yield prediction. This literature helps us understand the challenged that we face in the price dataset to identify the price and yield. According to the literature experimental results, we have come to know that random forest and back propagation helps in crop-price identification and helps to eradicate the challenges that we have inserted in dataset that comes popular the dataset and also the accuracy rate of this algorithm is far better than other algorithm. Through integrating this through extra departments like sericulture and other development initiatives on the village level, farmers are able to gain a deeper understanding of agricultural ecosystems. By incorporating this by additional departments such as sericulture, this can be enhanced in instruction to develop villages for the benefit of farmers. As we know farming is the backbone of a country like India, however, agricultural technology is too important to exclude agriculture. This article provides a system to help farmers estimate harvest yield founded on weather conditions and crop area. Using the system, farmers can decide to plant a particular crop or plant another crop if the forecast is bad. This research project can be occupied to a new level. We can make an agreement for agriculture and distribute it to the agriculturalists. Farmers can use it to decide which crops to plant in which season and thus earn more money. The system works well with the dataset. We may also use independent data in the future. This means that, regardless of the datatype, our crop system must work with the same functionality.

VII. REFERENCES

[1] F.F. Haque, A. Abdelgawad, V.P. Yanambaka and K. Yelamarthi, "Crop Yield Analysis Using Machine Learning Algorithms, "2020IEEE 6th World Forum on Internet of Things (WF-IOT), 2020, pp.1-2.
[2] M. Kavita and P. Mathur, "Crop Yield Estimation in India Using Machine Learning,"2020IEEE 5th International Conference on Computing Communication and Automation (ICCCA),2020, pp.220-224.
[3] P.S. Nishant, P. Sai Venkat, B.L. Avinash and B. Jabber, " Crop Yield Prediction based on Indian Agriculture using Machine Learning" 2020 International Conference for Emerging Technology (INCET),2020, pp.1-4.
[4] R. Kumar, M.P. Singh, P. Kumar and J.P. Singh, "Crop Selection Method to Maximize crop yield rate using machine learning technique,2019, pp.138-145.
[5] Analysis of crop yield prediction using data mining techniques, 2018.
[6] Effects of NP Fertilizer Placement Depth by Year Interaction on the Numeral of Maize plants afters emergence using the additive main effects and multiplicative interaction model,2021.
[7] Bajtor. C, Mousavi, S.M.N, Marton, Stability and adaptability of maize hybrids for precision crop production in a long-term field experiment in Hungary.Agronomy,2021.
[8] Ghutake L, Verma, R.K, Chaudhari, An intelligent crop price prediction using suitable machine learning algorithm, 2021.
[9] Ashoka, D.V., Abhilash C.B, A framework for enhancing the agriculture yield using cloud clusters,2019
Jain,Ayush & Marvaniya, Smit & Godbole, Towards context-based model selection for improved crop price forecasting,2022.
[10] Aiggpa;Crop price prediction using machine learning in Madhya Pradesh,2020.
[11] Ghutake L,Verma, R.K, Chaudhari, An intelligent crop price prediction using suitable machine learning algorithm, 2021.
[12] Ashoka, D.V., Abhilash C.B, A framework for enhancing the agriculture yield using cloud clusters,2019.

.