WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Plant Disease Detection Using Deep Learning-Fertilizer Suggestion

N.Srirekha[#1], R.Angelin Preethi [#2]

[#1] M.Sc, Department of Computer Science, Kamban College of Arts and Science for Women, Tiruvannamalai-606603.

[#2] Assistant professor, Department of Computer Science, Kamban College of Arts and Science for Women, Tiruvannamalai-606603.

ABSTRACT:

India is an Agriculture based economy whose most of the GDP comes from farming. The motivation of this project comes from the increasing suicide rates in farmers which may be due to low harvest in crops. Climate and other environmental changes have become a major threat in the agriculture field. Machine learning is an essential approach for achieving practical and effective solutions for this problem. Predicting yield of the crop from historical available data like weather, soil, rainfall parameters and historic crop yield. We achieved this using the machine learning algorithm. We did a comparative study of various machine learning algorithms, i.e., ANN, K Nearest Neighbour, Random Forest, SVM and Linear Regression and chose Random Forest Algorithm which gave an accuracy of 95%.In this project a mobile application has been developed which predicts the crop yield in general and also for a particular crop. Along with that, it also suggests the user if it is the right time touse the fertilizer or not.

1.INTRODUCTION:

Agriculture is the technique of cultivating the soil, growing crops, and raising livestock. It builds the preparation ofplant and animal products for people to use and their distribution to markets. Agriculture gives most of the nation'sfood and fabrics. Agriculture plays a very important role in the global economy. The world population is increasing ata very fast rate and with an increase in population, the need for food is also increasing. Traditional methods used byfarmers are not sufficient enough to serve the huge demand and so they have to hamper the soil by using harmfulpesticides in an intensified manner. This affects the agriculture practice a lot and in the end, the land remains barren with no fertility.

Machine Learning in agriculture is used to improve the product quality of the crops in the agriculture sector.Machine Learning is the scientific field which gives the machine the ability to learn without being intervention ofhuman being. The main aim of this research is to design the crop yield prediction and soil fertility analysis model bymachine learning supervised and deep neural network model. The real-time data of soil and crop are collected fromthe different online repository (Private). To assessed the model these datasets are used and computed the result. Thedatasets are divided into two category training dataset and testing dataset to build the model. Then differentMachine learning algorithms are applied to classify the soil whether the soil is fertile or not by using soil micro-nutrients and chemical features. The crop yield, crop sowing is also analyzed using this ML technique. The algorithm assessed involves Random forest which allare available in the machine learning model

3.Literature Survey

[1].S.Veenadhari In this paper, the study was aimed to develop a website for finding out the influence of climatic parameters on crop production in selected districts of Madhya Pradesh. The selection of districts has been made based on the area under that particular crop. Based on this criteria first top five districts in which the selected crop area is maximum were selected. The crops selected in the study were based on the predominant crops in the selected district. The selected crop included: Soybean, Maize, Paddy and Wheat. The yield of these crops was tabulated for continuous 20 years by collecting the information from secondary sources. Similarly for the corresponding years climatic parameters such as Rainfall, Maximum & Minimum temperature, Potential Evapotranspiration, Cloud cover, Wet day frequency were also collected from the secondary sources. The methodology adopted for analysis includes for values above the threshold were considered as one child and the remaining as another child.

[2]. **Rushika Ghadge** This paper states, most of the existing systems are hardware based which makes them expensive and difficult to maintain and lack to give accurate results. Some systems suggest crop sequence depending on yield rate and market price. In this paper, the system proposed tries to overcome these drawbacks and predicts crops by analyzing structured data. Being a totally software solution, it does not allow maintenance factor to be considered much. Also the accuracy level would be high as compared to hardware based solutions, because components like soil composition, soil type, pH value, weather conditions all come into picture during the prediction process.

[3]. **U.Muthaiah & M.Balamurugan** This paper uses R programming with Machine Learning techniques. R is the leading tool for statistics, data analysis, and machine learning. It is more than a statistical package; it's a programming language, so you can create your own objects, functions, and packages. It's platform-independent, so it can be used on any operating system and it's free. R programs explicitly document the steps of our analysis and make it easy to reproduce and/or update analysis, which means it can quickly try many ideas and/or correct issues. All the datasets used in the research were sourced from the openly accessible records of the Indian Government. This was sourced for the years 1997 to 2013 for different seasons like Kharif and Rabi of rice production. From the vast initial dataset, only a limited number of important factors which have the highest impact on agricultural yield were selected for the present research. The dataset contains the following parameters: rainfall, season, and temperature and crop production. This paper also compares the two machine learning algorithms: decision trees and Random Forest.

[4]. **Sriram Rakshith.K** This paper is mainly focused on the techniques and measures taken to improve farming by inculcating the technical knowledge and developments in order to make the agricultural sector more reliable and easy for the farmers by predicting the suitable crop by using Machine learning techniques by sensing parameters like-- soil, weather and market trends. Parameters considered are PH, Nitrogen-phosphate-potassium contents of soil, temperature, rainfall and humidity. They consider Artificial Neural Network, Information Fuzzy Network and other Data Mining Techniques.

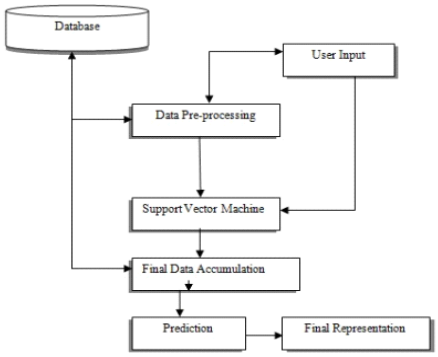
[5]. **S. Pavani** The paper says, vast research has been done and several attempts are made for application of Machine learning in agricultural fields. Major challenge in agriculture is to increase the production in the farm and deliver it to the end customers with best possible price and good quality. It is found that at least 50 percent of the farm produce never reaches the end consumer due to wastage and high-end prices. Machine learning based solutions developed to solve the difficulties faced by the farmers are being discussed in this work. The real time environmental parameters of Telangana District like soil moisture, temperature, rainfall, humidity are collected and crop yield is being predicted using KNN Algorithm.

There is a profound need to raise the farmer's income and ensure sustainable growth in Telangana to reduce poverty. Prediction of the yield of the crop in advance for a particular regiondepending on the climatic conditions and other factors which contribute to more yield

3.PROPOSED SYSTEM:

SYSTEM ARCHITECTURE DIAGRAM:

System design is the process of the defining the architecture, components, modules, interfaces, and data for a system to satisfy specified requirements .Systems design could be seen as the application of systems theory to product development. Objectoriented analysis and methods are becoming the most widely used methods for computer systems design. Systems design is therefore the process of defining and developing systems to satisfy specified requirements of the user. The UML has become the standard language in object oriented analysis and design



4. METHODOLOGY:

Data collection: The dataset used in this project is the data collected from reliable websites and merged to achieve the desired data set. The sources of our datasets are: https://en.tutiempo.net/for weather data and https://www.kaggle.com/srinivas1/agricuture-crops-production-in-india for crop yielddata. It consists of names of the crops, production, area, average temperature, average rainfall (mm), season, year, name of the states and the districts. _Production' is the dependent variable or the class variable. There are eight independent variables and 1 dependent variable.

Data Preprocessing: The purpose of preprocessing is to convert raw data into a form that fits machine learning. Structured and clean data allows a data scientist to get more precise results from an applied machine learning model. The technique includes data formatting, cleaning, and sampling. Here, data pre-processing focuses on finding the attributes with null values or invalid values and finding the relationships between various attributes as well. Data Pre-processing also helps in finding out the impact of each parameter on the target parameter. To preprocess our datasets we used EDA methodology. All the invalid and null values were handled by removing that record or giving the default value of that particular attribute based on its importance. **3.Dataset splitting**: A dataset used for machine learning should be partitioned into two subsets — training and test sets. We split the dataset into two with a split ratio of 80%i.e., in 100 records 80 records were a part of the training set and remaining 20 records

Random Forest

Random forest is a flexible, easy to use machine learning algorithm that produces, even without hyper-parameter tuning, a great result most of the time. It is also one of the most used algorithms, because of its simplicity and diversity. It can be used for both classification and regression tasks. Random forest builds multiple decision trees and merges them together to get a more accurate and stable prediction. One big advantage of random forest is that it can be used for both classification and regression problems, which form the majority of current machine learning systems. Another great quality of the random forest algorithm is that it is very easy to measure the relative importance of each feature on the prediction. Sklearn provides a great tool for this that measures a feature's importance by looking at how much the tree nodes that use that feature reduce impurity across all trees in the forest. It computes this score automatically for each featureafter training and scales the results so the sum of all importance is equal to one.

Model training: After a data scientist has preprocessed the collected data and split it intotrain and test can proceed with a model training. This process entails ―feeding‖ the algorithm with training data. An algorithm will process data and output a model that is able to find a target value (attribute) in new data an answer you want to get a predictive analysis. The purpose of model training is to develop a model. We trained our model using the random forest algorithm. On training the model it predicts the yield on giving the other attributes of the dataset as input.

Model evaluation and testing: The goal of this step is to develop the simplest model able to formulate a target value fast and well enough. A data scientist can achieve this goal through model tuning. That‘s the optimization of model parameters to achieve an algorithm‘s best performance.

5 CONCLUSION

Agriculture is the backbone of many countries including India. Since integrating the information technology with the agriculture will guide the farmer to improve the productivity. In this proposed work the system described works faster and gives better accuracy in prediction to predict the suitable crops and fertilizers for the field. It includes various parameters of soil to analyse the crop. This prediction makes the farmers to improve the productivity, growth, and quality of the plants

6. RESULTS

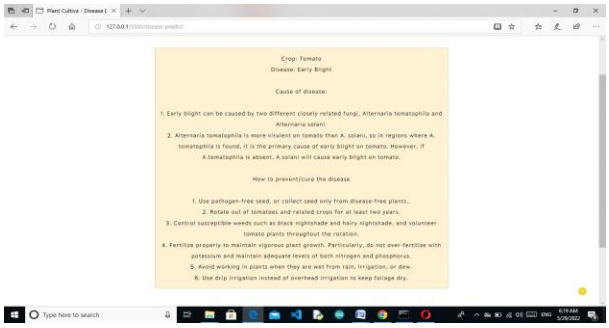


Fig 1: Analysis of Soil Fertilizer

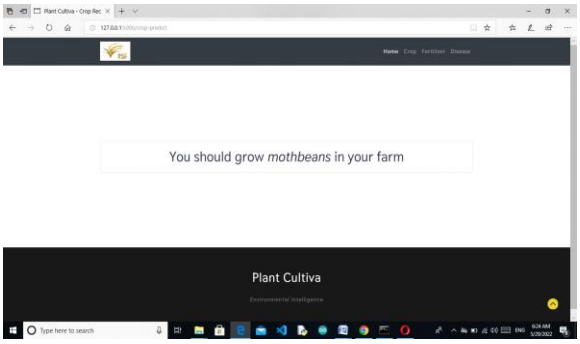


Fig 2: Classification Results

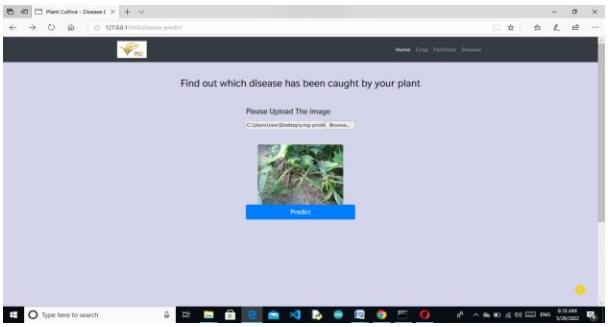


Fig:2 Plant Disease Detection

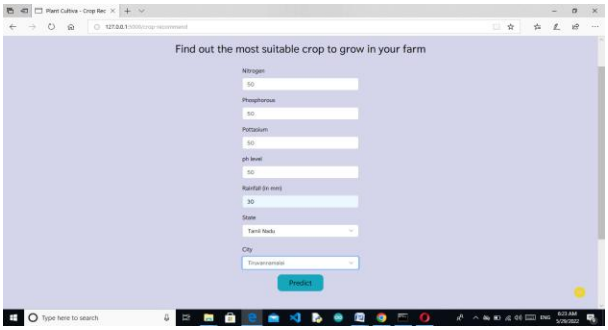


Fig 3: Crop Yield Range detection

REFERENCES:

1. National Research Council. Alternative Agriculture. National Academy Press. Washington, D.C. 1989.
2. P. Gruhn, F. Goletti, Yudelman M Integrated nutrient management, soil fertility, and sustainable agriculture: current issues and future challenges. International Food Policy Research Institute, Washington, D.C. 2000.

A. Mucherino, P. Papajorgji, P.M. Pardalos, A survey of data mining techniques applied to agriculture. Oper. Res. 9 (2009), 121-140.

3. K. Lamorski, Y. Pachepsky, C. Slawin ´ski, R.T. Walczak, Using support vector machines to develop pedotransfer functions for water retention of soils in Poland. Soil Sci. Soc. Amer. J. 72 (2008), 1243-1247.
4. M. Kovacevic, B. Bajat , B. Gajic, Soil type classification and estimation of soil properties using support vector machines, Geoderma. 154 (2009), 340-347.
5. Y. Liu, H. Wang, H. Zhang, and K. Liber, A comprehensive support vector machine-based classification model for soil quality assessment, Soil Tillage Res. 155 (2016), 19-26.
6. Y. Huang, Y. Lan, S.J. Thomson, A. Fang, W.C. Hoffmann, R.E. Lacey, Development of soft computing and applications in agricultural and biological engineering, Comput. Electron. Agric. 71 (2010), 107-127.
7. W. Wu, A.D. Li, X.H. He, et al. A comparison of support vector machines, artificial neural network and classification tree for identifying soil texture classes in southwest China. Comput. Electron. Agric. 144 (2018), 86-93
8. H. Li, W. Leng, Y. Zhou, F. Chen, Z. Xiu, D. Yang, Evaluation Models for Soil Nutrient Based on Support Vector Machine and Artificial Neural Networks, Sci. World J. 2014 (2014), 478569.
9. M.H. Yang, D.Y. Xu, S.C. Chen, et al. Evaluation of machine learning approaches to predict soil organic matter and pH using Vis-NIR spectra. Sensors. 19(2) (2019), 263.
10. M. Hosseini, S. Rajabi Agereh, Y. Khaledian, et al. Comparison of multiple statistical techniques to predict soil phosphorus. Appl. Soil Ecol. 114 (2017), 123-131.
11. K. Liakos, P. Busato, D. Moshou, S. Pearson, D. Bochtis, Machine Learning in Agriculture: A Review, Sensors. 18 (2018), 2674.
12. J.R. Romero, P.F. Roncallo, P.C. Akkiraju, et al. Using classification algorithms for predicting durum wheat yield in the province of Buenos Aires. Comput. Electron. Agric. 96 (2013), 173-179.
13. M. Sirsat, E. Cernadas, M. Fern ´andez-Delgado, M. Khan, Classification of agricultural soil parameters in India, Comput. Electron. Agric. 135 (2017), 269-279.

14. M.S. Sirsat, E. Cernadas, M. Fern ´andez-Delgado, S. Barro, Automatic prediction of village-wise soil fertility for several nutrients in India using a wide range of regression methods, Comput. Electron. Agric. 154 (2018), 120-133