CROP PREDICTION USING MACHINE LEARNING

[1]Ms. Afreen Subuhi ,[2]G. Manasa Devi ,[3]N. Ganesh Kumar & [4]P. Shravani

[1]Assistant Professor, Department of Information Technology, CMR College of Engineering & Technology

[2, 3, 4] B-Tech, Department of Information Technology, CMR College of Engineering & Technology

Abstract

As we know the fact that, India is the second largest population country in the world and majority of people in India have agriculture as their occupation. Farmers are growing same crops repeatedly without trying new verity of crops and they are applying fertilizers in random quantity without knowing the deficient content and quantity. So, this is directly affecting on crop yield and also causes the soil acidification and damages the top layer. So, we have designed the system using machine learning algorithms for betterment of farmers. Our system will suggest the best suitable crop for particular land based on content and weather parameters. And also, the system provides information about the required content and quantity of fertilizers, required seeds for cultivation. Hence by utilizing our system farmers can cultivate a new variety of crop, may increase in profit margin and can avoid soil pollution. As we all know that Agriculture is the primary source of Revenue. It becomes challenging for the farmers to satisfy our planets evolving requirements and the expectations of merchants, customers, etc. Some of the challenges the farmers face are: i) Dealing with climatic changes because of soil erosion and industry emissions ii) Nutrient deficiency in the soil, caused by a shortage of crucial minerals such as potassium, nitrogen & phosphorous can result in reduced crop growth. iii) Farmers make a mistake by cultivating the same crop over and over again without trying the new one. The project aims to discover the best model for crop prediction, which can help farmers decide the type of crop to grow based on the climatic conditions and nutrients present in the soil

INTRODUCTION:

Agriculture is one of the important occupations practiced in India. It is the broadest economic sector and plays a most important role in the overall development of the country. More than 60% of the land in the country is used for agriculture in order to suffice the needs of 1.3 billion people Thus adopting new agriculture technologies is very important. This will be leads the farmers of our country towards profit. Prior crop prediction and yield prediction was performed on the basis of farmers experience on a particular location. They will prefer the prior or neighbourhood or more trend crop in the

surrounding region only for their land and they don't have enough of knowledge about soil nutrients content such as nitrogen, phosphorus, potassium in the land. Being this as the current situation without the rotation of the crop and apply an inadequate amount of nutrients to soil it leads to reduce in the yield and soil pollution (soil acidification) and damages the top layer. Considering all these problems takes into the betterment of farmers. Why Crop prediction is important? Dealing with climatic changes because of soil erosion and industry emissions Nutrient deficiency in the soil, caused by a shortage of crucial minerals such as potassium, nitrogen & phosphorous can result in reduced crop growth. Farmers makes mistake by cultivating the same crop over and over again without trying the new one, so we recommend them about crops, disease and fertilizers. 1.2. Research Objective • Dealing with climatic changes because of soil erosion and industry emissions • Nutrient deficiency in the soil, caused by a shortage of crucial minerals such as potassium, nitrogen & phosphorous can result in reduced crop growth. • Farmers makes mistake by cultivating the same crop over and over again without trying the new one ,so we recommend them about crops, disease and fertilizers

IMPLEMENTATION OF CROP PREDICTION

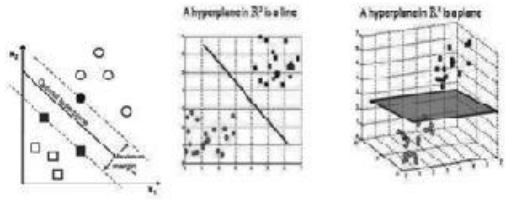

Figure 1: Maximum Margin and Hyper planes

In the proposed system, supervised learning algorithms are used to form a model which will help us in providing choices of the most feasible crops that can be cultivated in that region along with its estimated yield. Two of the algorithms used here is K-Nearest Neighbour and Support Vector Machine. The main stages involved in the process are dataset collection.

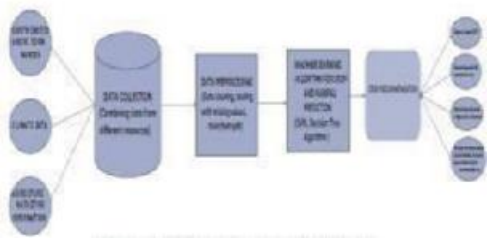

Fig:-2

Transforming the data The final step is transforming the selected data. The pre-processed data here is then transformed into data that is ready for machine learning algorithms by using various engineering features like scaling, feature aggregation and so on. There may be several features that can be combined into a single feature which would be more meaningful to the problem you are trying to solve. Figure 1

below shows the final data to be used by the classifiers.

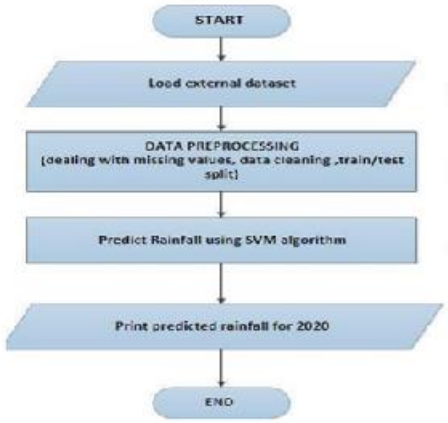

Fig:-3

PROPOSED MODEL

• The proposed solution is we are implementing an application where it will be helpful to farmers through improving agriculture techniques acknowledgement and increasing the cultivation methodologies.

• We want to help farmers for cultivation of crops and to make farmers increase their profit through cultivation.

• We use state-of-the-art machine learning and deep learning technologies to help you guide through the entire farming process. Make informed decisions to understand the demographics of your area, understand the factors that affect your crop and keep them healthy for a super awesome successful yield.

• Our project consists of three different modules like for crop recommendation, fertilizer recommendation and crop disease prediction.

• The Proposed system will predict the most suitable crop for particular land based on soil contents and weather parameters such as Temperature, Humidity, soil PH and Rainfall. 3.2. Data Collection and Performance metrics Recommending about type of crops to be cultivated based on respective conditions like climatic, soil fertility. Recommending about type of fertilizers suited for particular soil based on crops. Easily detecting the plant disease and suggesting curing the diseases

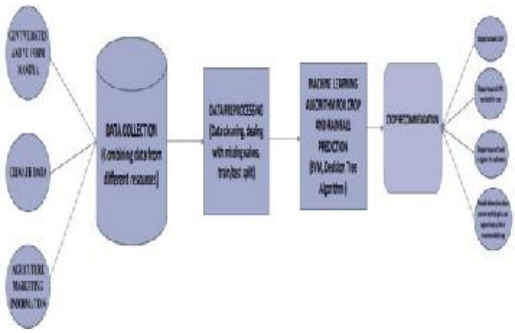

Fig:-3

RESULTS:

1.Decision Tree

The independent variables are used to create a decision tree, with each node having a conditio n over a feature. The algorithm begins at the tree's root node to predict the class of the specifi ed dataset. In layman's terms, decision trees are a series of if-else statements. It checks to see if the condition is satisfied, and if it is, it moves on to the next node in the decision chain. Bas ed on the condition, the nodes pick which node to travel to next. Output

is expected once the l eaf node is reached. The tree is efficient when the conditions are in the appropriate order. The criterion for selecting conditions in nodes is entropy/information gain. The tree structure is de rived using a recursive, greedy-based technique.

DecisionTrees's Accuracy is: 99.0

	precision	recall	f1-score	support
apple	1.00	1.00	1.00	13
banana	1.00	1.00	1.00	17
blackgram	0.59	1.00	0.74	16
chickpea	1.00	1.00	1.00	21
coconut	0.91	1.00	0.95	21
coffee	1.00	1.00	1.00	22
cotton	1.00	1.00	1.00	20
grapes	1.00	1.00	1.00	18
jute	0.74	0.93	0.83	28
kidneybeans	0.00	0.00	0.00	14
lentil	0.68	1.00	0.81	23
maize	1.00	1.00	1.00	21
mango	1.00	1.00	1.00	26
mothbeans	0.00	0.00	0.00	19
mungbean	1.00	1.00	1.00	24
muskmelon	1.00	1.00	1.00	23
orange	1.00	1.00	1.00	29
papaya	1.00	0.84	0.91	19
pigeonpeas	0.62	1.00	0.77	18
pomegranate	1.00	1.00	1.00	17
rice	1.00	0.62	0.77	16
watermelon	1.00	1.00	1.00	15
accuracy			0.90	440
macro avg	0.84	0.88	0.85	440
weighted avg	0.86	0.90	0.87	440

2.Guassian Naive Bayes

Naive Bayes classifiers are a collection of classification algorithms based on Bayes' Theore. I t is not a single algorithm but a family of algorithms where all of them share a common princi ple, i.e. every pair of features being classified is independent of each other. When we used the Naïve Bayes Algorithm the following results are obtained.

Naive Bayes's Accuracy is: 0.9999090909090091

	precision	recall	f1-score	support
apple	1.00	1.00	1.00	13
banana	1.00	1.00	1.00	17
blackgram	1.00	1.00	1.00	16
chickpea	1.00	1.00	1.00	21
coconut	1.00	1.00	1.00	21
coffee	1.00	1.00	1.00	22
cotton	1.00	1.00	1.00	20
grapes	1.00	1.00	1.00	18
jute	0.88	1.00	0.93	28
kidneybeans	1.00	1.00	1.00	14
lentil	1.00	1.00	1.00	23
maize	1.00	1.00	1.00	21
mango	1.00	1.00	1.00	26
mothbeans	1.00	1.00	1.00	19
mungbean	1.00	1.00	1.00	24
muskmelon	1.00	1.00	1.00	23
orange	1.00	1.00	1.00	29
papaya	1.00	1.00	1.00	19
pigeonpeas	1.00	1.00	1.00	18
pomegranate	1.00	1.00	1.00	17

Random Forest

Random Forest is a popular machine learning algorithm that belongs to the supervised learnin g technique. It can be used for both Classification and Regression problems in ML. It is based on the concept of ensemble learning, which is a process of combining multiple classifiers to s olve a complex problem and to improve the performance of the model. As the name suggests, "Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accura cy of that dataset." Instead of relying on one decision tree, the random forest takes the predict ion from each tree and based on the majority votes of predictions, and it predicts the final out put.

RF's Accuracy is: 0.9909090909090091

	precision	recall	f1-score	support
apple	1.00	1.00	1.00	13
banana	1.00	1.00	1.00	17
blackgram	0.94	1.00	0.97	16
chickpea	1.00	1.00	1.00	21
coconut	1.00	1.00	1.00	21
coffee	1.00	1.00	1.00	22
cotton	1.00	1.00	1.00	20
grapes	1.00	1.00	1.00	18
jute	0.90	1.00	0.95	28
kidneybeans	1.00	1.00	1.00	14
lentil	1.00	1.00	1.00	23
maize	1.00	1.00	1.00	21
mango	1.00	1.00	1.00	26
mothbeans	1.00	0.95	0.97	19
mungbean	1.00	1.00	1.00	24
muskmelon	1.00	1.00	1.00	23
orange	1.00	1.00	1.00	29
papaya	1.00	1.00	1.00	19
pigeonpeas	1.00	1.00	1.00	18
pomegranate	1.00	1.00	1.00	17
rice	1.00	0.81	0.90	16
watermelon	1.00	1.00	1.00	15
accuracy			0.99	440

Accuracy Comparison

```
plt.figure(figsize=[10,5],dpi = 100)
plt.title('Accuracy Comparison')
plt.xlabel('Accuracy')
plt.ylabel('Algorithm')
```

sns.barplot(x = acc,y = model,palette='dark')

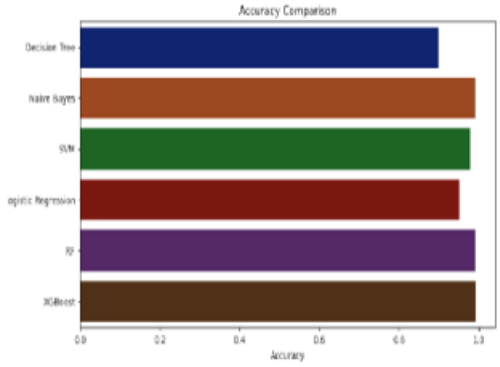

Fig:-4

CONCLUSION :

The comparative study of three different supervised machine learning models (KNN, Decisio n Tree, and Random Forest) is done to predict the best-suited crop for the particular land that can help farmers to grow crops more efficiently In completion, we concluded that the crop prediction dataset showed the best accuracy with R andom Forest Classifier both in Entropy and Gini Criterion with 99.32%. This project highlighted the limitations of current systems and their practical usage on yield p rediction. Then walks through a viable yield prediction system to the farmers, a proposed system provid es connectivity to farmers via a web application. The web application includes multiple features that users can leverage for the selection of a cr op. The inbuilt predictor system helps the farmers to predict the yield of a given crop. The inbuilt recommender system allows a user exploration of the possible crops and their yiel d to take more educated decisions. For yield to accuracy, various machine learning algorithm

FUTURE SCOPE :

We have to collect all required data by giving GPS locations of a land and by taking access fr om Rain forecasting system of by the government, we can predict crops by just giving GPS lo cation. Also, we can develop the model to avoid over and under crisis of the food.

REFERENCES

[1] Monali Paul, Santosh K. Vishwakarma, Ashok Verma, “Analysis of Soil Behavior an d Prediction of Crop Yield using Data Mining approach”, 2015 International Conferen ce on Computational Intelligence and Communication Networks.

[2] Abdullah Na, William Isaac, ShashankVarshney, Ekram Khan, “An IoT Based Syste m for Remote Monitoring of Soil Characteristics”, 2016 International Conference of I nformation Technology.

[3] Dr.N. Suma, Sandra Rhea Samson, S. Saranya, G. Shanmugapriya, R. Subhashri, “IO T Based Smart Agriculture Monitoring System”, Feb 2017 IJRITCC.

[4] N. Heemageetha, “A survey on Application of Data Mining Techniques to Analyze th e soil for agricultural purpose”, 2016IEEE.

[5] GiritharanRavichandran, Koteeshwari R S “Agricultural Crop Predictor and

Advisor u sing ANN for Smartphones", 2016 IEEE,

[6] GiritharanRavichandran, Koteeshwari R S "Agricultural Crop Predictor and Advisor u sing ANN for Smartphones", 2016 IEEE.

[7] R. Nagini, Dr. T.V. Rajnikanth, B.V. Kiranmayee, "Agriculture Yield Prediction Usin g Predictive Analytic Techniques, 2nd International Conference on Contemporary Co mputing and Informatics (ic3i),2016.

[8] Awanit Kumar, Shiv Kumar, "Prediction of production of crops using K-Means and F uzzy Logic", IJCSMC, 2015

[9] JeetendraShenoy, YogeshPingle, "IOT in agriculture", 2016 IEEE.

[10] Vinay, R., Soujanya, K. L. S., & Singh, P. (2019). Disease prediction by using deep learning based on patient treatment history. Int. J. Recent Technol. Eng, 7(6), 745-754.

[12] Madhavi, K. R., Kora, P., Reddy, L. V., Avanija, J., Soujanya, K. L. S., & Telagarapu, P. (2022). Cardiac arrhythmia detection using dual-tree wavelet transform and convolutional neural network. Soft Computing, 26(7), 3561-3571.

[13] Dabbu, M., Karuppusamy, L., Pulugu, D., Vootla, S.R., Reddyvari, V.R., 2022, Water atom search algorithm-based deep recurrent neural network for the big data classification based on spark architecture, International Journal of Machine Learning and Cybernetics, 10.1007/s13042-022-01524-8

[14] Venkateshwarlu, M., Rajagopal, K., Reddy, C.O., Reddy, Y.S., 2022, Investigation of Groundwater Contamination due to Landfill Leachate, Indian Journal of Environmental Protection, Volume: 42

[15] Suryanarayana, G., Prakash K, L., Mahesh, P.C.S., Bhaskar, T., 2022, Novel dynamic k-modes clustering of categorical and non categorical dataset with optimized genetic algorithm based feature selection, Multimedia Tools and Applications, 10.1007/s11042-022-12126-5

[16] Balamurugan, D., Aravinth, S.S., Reddy, P.C.S., Rupani, A., Manikandan, A., 2022, Multiview Objects Recognition Using Deep Learning-Based Wrap-CNN with Voting Scheme, Neural Processing Letters, 10.1007/s11063-021-10679-4

[17] Das, S., Nayak, S.C., Sahoo, B., 2022, Towards Crafting Optimal Functional Link Artificial Neural Networks with Rao Algorithms for Stock Closing Prices Prediction, Computational Economics, 10.1007/s10614-021-10130-9