e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
Volume:03/Issue:05/May-2021 Impact Factor- 5.354 www.irjmets.com

AUTOMATED CROP SUGGESTION SYSTEM

Somasunder.S[*1], Samritha Shivani.K[*2], Dr.C.Sathya[*3], Dr.P.Velvadivu[*4]

[*1,2]2nd Year, M.Sc Data Science(Integrated), Coimbatore Institute Of Technology, Coimbatore, Tamilnadu, India.

[*3,4]Assistant Professor, Department Of Data Science, Coimbatore Institute Of Technology, Coimbatore, Tamilnadu, India.

ABSTRACT

In Agriculture, crops are lost due to the mistakes made during crop selection. A lot of farmers are not really aware of the basic crop requirements like minerals, soil moisture, required temperature, rainfall, etc. Therefore, the problem of a farmer is addressed in our paper and we have solved it by creating an Automated Crop Recommendation System (Tamilnadu Regions only) which would recommend the most suitable crops for a given location based on the analysis of the present and past data. The analysis is performed using machine learning concepts like random forest model, gaussian naive bayes and decision tree classifier. The model with a high accuracy rate will be considered the best fit. The major parameters involve season , temperature, humidity, rainfall, soil type , pH of soil and based on previous year's crop yield and production. Apart from this, we have used the API (Application Programming interface) for accessing current weather data.

Keywords: Machine Learning, Random Forest, Gaussian Naive Bayes And Decision Tree, API, Soil Classification, Automation And Crop Recommendation.

I. INTRODUCTION

In Tamilnadu, almost 70% of the state population depends on agriculture and related works. Tamilnadu has an area of 130.33 Lakh Ha . From 2000-2020, the agricultural area was known to have used 30%-50% of the state's geographical area. In 2020, 5.7 Million Hectare(43.8%) area was used for agriculture in Tamil Nadu .

It is true that a farmer is the best decision maker for crop selection and cultivation, but at times, the decision made might not be accurate. However, machine learning techniques can be applied in this field for far greater accuracy and stability of crop selection. Various machine learning models are used for predicting crops depending on various factors. Machine Learning offers a system to self-learn and progress from experience. Hence, there is a necessity for effective crop prediction techniques for better crop production.

The currently existing crop prediction systems are a bit difficult to be used by farmers as it requires lots of inputs to be provided for selecting the suitable crops. But as a farmer, he cannot collect data on his own and provide it as input for the prediction process. In these types of situations, the automated system comes into play. The automated system detects the current location of the agricultural site and collects all the related data of that particular location required for the crop prediction process, thereby making it easier for the farmers to operate it.

II. METHODOLOGY

DATASET DESCRIPTION

Present Data (Crop Requirements):

In this dataset, the crops are mapped with the value of required factors. The following factors are type of soil , amount of rainfall , temperature, humidity, pH of soil , etc.the data set is taken from github.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
Volume:03/Issue:05/May-2021 Impact Factor- 5.354 www.irjmets.com

	crop	soil type	ph	rainfall	humidity	temperature
0	rice	Black Soils	6.502985	202.935536	82.002744	20.879744
1	rice	Red Soil	7.038096	226.655537	80.319644	21.770462
2	rice	Black Soils	7.840207	263.964248	82.320763	23.004459
3	rice	Red Soil	6.980401	242.864034	80.158363	26.491096
4	rice	Black Soils	7.628473	262.717340	81.604873	20.130175

Figure 1: Present dataset

Past Data(Previous year's production):

The datasets for this project were downloaded from tn.data.gov.in .The datasets were combined and modified into realtime datasets. which has tamilnadu's all past agricultural agriculture data from 1997 . The parameters are district names, year, season, soil type, crop name, area used , total production, etc.

	State_Name	District_Name	Crop_Year	Season	Crop	Area	Production	Yield
2557	Tamil Nadu	ERODE	2004	Kharif	Moong(Green Gram)	5644	2484.0	0.440113
10853	Tamil Nadu	TIRUNELVELI	2011	Whole Year	Black pepper	40	8.0	0.200000
6483	Tamil Nadu	PUDUKKOTTAI	2009	Whole Year	Sugarcane	7133	723920.0	101.488855
4926	Tamil Nadu	NAGAPATTINAM	2002	Whole Year	Sweet potato	15	205.0	13.666667
5370	Tamil Nadu	NAMAKKAL	2004	Kharif	Arhar/Tur	1307	1066.0	0.815608

Figure 2: Past dataset from 1977 to 2020

III. MODELING AND ANALYSIS

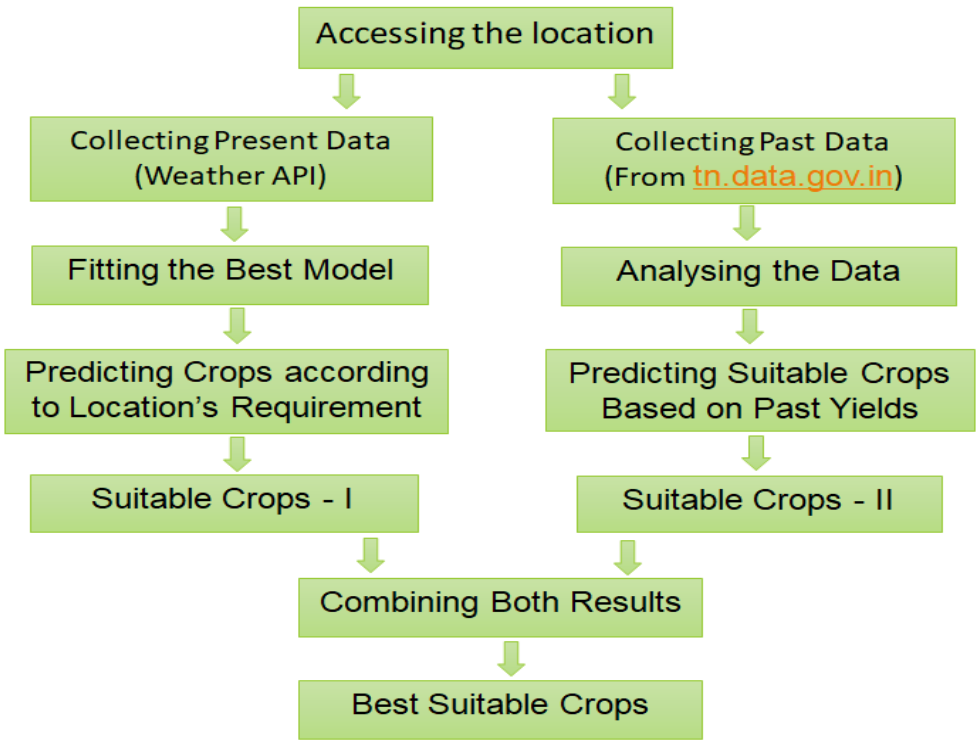

Figure 3: Depicts the overall methodology of the proposed system.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
Volume:03/Issue:05/May-2021 Impact Factor- 5.354 www.irjmets.com

WEATHER API

Weather APIs provide us with real time present and past weather related data like temperature, humidity, atmospheric pressure , rainfall , weather description and more. These data will be accessed and will be used for real time prediction of suitable crops based on given weather conditions. By this way the farmers automatically get suggestions on the types of crops that can be grown in their land without the hassle of providing inputs. [Shown in Fig.4]

DECISION TREE ALGORITHM

Decision Tree is a Supervised learning technique that is mostly used to solve problems based on classification. This machine learning technique trains the given data and creates a procedure which predicts the result for unknown data, which, in this case, is the suitable crop. This algorithm uses a tree shaped structure to classify the data (thereby forming a root node) into different categories based on the data values and their respective labels. From the root node, the data gets divided into two branches (based on a condition), which can be further divided or can end up with a leaf node which shows the final classification of the given data. [Result shown in Fig.5]

GAUSSIAN NAIVE BAYES ALGORITHM

Gaussian Naive Bayes is a type of Naive Bayes that follows normal distribution and supports continuous data. This algorithm works mostly with large databases. It assumes that variables do not depend on each other. The classifier calculates the probability of every possible factor which in this case are the crop requirements such as soil type, temperature, etc. Finally, the selected result would be the one with a higher probability. The crops that suit the factor with the highest probability are then put up for suggestion. [Result shown in Fig.6]

RANDOM FOREST ALGORITHM

Random forests are machine learning models which have the ability to perform both classification and regression. The training model of an RF algorithm requires a bunch of decision trees. These individual trees will be combined to form a single powerful model. After the model is created, we can predict the results from the RF classifier and suggest crops. The reason behind the combining process is that, when a model is employed individually, it is considered to be weak. But when the models are employed together, they are considered to be strong and stronger models have greater accuracy. [Result shown in Fig.7]

IV. RESULTS AND DISCUSSION

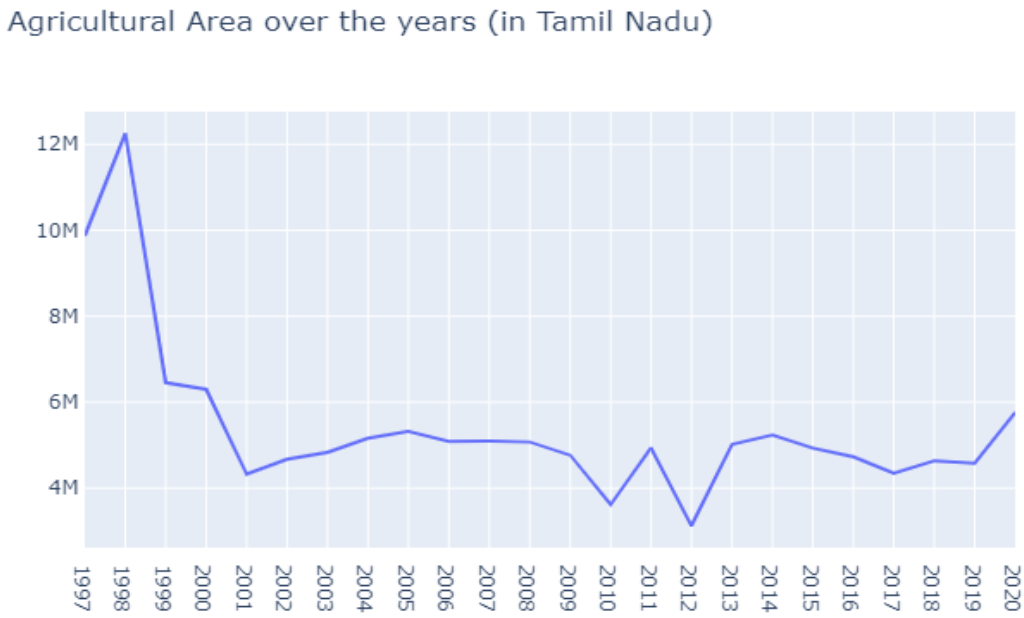

Figure 4: Agricultural Area

The major state area and population depends on the agriculture sector. Therefore, making it more accurate and efficient is the need of the hour.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
Volume:03/Issue:05/May-2021 Impact Factor- 5.354 www.irjmets.com

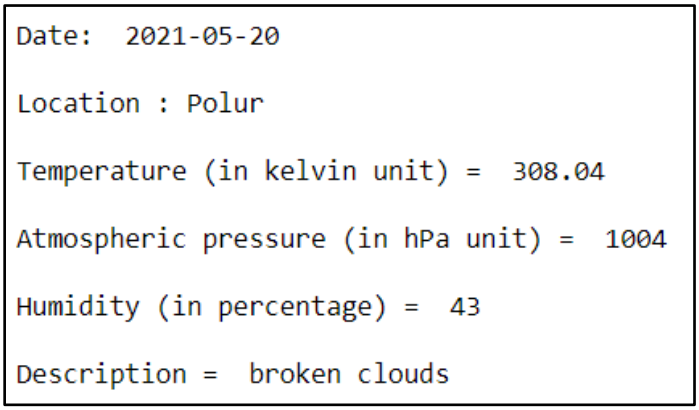

Figure 5: Current location details

The automated system detects the current location and accesses the details respective to the location [Shown in Fig 5] , which are required for the crop suggestion process.

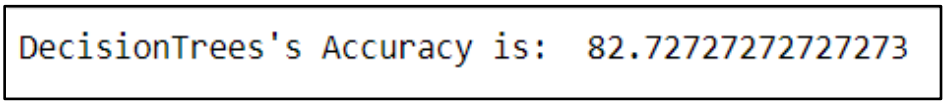

Figure 6: Decision Tree model accuracy

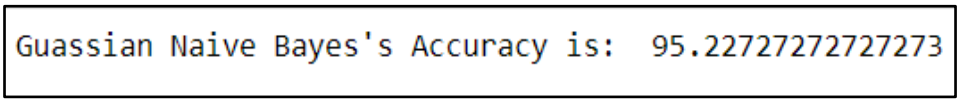

Figure 7: Gaussian Naive Bayes model accuracy

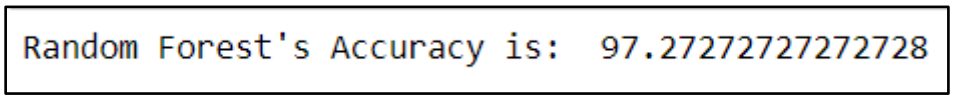

Figure 8: Random Forest model accuracy

[From Fig 6,7,8] , the accuracy for random forest model, gaussian naive bayes and decision tree classifier is found to be 97.3% , 95.2% and 82.7% . Random Forest model is considered to be the best fitted model as it has got a higher accuracy when compared to the other models.

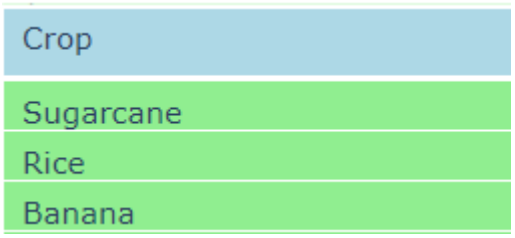

Figure 9: Final output (Suggested Crop)

The final output is produced through combining the output of past and present data analysis. [Last second step in Fig. 3]

Case 1: If "suitable crop I" and "suitable crop II" have similar data , then those crops will be suggested.

Case 2: If "suitable crop II" has profitable yields from past crops. Then those crops will be suggested first.

Case 3: If the "suitable crop II" has non-profitable yields from past crops. Then the crops from "Suitable crop I" will be suggested first.

V. CONCLUSION

Agriculture plays a vital role in contributing to the growth of our nation's economy. Our work would help the farmers to improve crop yield through easy access to the automated crop recommendation system. The goal of our future work is to improve our dataset by including many other factors that affect crop yield and also to make this automated system available across the country.

VI. REFERENCES

[1] https://tn.data.gov.in
[2] https://cms.tn.gov.in/sites/default/files/documents/agri_e_pn_2020_21.pdf
[3] https://openweathermap.org/api
[4] S. Dhamodaran, Allwyn Raja, "Prediction of Weather Condition Using Probability Function", International Journal of Applied Engineering Research, ISSN 0973- 4562 Volume 10, Number 4 (2015) pp.10665-10670.
[5] 2016 IEEE Eighth International Conference on Advanced Computing (ICoAC), "Crop Recommendation System for Precision Agriculture", S.Pudumalar*, E.Ramanujam*, R.HarineRajashreeń, C.Kavyań, T.Kiruthikań, J.Nishań.
[6] R.Nagini, Dr. T.V. Rajnikanth, B.V. Kiranmayee, "Agriculture Yield Prediction Using Predictive Analytic Techniques , 2nd InternationalConference on Contemporary Computing and Informatics (ic3i),2016
[7] 2017 International Conference on I2C2, "Agriculture decision support system using data mining", Prof. Rakesh Shirsath; Neha Khadke; Divya More.