DATA SCIENCE PROJECT

2022-2023

NARWADE SHRESHTA

1009-21-107-037

UNDER THE GUIDENCE OF

MR.TANVEER,

MRS. S. SRAVANTHI M.S.C. , M.TECH
(ASSISTANCE PROFESSOR)

CROP RECOMMENDATION USING ENSEMBLING TECHNIQUE

CONTENTS

ABSTRACT

The agriculture sector plays a crucial role in the development of the country's economy. The production of a particular farm depends upon soil characteristics, environmental characteristics, but major part goes to crop selection to get a better yield. Farmers sometimes lack the knowledge to choose the best crop for their land . Incorrect crop selection can lead to loss.

Applying modern technologies like machine learning algorithms assist in accurately predicting or recommending suitable crop to the farmer by considering the factors soil nutrients and some weather conditions like temperature ,rainfall ,humidity and pH.

Applied ensembling techniques like Gradient Boosting Classifier ,XGBClassifier ,AdaBoostClassifier and Random Forest Classifier in which Random Forest Classifier have achieved the highest accuracy.

INTRODUCTION

Agriculture, as we all know, is the foundation of the Indian economy.Agriculture is an important occupation in India. More than 60% of the country’s land is used for agriculture,which feeds 1.3 billion people.One of the main domains in the precision agriculture is the recommendation of crops to increase crop yield.

The features given to the model are characteristics of soil, rainfall ,humidity and temperature of environment.Based on those features provided as input , the farmers get suggested the most suitable crop.Machine learning techniques have created new apportunities to farmers by providing rich recommadation through customized information.

LITERATURE REVIEW

Year	Author/ publication	Title	Dataset/ features	Method/ Algorithm	Metrics	Challenge
2022	Mohamed Bouni, Badr Hssina, Khadija Douzi, Samira Douzi. Elsevier	Towards an Efficient Recommender systems in Smart Agriculture:A deep reinforcement learning approch	Soil data set Crop yield data set	Naïve Bayes, KNN, Decision tree, Random forest	Accuracy score	Overfitting problem
2022	S.P Raja, Barbara Sawicka, Zoran Stamenhovic, G.Mariammal IEEE	Crop prediction Based on Characteristics of the Agricultural environment using various feature	felin dataset soil ,rainfall, Temperature, humidity	Naïve Bayes, KNN, Decision tree, Random forest, Bagging	Accuracy, Recall, Precision, F1 score Accuracy 87.43%	The result depict that an ensemble technique offer better prediction accuracy than the existing classification

2021	Priyadarshini, Swapneel Chakraborty, Aayushkumar, Omen Rajendra pooniwala IEEE	Intelligent crop recommendation system using Machine learning	Crop yield data set Soil nutrients , Yield, model price of crops,rainfall, Temperature	Linear Regression , Neural Network	Accuracy score Accuracy 84.42%	They have taken less features and attained less accuaracy
2020	Amaury Dubois, Fabien Teytand, Sebastien Verel Elsevier	Short term soil moisture forecast for potato crop farming: a machine learning approach	Potato crop dataset Water pressure, Men temperature, Rainfall, Age , dry	Support vector machine, Random Forest, Neural Network	Root mean square error(RMSE), Mean absolute error(MAE), R squared (R^2) Accuracy 70%	With large dataset we will able to improve the performance of algorithms

2019	Neha Rale, Raxithumar Solanki,Doina bein, James andro vasko, Wolfgang IEEE	Prediction of crop cultivation	Winter wheat dataset, Windspeed dataset. Temperature, Average wind speed, precipitaion	Random Forest Regressor, Nearest neighbor regression, Support vecror machine, Gradient boosting trees	Root mean square error(RMSE) Accuracy 83%	overfitting
2018	Sk Al Zaminur Rahman, Kaushik Chandra Mitra, S.M Mohidul Islam IEEE	Soil Classifiaction using machine learning methods and crop suggestion based on soil series	Soil dataset Class label, Map unit , Upazilla code, Crop list	Weighted k-nearest neighbor (KNN), Bagged trees, Support vector machine (SVM)	Accuracy score Accuracy 92.93%	Small dataset Collected for only one district

EXISTING SYSTEM

- Farmers predict crops based on their own experience and observed weather condition. The climate is changing and the shifts from normal weather pattern are more frequent than before.
- Sometimes farmers were failed to choose the right crops based on the soil conditions, sowing season, and geographical location. This results in suicide, quitting the agriculture field, moving towards urban areas for livelihood.
- Farmers must make use of new innovative technologies to make use of existing soil, water and air conditions to obtain larger crops.

Proposed System

- The project aims to tackle the difficulties faced by the farmers in selecting the crops and recommendation of the best suitable crops in the area so that the farmer does not incur any loss.
- The proposed model recommend the crop to be selected by the farmer for growing, by considering some soil nutrients and whether conditions like temperature, rainfall, humidity,PH.
- The crop is recommended based on the features by using the machine learning ensembling techniques. In this we applied Random Forest, Gradient Boosting Classifier, AdaBoost Classifier, XGB Classifier.
- The ensembling techniques is used to build a model that combines the predictions of multiple machine learning models together to recommend the right crop based on characteristics with high accuracy.

METHODOLOGY

1.Collection of the dataset

2.Preprocessing of the data

3.Splitting the data in to training and testing set

4.Applying the machine learning algorithms on the dataset

5.Predicting the accuracy

6.Evaluating the results and recommending the crop

soil nutrients

weather condition

Data preprocessing

Machine learning algorithms

Training models

Crop recommendation

DATASET

Crop recommended dataset

Soil nutrients dataset

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	N	P	K	EC	OC	S	Zn	Fe	Cu	Mn	B	temperature	humidity	ph	rainfall	label
2	90	42	43	0.62	0.7	5.9	0.24	0.31	0.77	8.71	0.11	20.8797437	82.0027442	6.50298529	202.935536	rice
3	85	58	41	0.75	1.06	25.4	0.3	0.86	1.54	2.89	2.29	21.7704617	80.3196441	7.03809636	226.655537	rice
4	60	55	44	0.51	1.11	14.3	0.3	0.86	1.57	2.7	2.03	23.0044591	82.3207629	7.84020714	263.964248	rice
5	74	35	40	0.58	0.94	26	0.34	0.54	1.53	2.65	1.82	26.4910963	80.1583626	6.9804009	242.864034	rice
6	78	42	42	0.4	0.86	11.8	0.25	0.76	1.69	2.43	2.26	20.1301748	81.6048729	7.62847289	262.71734	rice
7	69	37	42	0.65	0.72	11.7	0.37	0.66	0.9	2.19	1.82	23.0580487	83.3701177	7.0734535	251.055	rice
8	69	55	38	0.43	0.81	7.4	0.34	0.69	1.05	2	1.88	22.708838	82.6394139	5.70080568	271.32486	rice
9	94	53	40	0.59	0.69	7.6	0.32	0.68	0.62	2.43	1.68	20.2777436	82.8940862	5.71862718	241.974195	rice
10	89	54	38	0.44	0.67	7.3	0.63	0.66	0.94	2.43	1.79	24.5158807	83.5352163	6.68534642	230.446236	rice
11	68	58	38	0.33	0.78	9	0.69	0.41	1.15	2.75	2	23.2239739	83.0332269	6.33625353	221.209196	rice
12	91	53	40	0.45	0.97	9.6	0.71	0.38	1.33	2.79	2.41	26.5272351	81.4175385	5.38616779	264.61487	rice
13	90	46	42	0.73	0.89	9.2	0.63	0.47	1.03	2.79	2.38	23.9789822	81.450616	7.50283396	250.083234	rice
14	78	58	44	0.6	0.78	9.7	0.73	0.36	1.32	3.32	2.12	26.800796	80.8868482	5.10868179	284.436457	rice
15	93	56	36	0.53	0.81	10.2	0.51	0.56	1.26	2.9	2.29	24.0149762	82.0568718	6.98435366	185.277339	rice
16	94	50	37	0.77	0.72	9.7	0.58	0.47	1.02	3.77	2.56	25.6658521	80.6638504	6.94801983	209.586971	rice
17	60	48	39	0.34	0.67	10.6	0.77	0.41	1.28	3.04	2.79	24.2820941	80.3002559	7.04229907	231.086335	rice
18	85	38	41	0.88	0.75	11	0.46	0.38	1.16	2.96	1.32	21.5871178	82.7883708	6.24905066	276.655246	rice
19	91	35	39	0.55	0.67	10.2	0.28	0.44	1.26	7.75	2.56	23.7939196	80.4181796	6.97085975	206.261186	rice
20	77	38	36	0.78	0.61	10.5	0.3	0.49	0.66	7.74	1.85	21.8652524	80.1923008	5.95393328	224.555017	rice
21	88	35	40	0.63	0.78	11.6	0.29	0.43	0.57	7.73	0.74	23.5794363	83.5876032	5.85393208	291.298662	rice
22	89	45	36	0.62	0.75	11	0.32	0.5	0.81	4.99	2.65	21.3250416	80.474764	6.44247537	185.497473	rice
23	76	40	43	0.51	0.69	23.6	0.28	0.93	1.04	2.17	1.97	25.1574553	83.1171348	5.07017567	231.384316	rice

FEATURES

- Nitrogen (N) – Nitrogen is largely responsible for the growth of leaves on the plant.
- Phosphorus (P) – Phosphorus is largely responsible for root growth and flower and fruit development.
- Potassium (K) – Potassium is a nutrient that helps the overall functions of the plant perform correctly.
- Temperature - temperature in degree Celsius
- humidity - relative humidity in %
- Potential of hydrogen(ph) - ph value of the soil
- rainfall - rainfall in mm

ALGORITHMS/METRICS

- Random forest : Random Forest models combines the output of multiple decision tree to reach a single output.Decide where to split based on a random selection of features and how to make the decisions.
- Gradiant Boosting Classifier:It gives a prediction model in the form of an ensrmble of weak prediction models, which are typically decision trees. Gradient boosting is one of the boosting algorithms it is used to minimize bias error of the model.
- AdaBoost Classifier : Aaptive boosting , these algorithm improve the prediction power by converting a number of weak learners to strong learners.Adadaptive boosting has immunity from overfitting of data as it runs each model in a sequence and has a weight associated with them.
- XGB Classifier : XGBoost stands for “Extreme Gradient Boosting” has the ability to train large datasets and One of the key features of XGBoost is its efficient in handling of missing values in the dataset.
- Metrics : Accuracy score

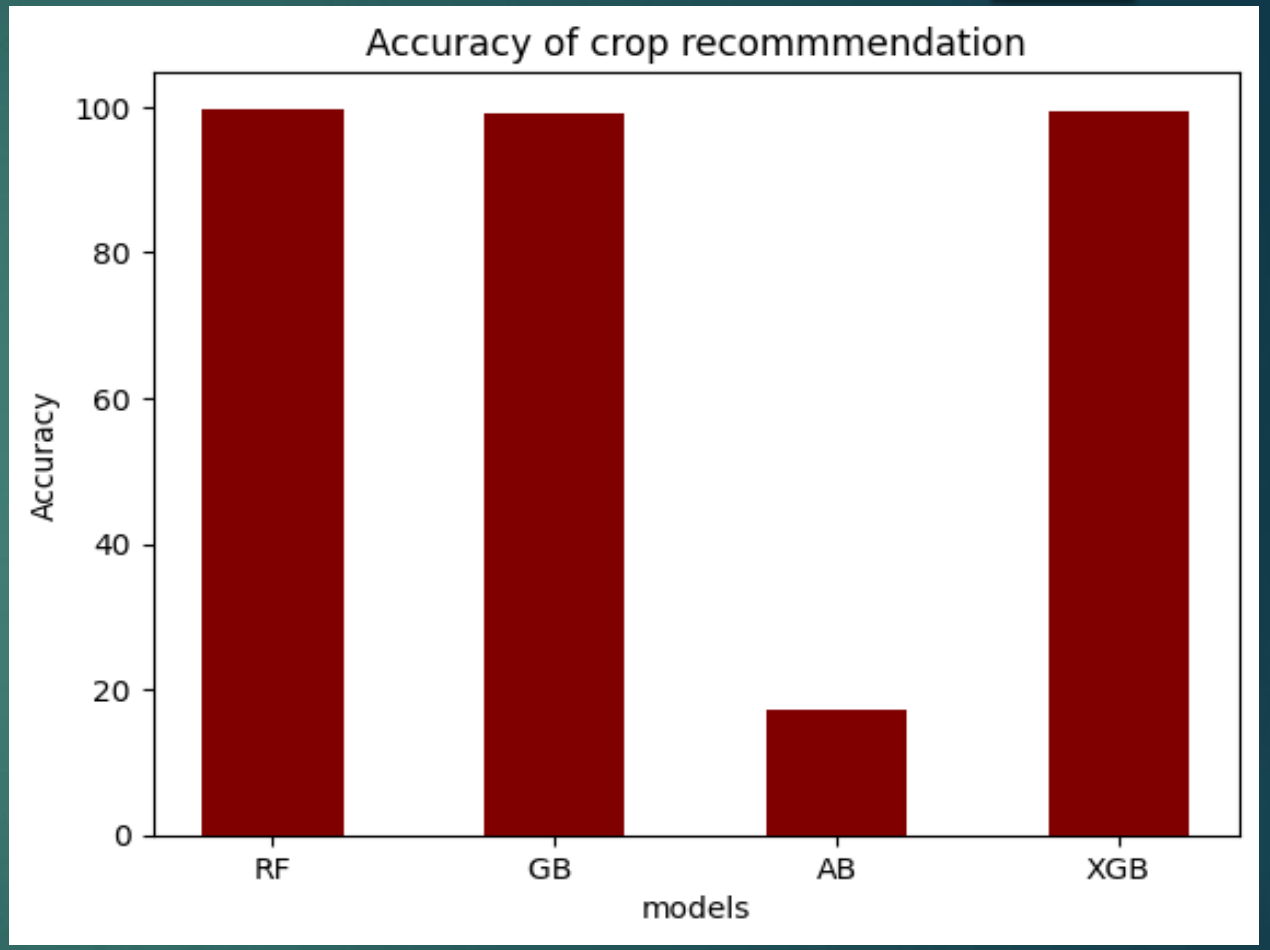
RESULTS

Random Forest(RF) :99.77%

Gradient Boosting(GB) :99.09%

AdaBoost(AB) :17.27%

XGBoost(XGB):99.32%



FUTURE WORK

- We can use deep learning techniques like Convolutional neural networks ,Recurrent neural network to make this model more reliable and accurate.

- We can develop this model for fertilizer recommendation , Pest control by adding some more features.

Conclusion

This system would assist farmers in making an informed decision about which crop to grow depending on a
variety of environmental and geographical factors.

The proposed ensemble recommendation system helps the farmers to choose right crop and thereby decreasing the chance of crop failure and increase the productivity.

Among all ensembling techniques RandomForest got highest accuracy of 99.77%

REFERENCES

- [1] Mohamed Bouni et.al.," Towards an Efficient Recommender systems in Smart Agriculture:A deep reinforcement learning approach", Elsevier,pp.826-830,2022.
- [2] Sk Al Zaminur Rahman et.al.," Crop prediction Based on Characteristics of the Agricultural environment using various feature selection Techniques and classifiers",IEEE,vol 10,pp.23625-23641,2022.
- [3] Priyadarshin et.al.," Intelligent crop recommendation system using Machine learning", IEEE,pp.843-848,2021.
- [4] Amaury Dubois et.al.," Short term soil moisture forecast for potato crop farming: a machine learning approach",Elsevier,pp.1-40,2020.
- [5] Neha Rale et.al.," Prediction of crop cultivation",IEEE,pp.0227-0232,2019.

[6] Sk Al Zaminur Rahman et.al.,” Soil Classifiaction using machine learning methods and crop suggestion based on soil series”,IEEE,2018.
[7]Vaishnavi et.al.,”Agricultural Crop Recommendations based on Productivity and season”,IEEE,pp.883-886,2021.

[8]Apoorva Chaudhari et.al.,”Crop Recommendation and its Optimal pricing using ShopBot”,IEEE,pp.36-41,2020.

[9]Avinash Kumar et.al.,”Recommendation System for crop Identification and Pest control Technique in Agriculture”,IEEE,pp.0185-0189,2019.

[10]Sonal Jain et.al.,”Machine Learning convergence for weather based crop selection”,IEEE,2020

THANK YOU