INTRODUCTION

In India even after so many developments we are very behind in our agriculture developments despite agriculture being the biggest contributor in our GDP. It is surprising to see that a country which has placed such importance on agriculture does not use the latest technology and methods to improve the sector. There can be many reasons this happens. One of the reasons could be that farmers are not aware of various technologies or there is not enough tech to support our farmers. We have used our tech to change different industries and the effects have been phenomenal to say the least. So the next obvious question arises to why not introduce technology so that farmers can also benefit.

In this project we tried to address this very problem by introducing various ML models that have helped us to predict which crops farmers should plant depending upon the various factors like the weather concentration of N,P,K rainfall and other factors which affect the crops. Apart from selecting the crops we have also worked on selecting the crops based on profit for which we have used data sets which have already defined what are the various parameters of the profit. This helps the farmers to select the best crop that is not only best under the given conditions but also ensures them the maximum profit from this.

1. LITERATURE SURVEY

1.1. Existing problem

1.1.1. Agribusiness adds to 18% of India's Gross Domestic Product (GDP) and offers work to half of the population in the country. Every individual has been rehearsing Agriculture for quite a long time yet the outcomes are rarely fulfilling due to different variables that influence the harvest yield. To satisfy the necessities of around 1.2 billion individuals, have a decent yield of harvests. Due to factors like soil type, precipitation, seed quality, absence of specialized resources, the harvest yield is directly impacted.

1.2. Proposed solution

1.2.1. We center around carrying out crop yield forecast frameworks by utilizing Machine learning techniques by exploring and evaluating different datasets. For assessing performance, accuracy is utilized as one of the components. The classifiers are additionally compared over the upsides of Precision, Recall, and F1 Score. The lesser the value of inaccuracy, the more precisely the calculation will work. The outcome depends on a correlation among the classifiers.

2. THEORETICAL ANALYSIS

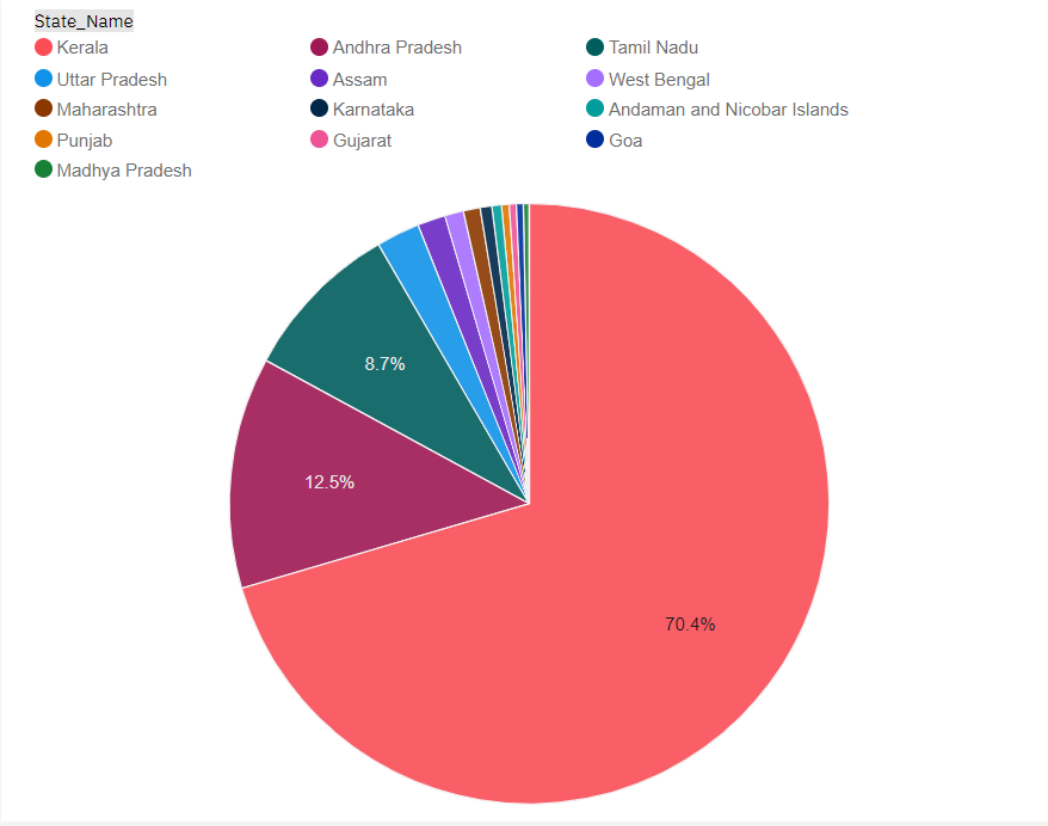
2.1. We have used watson studio and have used auto ai to generate the models and selected the best algo that was working for the given data sets. We ran all the datasets in the auto ai experiment which tested various algorithms based on the dataset and selected the best algo. It also worked on identifying the problem and classifying it into a regression or classification problem depending on the dataset.

3. EXPERIMENTAL INVESTIGATIONS

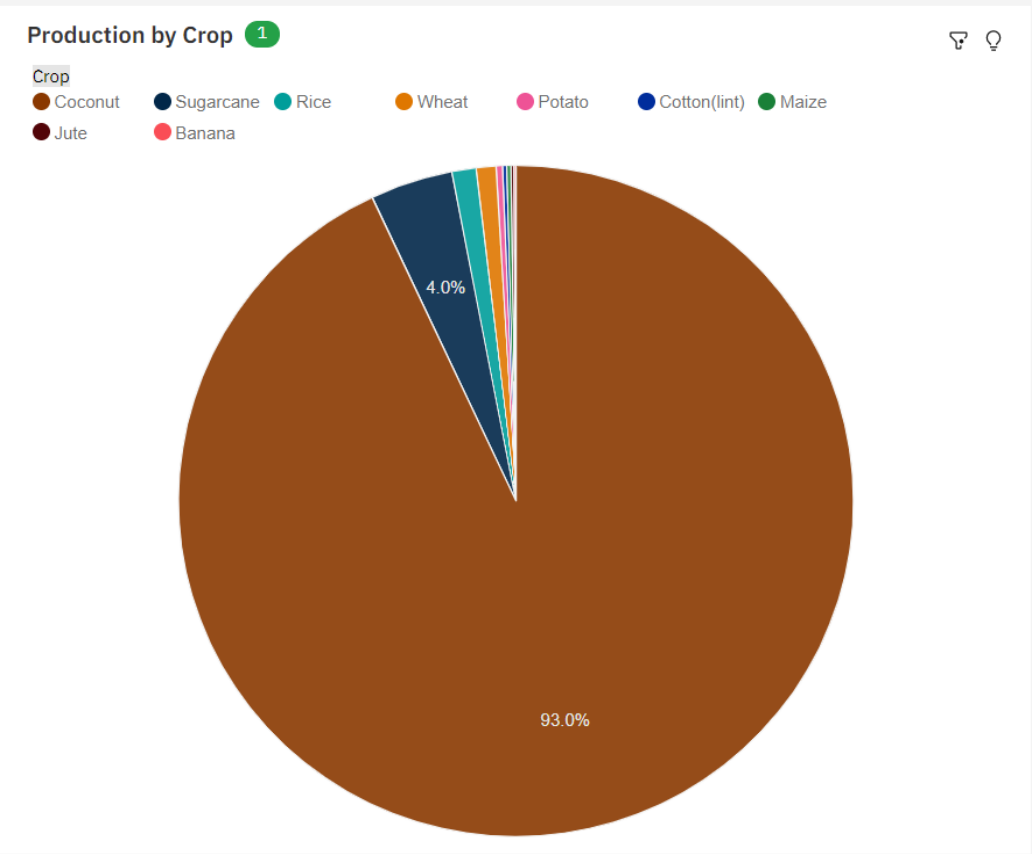
3.1. In the framework, we have proposed a trial of numerous Algorithms, and by contemplating the order report we analyze these Algorithms and pick the best one. It needs to discover the precision of the training dataset, accuracy of the testing dataset, particularly, False Positive rate, precision , and review by comparing algorithms utilizing python code.

4. FLOWCHART

4.1. The sum of **Production** for all values of **State_Name** is 138,935,794,024.The value of **Production** is unusually high when **State_Name** is Kerala.The summed values of **Production** range from a minimum of 448,840,739 (when **State_Name** is Madhya Pradesh) to a maximum of 97,880,045,376 (when **State_Name** is Kerala).

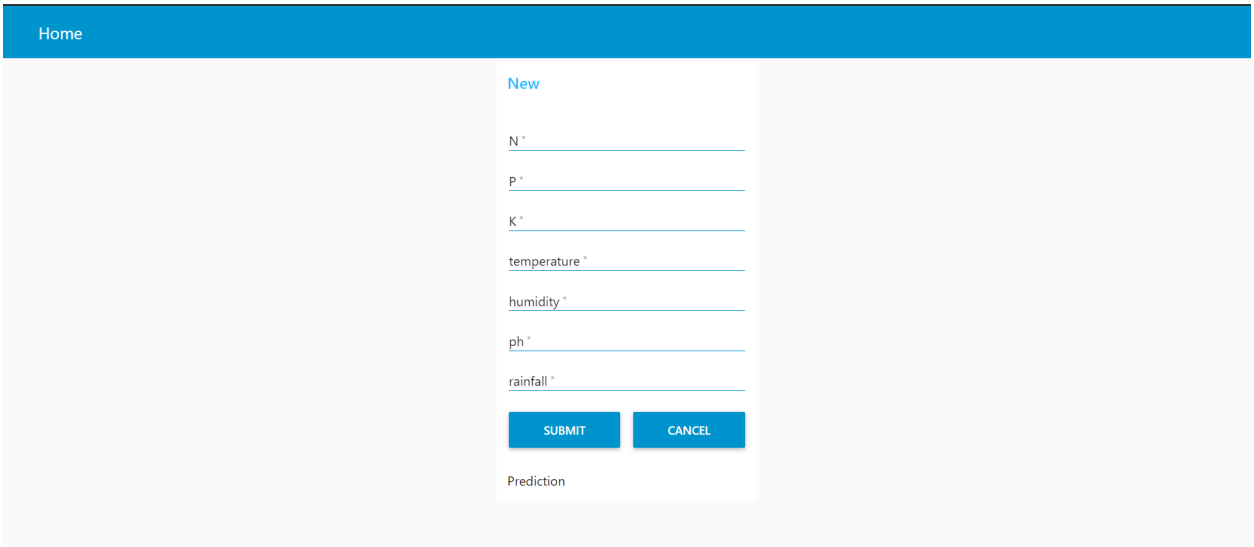


4.2. The sum of **Production** for all values of **Crop** is 139,778,465,803.The value of **Production** is unusually high when **Crop** is Coconut.The summed values of **Production** range from a minimum of 146,132,680 (when **Crop** is Banana) to a maximum of 129,981,629,216 (when **Crop** is Coconut).

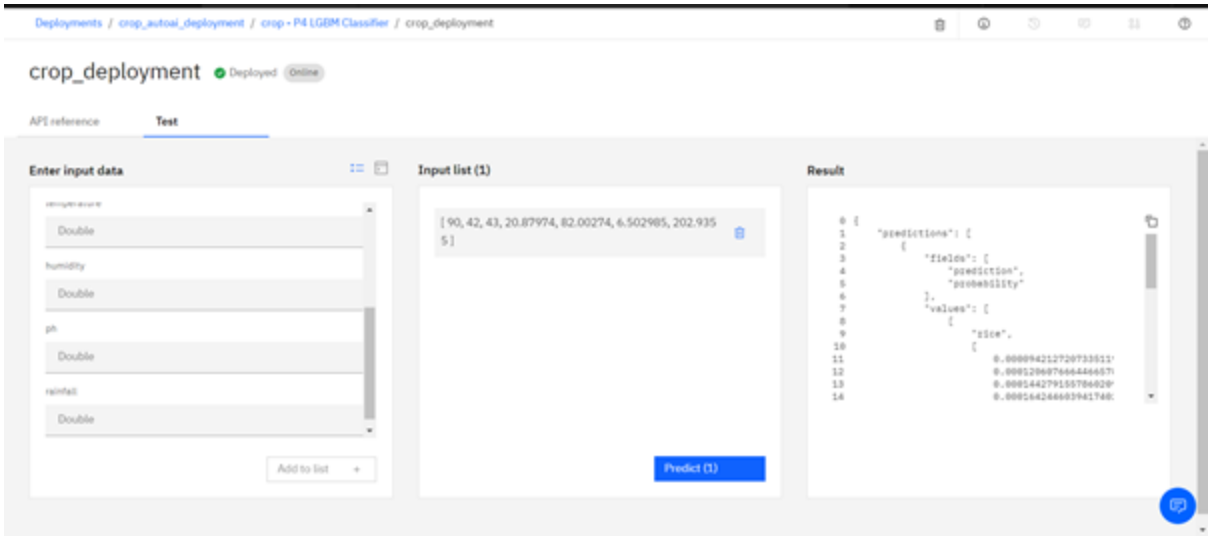


5. RESULT

For the crop recommendation system we used LGBM classifier to recommend crops which takes the inputs as follows



And has shown to give the crop which will suit the conditions. Following is an example of the output that was predicted by the model.



6. ADVANTAGES & DISADVANTAGES

6.1. Given advantages :

6.1.1. Providing an interface where our system will recommend to the farmers which crops he should cultivate in order to maximize his/her productivity and profit.

6.1.2. Providing an initial level of technological edge to smallholders who are unable to use or have the technological support to improve their yields.

6.1.3. Early crop prediction will give the farmers time to equip themselves with the right technology that would be useful to them in improving their harvest yield.

6.1.4. Through our model we would be able to give them the information that they need to maximize their profits.

6.2. The model's dataset relies upon the harvest's greatest expected yield, and average climates and primarily produced crops in an region which may not be significant under the following overlooked situations:

6.2.1. Any befuddle in the yield's environment and soil transformations contrasted with the real environment/soil conditions that it's filled in.

6.2.2. The harvest's compost proficiency and different variables may increase or lessen the powerful manure productivity.

6.2.3. Bugs and infections contrasted with the yield's overall opposition and perseverance against those illnesses.

6.2.4. Other pressure factors like dry season pressure, flood pressure, and whatever temperature conditions may win during a given year, contrasted with the climate averages.

7. APPLICATIONS (ad/av)

7.1. The areas where this solution can be applied:

7.1.1. Predicting the harvest yield is a fundamental duty of the leaders at **public and provincial sectors** (e.g., the EU level) for quick dynamics. A precise harvest yield prediction model can assist people with settling on what to develop and when to develop.

7.1.2. Analyzing all these issues and problems like weather, temperature and other factors, to develop proper technologies to overcome the situation faced by the **farmers and smallholders**. Thus lowering the losses and boosting the efficiency and profits in the sector.

8. CONCLUSION

We have considered 2 datasets. One finds out the profit and classifies it if there is profit or loss. The second dataset predicts the production

datainput - DataFrame

Index	Crop	State	ıltivation (`/Hecta	Cultivation (`/Hec	Production (`/Qui	ld (Quintal/ Hecta	Support price
0	ARHAR	Uttar Pradesh	9794.05	23076.7	1941.55	9.83	6000
1	ARHAR	Karnataka	10593.1	16528.7	2172.46	7.47	6000
2	ARHAR	Gujarat	13468.8	19551.9	1898.3	9.59	6000
3	ARHAR	Andhra Pradesh	17051.7	24171.7	3670.54	6.42	6000
4	ARHAR	Maharashtra	17130.5	25270.3	2775.8	8.72	6000
5	COTTON	Maharashtra	23711.4	33116.8	2539.47	12.69	5515
6	COTTON	Punjab	29047.1	50828.8	2003.76	24.39	5515
7	COTTON	Andhra Pradesh	29140.8	44756.7	2509.99	17.83	5515
8	COTTON	Gujarat	29616.1	42070.4	2179.26	19.05	5515
9	COTTON	Haryana	29919	44018.2	2127.35	19.9	5515

We combined data from different sources.
The data contains columns: - Crops - State - Cost of Cultivation (`/Hectare) A2+FL - Cost of Cultivation (`/Hectare) C2 - Cost of Production (`/Quintal) C2 - Yield produced.

The profit for each row was calculated using the formula C1 -> Cost of cultivation(`/Hectare) A2+FL C2 -> Cost of Cultivation (`/Hectare) C2 Cp -> Cost of Production (`/Quintal) Profit = (Yield *Support Price) - (C1 + C2 + (Yield*Cp))

The govt. fixes support prices[2] per Quintal for various commodities, for example various Kharif and Rabi crops. If the yield produced will result in profit based on support prices declared by the government, class 1 was allotted; else it was classified as class 0.

Advantages - This dataset is compiled by using data from an official government site which proves its authenticity. - Farmers can directly find out if the crop they are about to sow will result in profit after cultivation Disadvantage - Does not have many instances

9. FUTURE SCOPE

The scope to improve the efficiency of this model is immense. We can add some parameters that we feel are important and help us make the model even more efficient. Deciding on the parameters would help us to collect data even more effectively and make our model better. One of the things that we can do is also connect a weather api which can give forecasts to the farmers and they can use it to plan how to plant their crops after predicting the weather by themselves.

10. BIBLIOGRAPHY

11.

11.1. https://data.gov.in/

11.2. https://farmer.gov.in/mspstatements.aspx

11.3. N. P. Sastra and D. M. Wiharta, —Environmental monitoring as an IoT application in building smart campus of UniversitasUdayana,‖ in Proc. Int. Conf. Smart Green Technol. Elect. Inf. Syst. (ICSGTEIS), Oct. 2016, pp. 85–88.

11.4. M. Suganya., Dayana R and Revathi.R, Crop Yield Prediction Using Supervised Learning Techniques, International Journal of Computer Engineering and Technology, 11(2), 2020, pp. 9-20

11.5. CH. Vishnu Vardhanchowdary, Dr.K.Venkataramana, Tomato Crop Yield Prediction using ID3, March 2018,IJIRT Volume 4 Issue 10 pp,663-62.

11.6. R. Sujatha and P. Isakki, A study on crop yield forecasting using classification techniques 2016 International Conference on Computing Technologies and Intelligent Data Engineering (ICCTIDE'16), Kovilpatti, 2016, pp. 1-4.

11.7. N. Gandhi, L. J. Armstrong, O. Petkar and A. K. Tripathy, Rice crop yield prediction in India using support vector machines 2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE), KhonKaen, 2016, pp. 1-5

12. APPENDIX A. Source Code Attach the code for the solution built.