

# **Profitability and Yield prediction on agricultural crops of India**

—

# INTRODUCTION & MOTIVATION

Across the globe India is the second largest country having a population of more than 1.3 Billion. Many people are dependent on agriculture but the sector lacks efficiency and technology especially in our country. By bridging the gap between traditional agriculture and data science, effective crop cultivation can be achieved.

It's is very important to have a good yield of crops. Due to factors like soil type, precipitation, seed quality, lack of technical facilities etc. the crop yield is directly influenced.

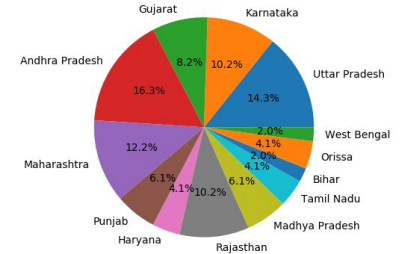
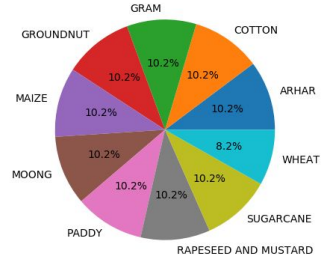
We focus on implementing crop yield and profitability prediction systems by using Machine learning techniques on agriculture/crop dataset.

## Objectives

- ❑ 1. Data Preprocessing
- ❑ 2. Data Visualization
- ❑ 3. Using various algorithms and comparing the accuracy

# DATASET-1

X - DataFrame							Profit - NumPy array	
Index	Crop	State	Cost of Cultivation (/Hectare) A2+FL	Cost of Cultivation (/Hectare) C2	Cost of Production (/Quintal) C2	Support price		
0	ARHAR	Uttar Pradesh	9794.05	23076.7	1941.55	6000	0	1
1	ARHAR	Karnataka	10593.1	16528.7	2172.46	6000	1	1
2	ARHAR	Gujarat	13468.8	19551.9	1898.3	6000	2	1
3	ARHAR	Andhra Pradesh	17051.7	24171.7	3670.54	6000	3	0
4	ARHAR	Maharashtra	17130.5	25270.3	2775.8	6000	4	0
5	COTTON	Maharashtra	23711.4	33116.8	2539.47	5515	5	0
6	COTTON	Punjab	29047.1	50828.8	2003.76	5515	6	1
7	COTTON	Andhra Pradesh	29140.8	44756.7	2509.99	5515	7	0
8	COTTON	Gujarat	29616.1	42070.4	2179.26	5515	8	0
9	COTTON	Haryana	29919	44018.2	2127.35	5515	9	0
10	GRAM	Rajasthan	8552.69	12610.9	1691.66	5100	10	1



Dataset 1 - Classifies it if would be profit or loss.

# DATASET-2

datainput - DataFrame

Index	State_Name	District_Name	Crop_Year	Season	Crop	Area	Production
0	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Arecanut	1254	2000
1	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Other Kharif pulses	2	1
2	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Rice	102	321
3	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Banana	176	641
4	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Cashewnut	720	165
5	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Coconut	18168	6.51e+07
6	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Dry ginger	36	100
7	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Sugarcane	1	2
8	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Sweet potato	5	15
9	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Tapioca	40	169
10	Andaman and Nicobar Islands	NICOBARS	2001	Kharif	Arecanut	1254	2061
11	Andaman and Nicobar Islands	NICOBARS	2001	Kharif	Other Kharif pulses	2	1

Dataset 2 - The second dataset predicts the production.

# PROPOSED SOLUTION

- ❑ The system we propose, tests many algorithms and by studying the classification report we compare the algorithms and choose the best one.
- ❑ It has to find Accuracy of the training dataset, Accuracy of the testing dataset, Specification, False Positive rate, precision and recall by comparing algorithms using python code.
- ❑ The following proposed steps are,
  1. Define a problem
  2. Preparing data
  3. Evaluating algorithms
  4. Improving results
  5. Predicting results

# RESULTS

## Results for Dataset 1

Algorithm	Precision		Recall		F1 Score		Accuracy
	Class 0	Class 1	Class 0	Class 1	Class 0	Class 1	
Logistic Regression	1.0	0.89	0.86	1.0	0.92	0.94	0.93
Decision Tree	0.78	1.0	1.0	0.75	0.88	0.86	0.87
Random forest	0.86	0.75	0.75	0.86	0.80	0.80	0.80
K nearest	0.50	0.86	0.75	0.67	0.60	0.75	0.69

# RESULTS

## Results for Dataset 2

Algorithm	R2 score	Mean absolute error
Decision Tree	0.84	167163.3086041714
Random Forest	0.91	155503.99436675265

# References

[1] <https://data.gov.in/>

[2] <https://farmer.gov.in/mspstatements.aspx>

[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.

[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



Thank You!