प्रकृति 2021
agri and food innovation fest

# Problem Statement : Data Analytics

## Dataset

The greatest challenge coming ahead for the world is to meet the growing food demand. We must bring new innovations that can supply ample '**Food for Future**' in an eco-friendly way. In order to develop strategies to ensure food security in the future , it is very important to understand the current food habits of the people.

The attached dataset contains information about some important regional foods of India. The information about their ingredients, preparation time, cooking time, flavor, food course type, region and price per unit is provided in the dataset.

प्रकृति 2021
agri and food innovation fest

# QUESTIONS

1. Use the appropriate method to input the missing data.
2. Split the dataset into training and test set (take test size = 20%)
2. Classify the food course type using all other variables except name via random forest, support vector machine, and gradient boosting.
3. Produce the confusion matrix, overall accuracy, producer's accuracy, users' accuracy, and kappa coefficient of the classification using all three methods.
4. If we classify flavor using all other variables except name, how much increase or decrease in classification accuracy we will get as compared to the previous classification? Produce confusion matrix, user's accuracy, producer's accuracy, and kappa coefficient using all three methods.

**Notes**

- Be systematic and give some reasons behind every task that you perform, make sure use comments to highlight your logics.
- If possible, incorporate some visualization may be using some simple plots.

प्रकृति 2021
agri and food innovation fest

# Rules

- Only teams whose members have registered online would be able to participate as teams.
- Each team can contain a maximum of four(4) participants.
- Students from different colleges can form a single team.
- All Teams must submit the following files to this google form link before the deadline (17th March,11:59 pm) from the beginning of the event.

1) Code: In any language

2)Presentation and appendix( word/PDF document) containing the technical explanation in short.

- All the reports and presentations submitted by the team must contain the team name, team members name, their contact number, and the name of the college they belong to.
- On the basis of abstract ,teams will be selected for the second round(Online Presentation and Question and Answer round on 21st March) .

प्रकृति 2021
agri and food innovation fest

- The presentation time must not exceed 6 minutes followed by the Q&A round of 2minutes, exceeding this may yield negative points.
- The whole analysis of the problem statement should be done by the participants themselves. Plagiarism will lead to the cancellation of participation.
- The decision of judges is final and binding. Team Prakriti reserves the right to disqualify any team in case of violation of any of the above rules.
- Total Prize Money: INR 8,000.