

## **Project 1: Exploratory Data Analysis on Indian Food and Cuisine Dataset**

Indian cuisine consists of a variety of regional and traditional cuisines native to the Indian subcontinent. Given the diversity in soil, climate, culture, ethnic groups, and occupations, these cuisines vary substantially and use locally available spices, herbs, vegetables, and fruits. Indian food is also heavily influenced by religion, particularly Hinduism, cultural choices and traditions.

This dataset consists of information about various **Indian dishes**, their **ingredients**, their **place of origin**, etc.

### **Column Description:**

**name:** name of the dish

**ingredients:** main ingredients used

**diet:** type of diet - either vegetarian or non-vegetarian

**prep time:** preparation time

**cook time:** cooking time

**flavor profile:** flavor profile includes whether the dish is spicy, sweet, bitter, etc.

**course:** course of meal - starter, main course, dessert, etc.

**state:** state where the dish is famous or is originated

**region:** region where the state belongs

The presence of -1 in any of the columns indicates NaN value.

### **KPIs:**

- Visualize the preferred flavors of Indian dishes by customers using a Bar graph
- Visualize the preferred flavors of Indian dishes by customers using a Pie chart
- Visualize the top 10 most requested meals
- Visualize the total vegan and non-vegan orders by different regions

## **Project 2: Image Classification using “Alien and Predators” images**

The "Alien vs Predator" image classification dataset is used to build a machine learning model capable of distinguishing between images of aliens and predators. This binary classification problem involves training a neural network on labeled images to identify the category of unseen images correctly.

The dataset comprises images in JPG format, with varying thumbnail sizes, typically around 250 x 250 pixels.

The dataset is organized into three main sets:

Training Set: Used to train the machine learning model.

- Aliens: 247 images of aliens
- Predators: 247 images of predators

Validation Set: Used to tune the model and evaluate its performance during training.

- Aliens: 100 images of alien
- Predators: 100 images of predators

## **Project 3: Multiclass classification of Liver cirrhosis dataset**

Create predictive models to predict the stage of liver Cirrhosis using 18 clinical features. Cirrhosis damages the liver from various causes leading to scarring and liver failure.

Hepatitis and chronic alcohol abuse are frequent causes of the disease. Liver damage caused by cirrhosis can't be undone, but further damage can be limited. Treatments focus on the underlying cause. In advanced cases, a liver transplant may be required. Predicting the stage of cirrhosis and beginning the treatment before it's too late can prevent the fatal consequences of the disease.

- Train Dataset: It has 6801 data points.
- Test Dataset: You must predict the stage of cirrhosis of 3201 datapoints.

Column description:

ID: Unique Identifier

N\_Days: number of days between registration and the earlier of death, transplantation, or study analysis time.

Status: status of the patient C (censored), CL (censored due to liver tx), or D (death)

Drug: type of drug. D-penicillamine or placebo

Age: age in [days]

Sex: M (male) or F (female)

Ascites: presence of ascites N (No) or Y (Yes)

Hepatomegaly: the presence of hepatomegaly N (No) or Y (Yes)

Spiders: the presence of spiders N (No) or Y (Yes)

Edema: the presence of edema N (no edema and no diuretic therapy for edema), S (edema present without diuretics, or edema resolved by diuretics), or Y (edema despite diuretic therapy)

Bilirubin: serum bilirubin in [mg/dl]

Cholesterol: serum cholesterol in [mg/dl]

Albumin: albumin in [gm/dl]

Copper: urine copper in [ug/day]

Alk\_Phos: alkaline phosphatase in [U/liter]

SGOT: SGOT in [U/ml]

Triglycerides: triglycerides in [mg/dl]

Platelets: platelets per cubic [ml/1000]

Prothrombin: prothrombin time in seconds [s]

Stage: histologic stage of disease (1, 2, 3, or 4)

By-

Aashutosh Kumar Mishra

Data Science & AI Consultant and Instructor