DA Project – Problem Statement and Queries

Name: Anway Pimpalkar

MIS: 111907066 | **Batch:** D

Branch: Electronics and Telecommunications

Problem Statement:

Classifying mushrooms as edible and poisonous based on the UCI Mushroom Dataset.

Visualization Queries:

1. Number of poisonous and edible mushrooms in the dataset.
2. Visualize the distribution between various habitats and the edibility of mushrooms.
3. Population parameters of the instances of edible and poisonous mushrooms.
4. Plot a treemap showing the distribution of different gill colors of mushrooms.
5. Visualize the occurrence of different ring types, and the number of rings on such mushrooms.

Machine Learning Queries:

1. Test the accuracy of various classification models on the dataset to build a accurate prediction model for edibility of mushrooms.
2. Run the most accurate method of classification on an unseen dataset and cross-check it's accuracy.