RAVINDRA V

DPR (detailed project report)

PROJECT : MUSHROOM CLASSIFICATION

OBJECTIVE:

To build a model to predict whether a mushroom is edible or poisonous , by looking at 23 species of gilled mushrooms we have to identify whether a mushroom is edible or poisonous

BENEFITS :

- we will be able to detect the poisonous mushroom
- we can look at different features and components of mushroom and predict that mushroom is edible
- we can find most indicative of a poisonous mushroom

DSA(DATA SHARING AGREEMENT):

- This dataset includes descriptions of hypothetical samples corresponding to 23 species of gilled mushrooms in the Agaricus and Lepiota Family Mushroom drawn from The Audubon Society Field Guide to North American Mushrooms (1981)

- ['class', 'cap-shape', 'cap-surface', 'cap-color', 'bruises', 'odor',
  'gill-attachment', 'gill-spacing', 'gill-size', 'gill-color',
  'stalk-shape', 'stalk-root', 'stalk-surface-above-ring',
  'stalk-surface-below-ring', 'stalk-color-above-ring',
  'stalk-color-below-ring', 'veil-type', 'veil-color', 'ring-number',
  'ring-type', 'spore-print-color', 'population', 'habitat']

- each of the columns in above should be provided for model training and prediction

DATA VALIDATION:

- first we have to replace all the alphabets provided in csv file into readable format
- if any of the columns contains null values or missing values than that data is not accepted

MODEL TRAINING:

- The data contains 8124 rows and 23 columns , the data is in csv format and this data is used for model
- After replacing the every single alphabets in each row into readable format we have to convert the data into numeric form

DATA PRE-PROCESSING:

- By performing EDA we will be able to identify and predict the features and labels ,so that we will be able to find in every column which feature are indicative to poisonous mushrooms
- after data pre-processing we have to find the best model for prediction in classification algorithms

PREDICTION:

- after finding the best model we will be able to fit the model to training and testing the data so we we can get best accuracy score
- after classifying different features into testing model we will be able to predict whether the mushroom in the classification is edible or poisonous