# PROJECT : MUSHROOM CLASSIFICATION

## HIGH LEVEL DESIGN

By *Suyash Waykar*

# Introduction :

Mushroom classification is a technique by which we can classify whether the mushroom species is edible or poisonous.

# Problem Statement :

The Audubon Society Field Guide to North American Mushrooms contains descriptions of hypothetical samples corresponding to 23 species of gilled mushrooms in the Agaricus and Lepiota Family Mushroom (1981). Each species is labelled as either definitely edible, definitely poisonous, or maybe edible but not recommended. This last category was merged with the toxic category. The Guide asserts unequivocally that there is no simple rule for judging a mushroom's edibility, such as "leaflets three, leave it be" for Poisonous Oak and Ivy. The main goal is to predict which mushroom is poisonous & which is edible.

# Approach :

The classical machine learning tasks like Data Exploration, Data Cleaning, Feature Engineering, Model Building and Model Testing as been done on the project. Tried with different machine learning algorithms such as Logistic Regression, SVM, Gradient Boosting, KNN, Random Forest and found out best fit model in the project as Random Forest.

# Dataset :

The dataset includes categorical characteristics on 8,124 mushroom samples from various species of gilled mushrooms.

- The target variable assessed was a class distinction of 'edible' or 'poisonous'.
- The explanatory variables covered a range of descriptive and visual characteristics on the structure of each observed mushroom - such as, cap color, odor, ring number and stalk shape.

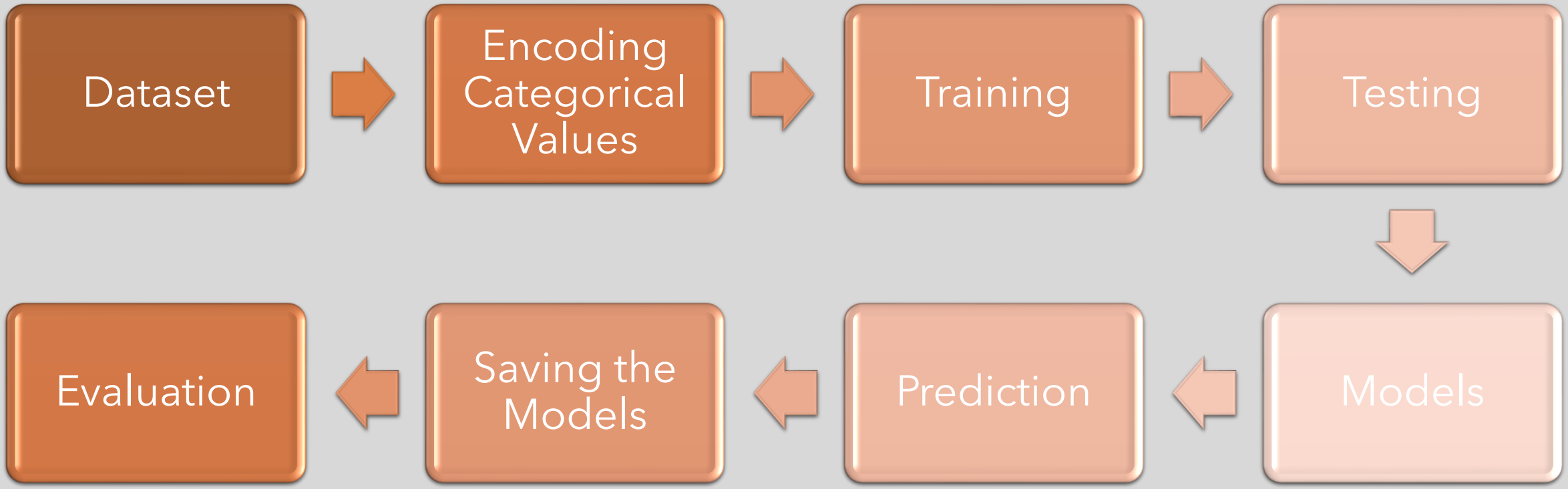
# Tools Used :

Python Programming languages and libraries such as –

- NumPy
- Pandas
- Matplotlib
- Seaborn
- Scikit learn

were used to build the whole model and Flask were used for a web framework.

# Design Flow :


Dataset
Encoding Categorical Values
Training
Testing
Models
Prediction
Saving the Models
Evaluation

# Conclusion :

Our tuned classification models, all performed really well with the dataset. Random Forest Classifier, which had a score of 99% was a great choice & much better suited to classify mushrooms.

Since our models performed so well, it was clear to us that they were able to identify specific traits that greatly influenced the classification of an edible versus poisonous mushroom. And that was exactly what we were hoping for!!

*Thanks!*