INEURON.AI

MUSHROOM CLASSIFICATION REPORT

BY SAEED SHAIKH

30-08-2023

❖ Abstract

This study explores the use of machine learning for mushroom classification. Traditional methods are time-consuming and require expert knowledge. The researchers collected a dataset of mushroom samples, labeled them with their classes, and used feature extraction techniques. Various machine learning models were trained and evaluated, demonstrating higher accuracy compared to traditional methods. The results highlight the potential for automated mushroom identification and quality control in the food industry.

❖ Objective:

➢ Development of a predictive model for Mushroom classification determine whether a mushroom is edible or poisonous .

❖ Benefits:

➢ Detection of upcoming poisonous mushrooms.

➢ Gives better insight of edible mushrooms.

➢ Detection of upcoming poisonous mushrooms.

❖ Data Description

cap-shape: bell=b,conical=c,convex=x,flat=f, knobbed=k,sunken=s

cap-surface: fibrous=f,grooves=g,scaly=y,smooth=s

cap-color: brown=n,buff=b,cinnamon=c,gray=g,green=r,pink=p,purple=u,red=e,white=w,yellow=y

bruises: bruises=t,no=f

odor: almond=a,anise=l,creosote=c,fishy=y,foul=f,musty=m,none=n,pungent=p,spicy=s

gill-attachment: attached=a,descending=d,free=f,notched=n

gill-spacing: close=c,crowded=w,distant=d

gill-size: broad=b,narrow=n

gill-color: black=k,brown=n,buff=b,chocolate=h,gray=g,
green=r,orange=o,pink=p,purple=u,red=e,white=w,yellow=y

talk-shape: enlarging=e,tapering=t

stalk-root: bulbous=b,club=c,cup=u,equal=e,rhizomorphs=z,rooted=r,missing=?

Stalk-surface-above-ring: fibrous=f,scaly=y,silky=k,smooth=s

stalk-surface-below-ring: fibrous=f,scaly=y,silky=k,smooth=s

stalk-color-above-ring:brown=n,buff=b,cinnamon=c,gray=g,orange=o,pink=p,red=e,white=w,y
ellow=y

stalk-color-below-ring:
brown=n,buff=b,cinnamon=c,gray=g,orange=o,pink=p,red=e,white=w,yellow=y

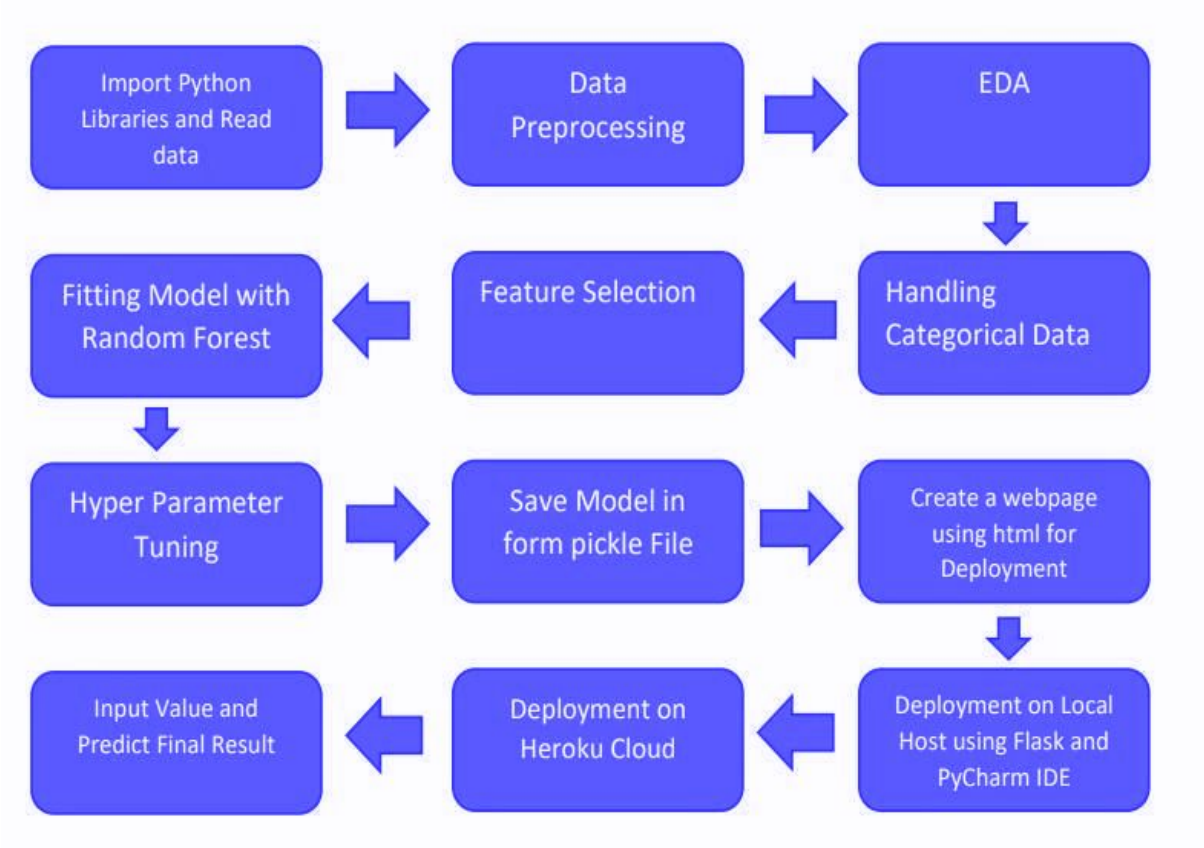
veil-color: brown=n,orange=o,white=w,yellow=y
ing-number: none=n,one=o,two=t

ring-type: cobwebby=c,evanescent=e,flaring=f,large=l,none=n,pendant=p,sheathing=s,zone=z

spore-print-color:
black=k,brown=n,buff=b,chocolate=h,green=r,orange=o,purple=u,white=w,yellow=y

population: abundant=a,clustered=c,numerous=n,scattered=s,several=v,solitary=y

❖ ARCHITECTURE



❖ Model Training:

- ➢ Data Preprocessing o Performing EDA to get insight of data like identifying distribution , outliers, trend among data etc.
- ➢ Check for null values in the columns. If present impute the null values.
- ➢ Encode the categorical values with numeric values.
- ➢ Perform Standard Scalar to scale down the values.

❖ Feature Selection:

- ➢ Use feature selection techniques Relevant information as possible. This helps to simplify the classification task and avoid overfitting.

❖ Model Selection:

- ➢ Evaluates classification models using Logistic Regression ,Random forest , Decision Tree , through exhaustive Random search, and Area Under Curve (AUC) & F1 score.

- ➢ Compute metrics and generate graphs for model evaluation and importance analysis.

- ➢ We view AUC and F1 score values for each model and found out the Gradient boosting classifier perform well

- ➢ Finally, we fit the Gradient Boosting Classifier model with optimal tuning parameters on the entire dataset. We then could use this model to predict whether Mushroom is edible or Poisonous .

- **Prediction :**

  - **The testing file is used and we perform the same validation operations, data transformation and data insertion on them.**

  - **The accumulated data from database is exported in csv format for prediction.**

  - **We perform data pre-processing techniques in it.**

- **Deployment :**

  - **We have deployed the application on AWS elastic beanstalk.**