Capstone Project: Classification

Capstone Project

Classification

- Data source: https://www.kaggle.com/datasets/uciml/mushroom-classification

- Target:
  - Predict class
  - edible / poisonous
- 22 features

Capstone Project

Features

Name	Properties
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
stalk-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,yellow=y
stalk-color-below-ring:	brown=n,buff=b,cinnamon=c,gray=g,orange=o,pink=p,red=e,white=w,yellow=y
veil-type:	partial=p,universal=u
veil-color:	brown=n,orange=o,white=w,yellow=y
ring-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
habitat:	grasses=g,leaves=l,meadows=m,paths=p,urban=u,waste=w,woods=d

Capstone Project

Steps

Categorical Variable Treatment	Find a way to encode all categories
Data Prep	Separate dependent/independent, data splitting, …
Modeling	Check out different models
Model Evaluation	confusion matrix, ROC curve