

Dataset name: Mushrooms

Dataset file: mushrooms.csv

Response variable: class

Scenario

The practice of mushroom picking is still enjoyed by many people, but wild mushrooms can be very dangerous to eat without knowing exactly which are safe, and which are poisonous. Therefore, building a model to help in the identification of poisonous mushrooms, based on the features of the mushrooms themselves, may help people to stay safe and well. The dataset presents data on a variety of features representing the physical and environment characteristics of mushroom species. Your task is to build a classification model capable of predicting the class of the mushrooms (i.e., whether they are poisonous or edible), and to attempt to identify which features are most indicative of poisonous mushrooms.

Data Dictionary

Feature ID	Feature label	Category key	Data type	Response feature
cap-shape	The shape of the cap.	bell=b, conical=c, convex=x, flat=f, knobbed=k, sunken=s	Categorical	
cap-surface	The surface texture of the cap.	fibrous=f, grooves=g, scaly=y, smooth=s	Categorical	
cap-color	The colour of the cap.	brown=n, buff=b, cinnamon=c, gray=g, green=r, pink=p, purple=u, red=e, white=w, yellow=y	Categorical	
bruises	Are bruises present?	True = t, false, = f	Boolean	
odor	The odour of the mushroom.	almond=a, anise=l, creosote=c, fishy=y, foul=f, musty=m, none=n, pungent=p, spicy=s	Categorical	
gill-attachment	The attachment of the gills.	attached=a, descending=d, free=f, notched=n	Categorical	
gill-spacing	The spacing of the gills.	close=c, crowded=w, distant=d	Categorical	
gill-size	The size of the gills.	broad=b, narrow=n	Categorical	

gill-color	The colour of the gills.	black=k, brown=n, buff=b, chocolate=h, gray=g, green=r, orange=o, pink=p, purple=u, red=e, white=w, yellow=y	Categorical	
stalk-shape	The shape of the stalk.	enlarging=e, tapering=t	Categorical	
stalk-root	The type of root.	bulbous=b, club=c, cup=u, equal=e, rhizomorphs=z, rooted=r	Categorical	
stalk-surface-above-ring	The texture of the stalk surface (above the ring).	fibrous=f, scaly=y, silky=k, smooth=s	Categorical	
stalk-surface-below-ring	The texture of the stalk surface (below the ring).	fibrous=f, scaly=y, silky=k, smooth=s	Categorical	
stalk-color-above-ring	The colour of the stalk (above the ring).	brown=n, buff=b, cinnamon=c, gray=g, orange=o, pink=p, red=e, white=w, yellow=y	Categorical	
stalk-color-below-ring	The colour of the stalk (below the ring).	brown=n, buff=b, cinnamon=c, gray=g, orange=o, pink=p, red=e, white=w, yellow=y	Categorical	
veil-type	Type of veil.	partial=p, universal=u	Categorical	
veil-color	The colour of the veil.	brown=n, orange=o, white=w, yellow=y	Categorical	
ring-number	Number of rings present.	none=n, one=o, two=t	Categorical	
ring-type	Type of ring(s)	cobwebby=c, evanescent=e, flaring=f, large=l, none=n, pendant=p, sheathing=s, zone=z	Categorical	
spore-print-color	Colour of the spores.	black=k, brown=n, buff=b, chocolate=h, green=r, orange=o, purple=u, white=w, yellow=y	Categorical	
population	The pattern of the mushrooms in the location(s) found.	abundant=a, clustered=c, numerous=n, scattered=s, several=v, solitary=y	Categorical	
habitat	The habitat where the mushrooms are found.	grasses=g, leaves=l, meadows=m, paths=p, urban=u, waste=w, woods=d	Categorical	
class	Class of the mushroom; poisonous or edible.	edible = e, poisonous = p	Categorical	Yes