The aim of this study is to apply Explainable Artificial Intelligence (XAI) techniques to analyze data related to mushrooms. XAI allows for understanding how artificial intelligence models, such as classifiers, make decisions. This is particularly important in fields where the accuracy and reliability of predictions are crucial, such as in distinguishing edible mushrooms from poisonous ones.

Soon, we will ask you to interpret several visualizations (mainly charts) that show how artificial intelligence (AI) predicts whether a wild mushroom with certain characteristics is either edible, inedible, or poisonous.

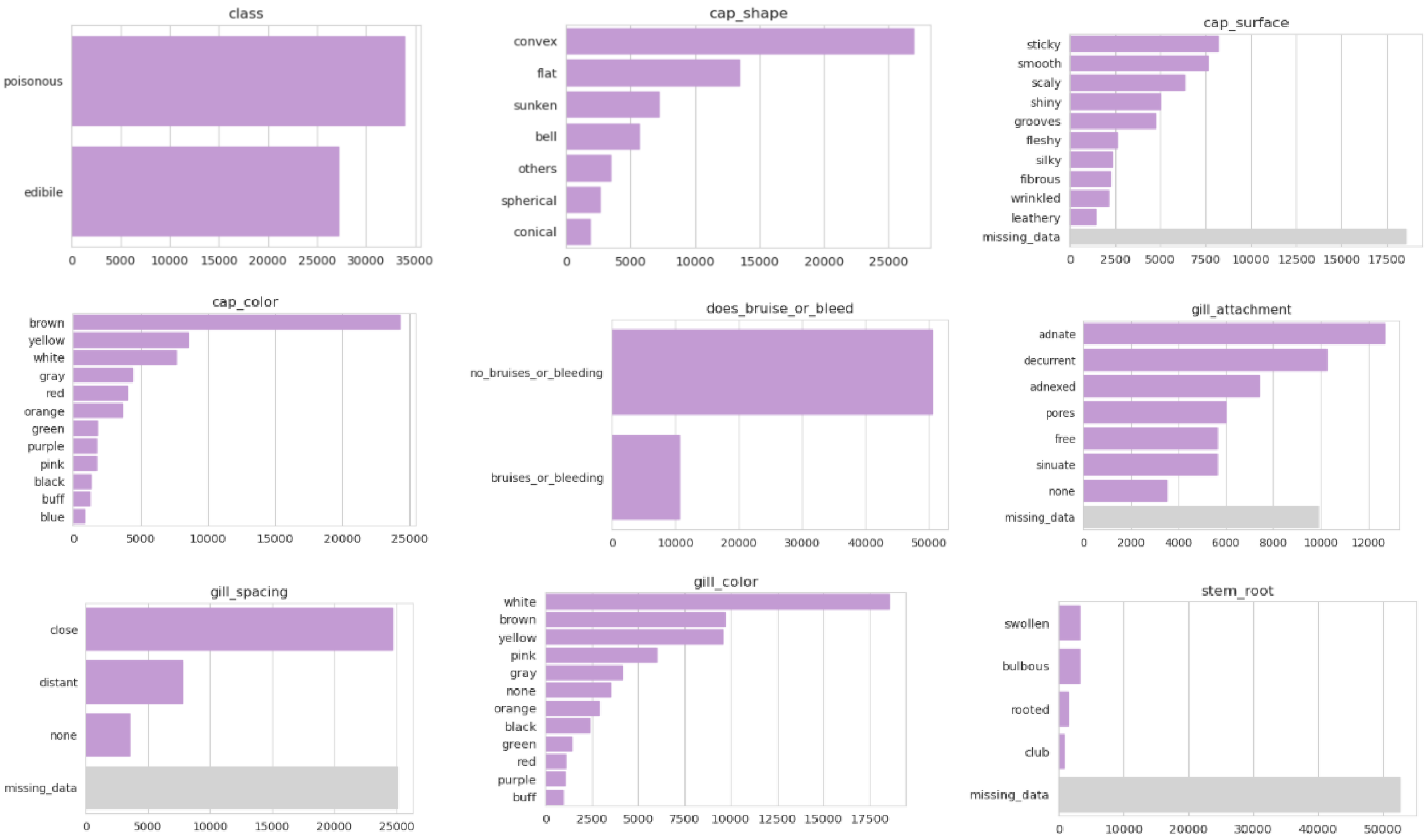
The original dataset that the AI worked on comes from the UC Irvine Machine Learning Repository, which is a collection of practice datasets from the University of California, Irvine, intended for training (improving) machine learning algorithms.

Link to this dataset:
https://archive.ics.uci.edu/dataset/848/secondary+mushroom+dataset

The dataset contains information about 61,069 fruiting bodies of 173 species of mushrooms, classified as either edible, non-edible, or poisonous. Mushrooms with unknown edibility were classified as non-edible or poisonous.

**These are exclusively cap mushrooms with a stem and lamellar hymenophore.**

Some of the data is simulated, hypothetical data, meaning artificially generated based on a smaller set of real observations of mushrooms found in nature.


class
poisonous
edibile
0 5000 10000 15000 20000 25000 30000 35000
cap_shape
convex
flat
sunken
bell
others
spherical
conical
0 5000 10000 15000 20000 25000
cap_surface
sticky
smooth
scaly
shiny
grooves
fleshy
silky
fibrous
wrinkled
leathery
missing_data
0 2500 5000 7500 10000 12500 15000 17500
cap_color
brown
yellow
white
gray
red
orange
green
purple
pink
black
buff
blue
0 5000 10000 15000 20000 25000
does_bruise_or_bleed
no_bruises_or_bleeding
bruises_or_bleeding
0 10000 20000 30000 40000 50000
gill_attachment
adnate
decurrent
adnexed
pores
free
sinuate
none
missing_data
0 2000 4000 6000 8000 10000 12000
gill_spacing
close
distant
none
missing_data
0 5000 10000 15000 20000 25000
gill_color
white
brown
yellow
pink
gray
none
orange
black
green
red
purple
buff
0 2500 5000 7500 10000 12500 15000 17500
stem_root
swollen
bulbous
rooted
club
missing_data
0 10000 20000 30000 40000 50000


gill_color
white
brown
yellow
pink
gray
none
orange
black
green
red
purple
buff
0 2500 5000 7500 10000 12500 15000 17500
stem_root
swollen
bulbous
rooted
club
missing_data
0 10000 20000 30000 40000 50000
stem_surface
smooth
scaly
fibrous
sticky
grooves
silky
none
shiny
missing_data
0 5000 10000 15000 20000 25000 30000 35000 40000
stem_color
white
brown
yellow
gray
orange
red
purple
none
pink
black
green
blue
buff
0 5000 10000 15000 20000
veil_type
universal
missing_data
0 10000 20000 30000 40000 50000 60000
veil_color
white
yellow
brown
purple
black
red
missing_data
0 10000 20000 30000 40000 50000
has_ring
none
ring
0 10000 20000 30000 40000
ring_type
none
evanescent
zone
large
flaring
pendant
grooved
movable
missing_data
0 10000 20000 30000 40000 50000
spore_print_color
black
pink
white
brown
gray
purple
green
missing_data
0 10000 20000 30000 40000 50000

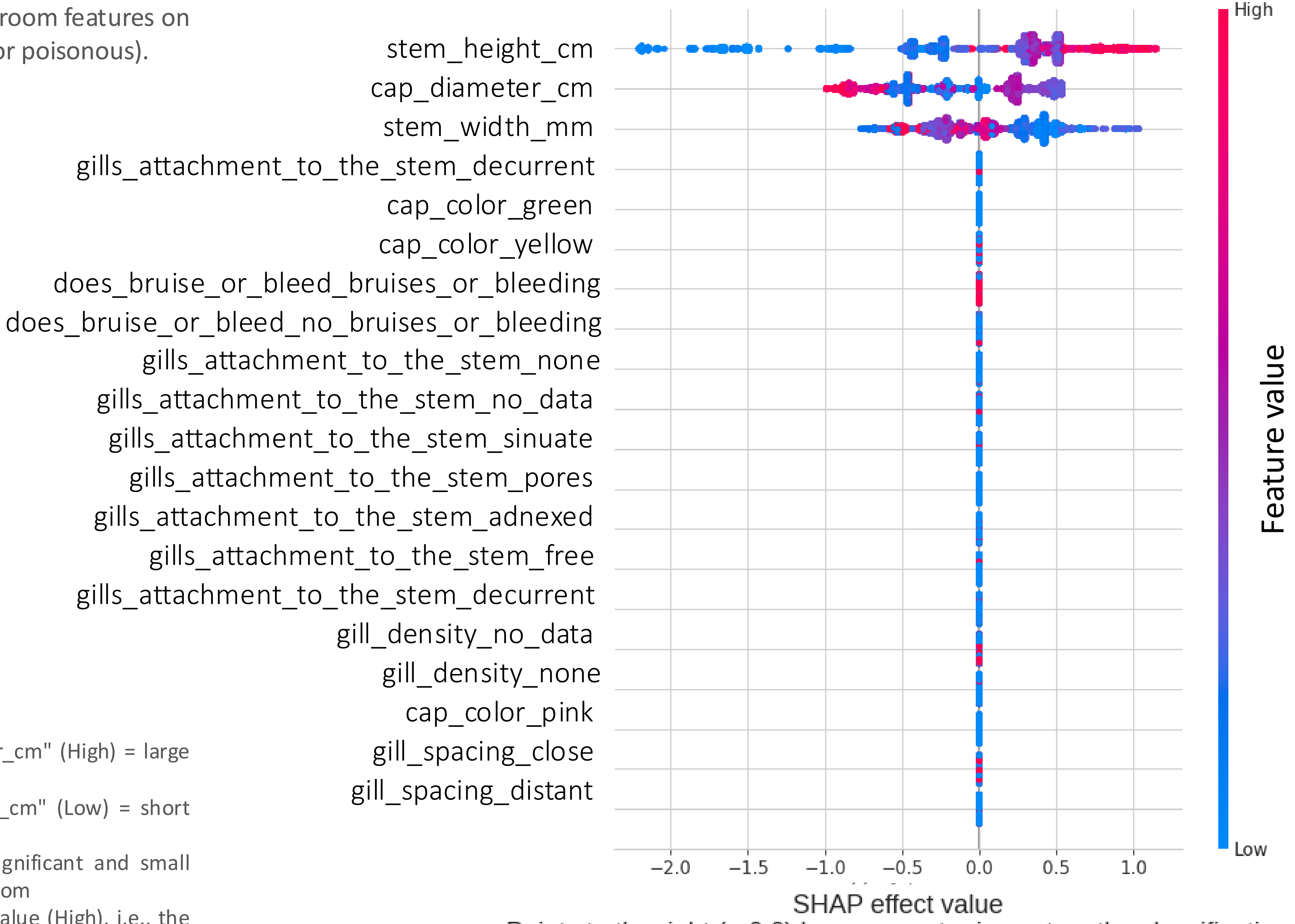
Descriptive statistics:

| | cap_diameter_cm | stem_height_cm | stem_width_mm |
|---|---|---|---|
| count | 61069.000000 | 61069.000000 | 61069.000000 |
| mean | 6.733854 | 6.581538 | 12.149410 |
| std | 5.264845 | 3.370017 | 10.035955 |
| min | 0.380000 | 0.000000 | 0.000000 |
| 25% | 3.480000 | 4.640000 | 5.210000 |
| 50% | 5.860000 | 5.950000 | 10.190000 |
| 75% | 8.540000 | 7.740000 | 16.570000 |
| max | 62.340000 | 33.920000 | 103.910000 |

Missing data (count and percent):

| | column: | count: | percent: |
|---|---|---|---|
| 14 | veil-type | 57892 | 94.797688 |
| 18 | spore-print-color | 54715 | 89.595376 |
| 15 | veil-color | 53656 | 87.861272 |
| 11 | stem-root | 52597 | 86.127168 |
| 12 | stem-surface | 38124 | 62.427746 |
| 7 | gill-spacing | 25063 | 41.040462 |
| 3 | cap-surface | 18552 | 30.378752 |
| 6 | gill-attachment | 9884 | 16.184971 |
| 17 | ring-type | 2471 | 4.046243 |
| 0 | class | 0 | 0.000000 |
| 13 | stem-color | 0 | 0.000000 |
| 19 | habitat | 0 | 0.000000 |
| 16 | has-ring | 0 | 0.000000 |
| 10 | stem_width_mm | 0 | 0.000000 |
| 1 | cap_diameter_cm | 0 | 0.000000 |
| 9 | stem_height_cm | 0 | 0.000000 |
| 8 | gill-color | 0 | 0.000000 |
| 5 | does-bruise-or-bleed | 0 | 0.000000 |
| 4 | cap-color | 0 | 0.000000 |
| 2 | cap-shape | 0 | 0.000000 |
| 20 | season | 0 | 0.000000 |

**"Bee swarm"** – the impact of individual mushroom features on the prediction of its edibility (edible/inedible or poisonous).



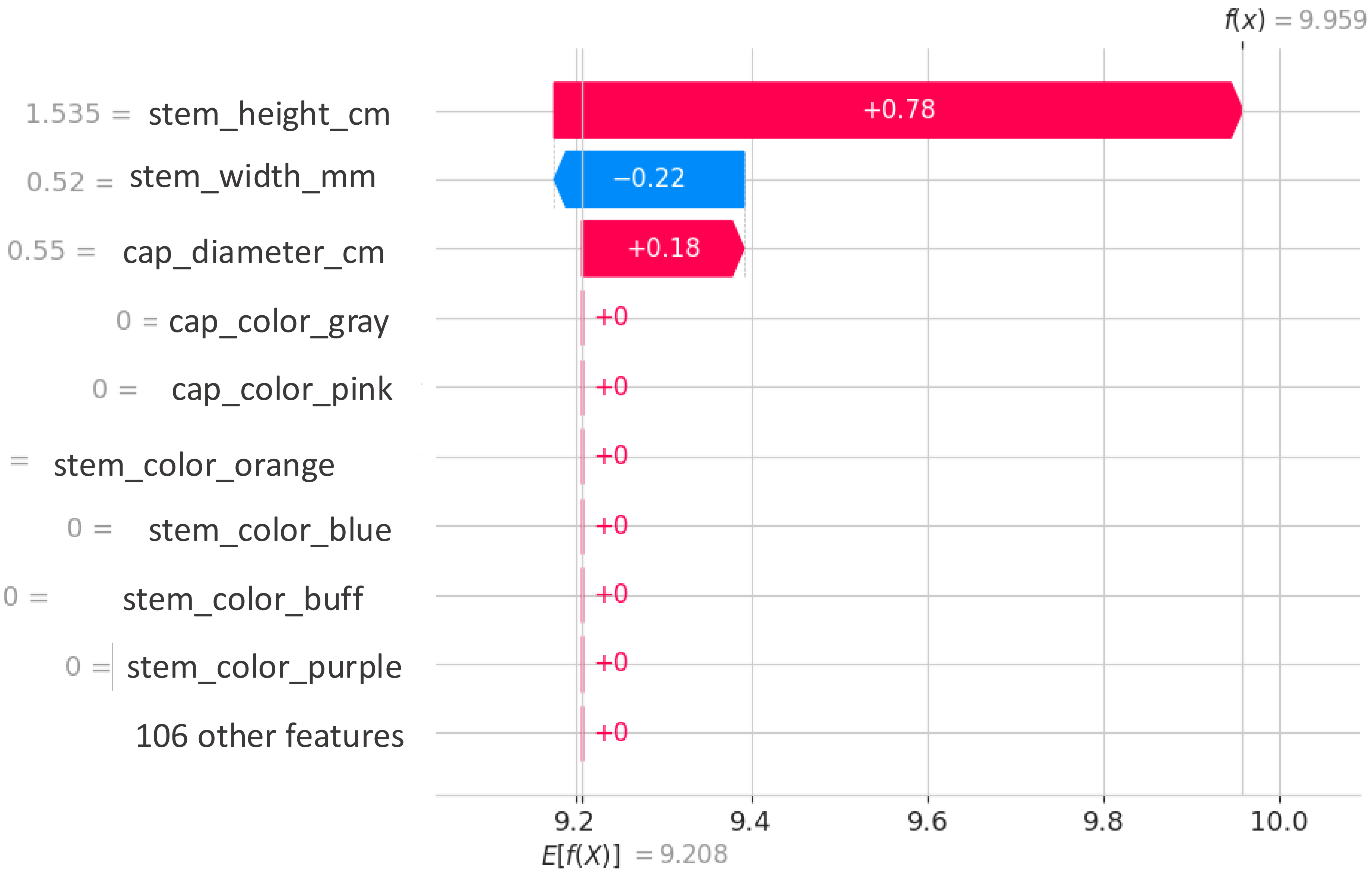

Points to the right (> 0.0) have a greater impact on the classification of the mushroom as poisonous by the model, to the left as edible

**Legend:**
High = high value of the feature, e.g., "cap_diameter_cm" (High) = large cap diameter in cm
Low = low value of the feature, e.g., "stem_height_cm" (Low) = short stem
"0" = the boundary between features that have significant and small importance in determining the edibility of the mushroom
For binary features (either present or absent), high value (High), i.e., the red color, means: "the feature is present."

"Waterfall" plot:
**the impact of the features of a given mushroom on the prediction of its toxicity.**

**Analysis of Feature Impact on Prediction for the 'Non-edible/Poisonous' Class (for a mushroom that is actually poisonous)**

The contribution of individual features to the model's prediction of the mushroom's class.
Positive values (to the right) indicate an increase in the probability of classification as "non-edible/poisonous" according to the model, while negative values (to the left) indicate a decrease.
**E[f(x)]** represents the average model output, and **f(X)** is the prediction for this specific observation.

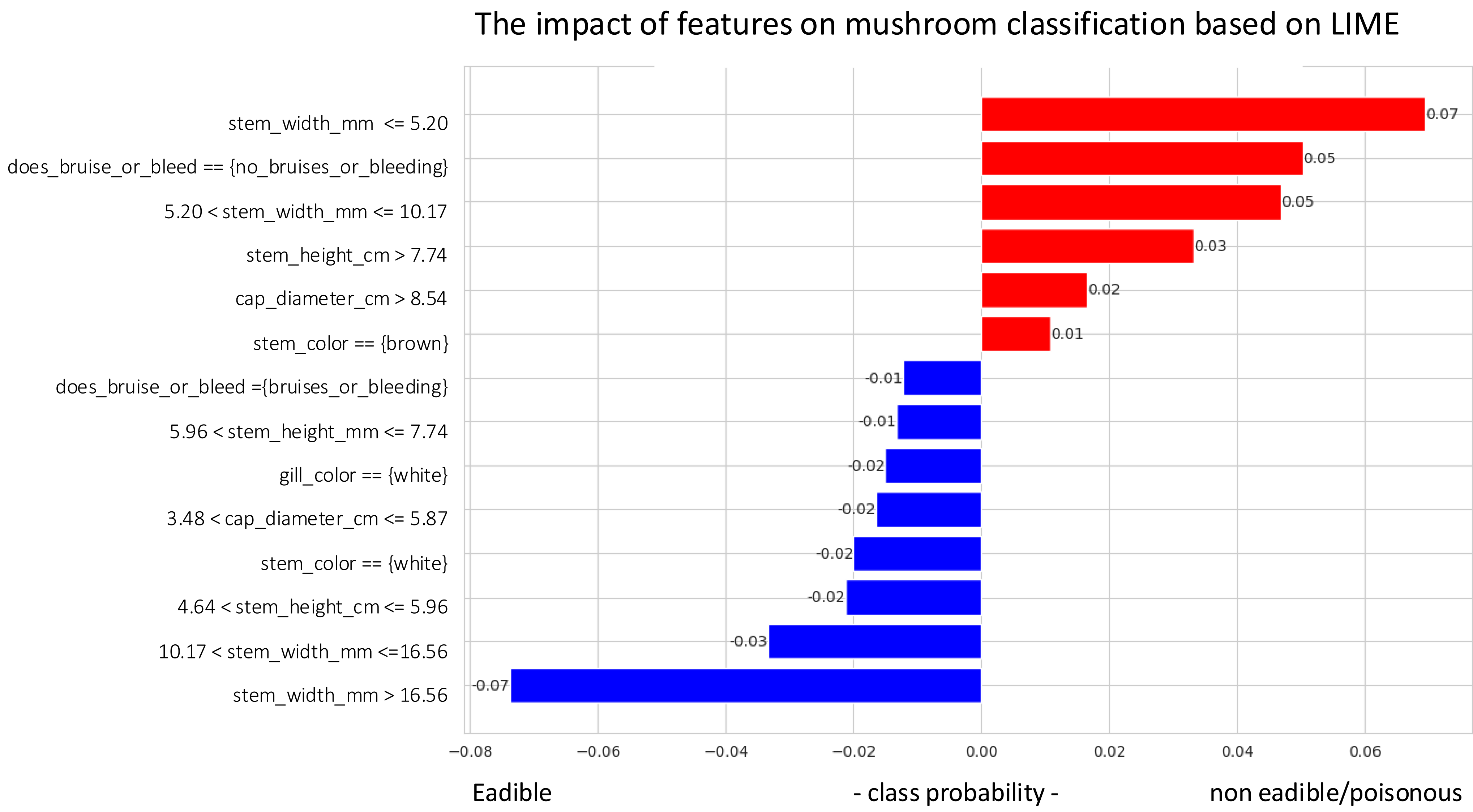



Legend:

E[f(X)]: baseline value = the average model prediction for all observations
Colored bands: the impact of a given feature on the toxicity prediction for a specific mushroom

f(x): final value = the model's prediction for this specific mushroom

**LIME plot: feature values of mushrooms that decrease or increase the probability of predicting edible/non-edible or poisonous.**



```
Comparison with instance #0:
Feature                                      Original Value                 Modified Value
-------------------------------------------------------------------------------------------------
  cap_diameter_cm                            9.41
  cap_shape                                  other
  cap_surface                                sticky
  cap_color                                  yellow
  changes_color_or_releases_
   _milk_in_reaction_to_damage               does_not_change_color_or_no_milk
  gill_attachment_to_stem                    no_data
  gill_density                               dense
  gill_color                                  orange
  stem_height_cm                             1.43                        -> 23.5
  stem_width_mm                              16.03
  stem_base                                   no_data
  stem_surface                                no_data
  stem_color                                  brown
  hymenophore_type                           no_data
  hymenophore_color                          no_data
  spore_color                                no_data
  habitat                                     forests
  season                                      autumn

Comparison with instance #1:
Feature                                      Original Value                 Modified Value
-------------------------------------------------------------------------------------------------
  cap_diameter_cm                            9.41
  cap_shape                                  other
  cap_surface                                sticky
  cap_color                                  yellow
  changes_color_or_releases_
  _milk_in_reaction_to_damage               does_not_change_color_or_no_milk
  gill_attachment_to_stem                    no_data
  gill_density                                dense
  gill_color                                  orange
  stem_height_cm                             1.43                        -> 23.5
  stem_width_mm                              16.03
  stem_base                                   no_data
  stem_surface                                no_data
  stem_color                                  brown
  hymenophore_type                           no_data
  hymenophore_color                          no_data
  spore_color                                no_data
  habitat                                     forests
  season                                      autumn
```

**Counterfactual analysis:**
**how the AI's filling in of missing data**
**in the description of a specific mushroom affects the change in its edibility prediction.**

Clarification:

**The mushroom described in the visualization, studied in nature, was NON-EDIBLE/POISONOUS.**

The AI model correctly classified it as NON-EDIBLE/POISONOUS.

The two counterfactual analyses presented in the image show what data needs to be changed for this mushroom to receive a prediction of **EDIBLE**.

**Legend:**

Original value: the data taken for the specific mushroom from the original dataset

Changed value: the data modified or completed by the AI model

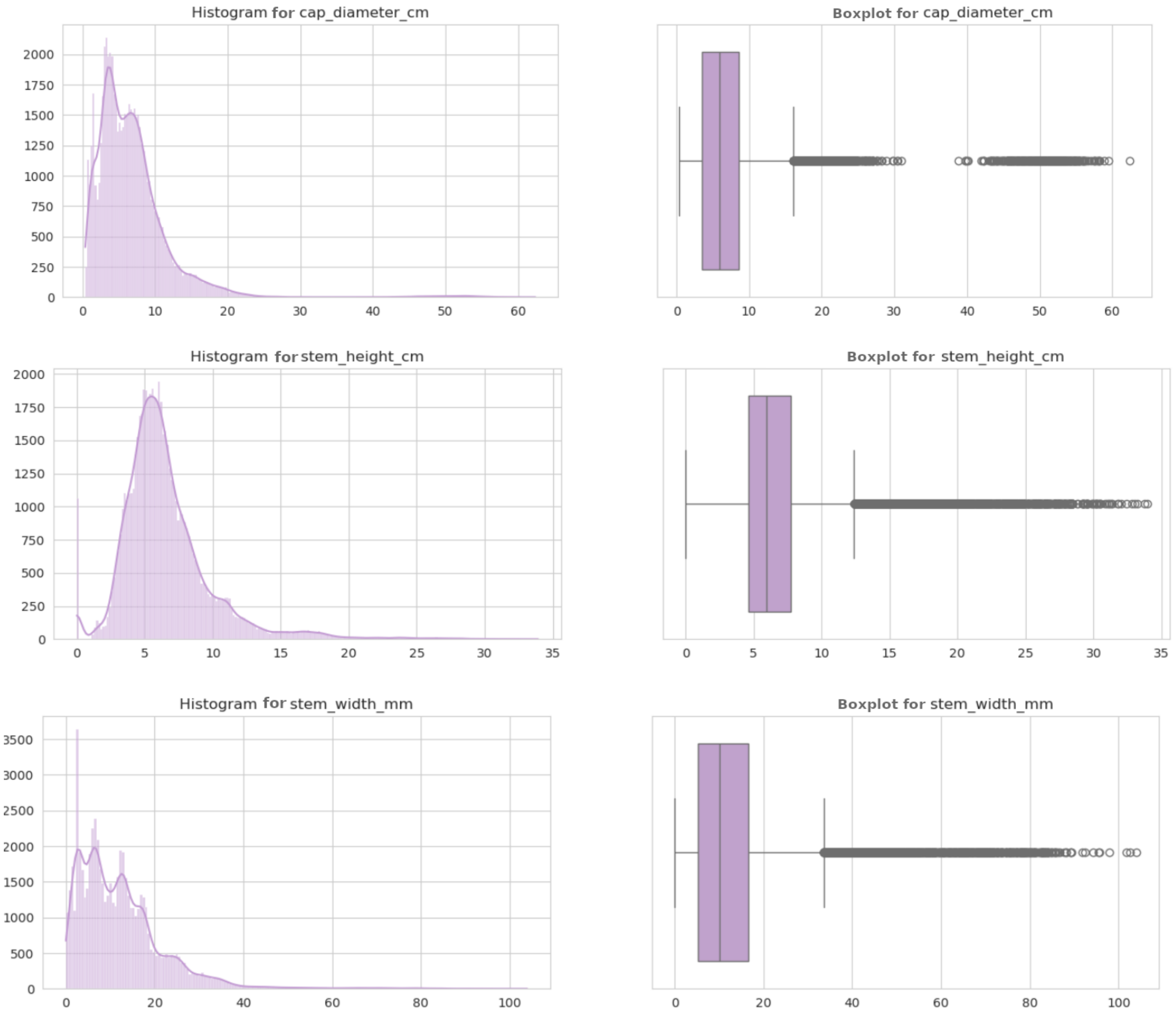
The study used an advanced machine learning model, known as the **Gradient Boosting Classifier** (shortened to "XGBClassifier"). This model achieved an accuracy of **99.97%**, which indicates its high effectiveness in distinguishing edible mushrooms from inedible or poisonous ones.

**High accuracy:** The model effectively identifies whether a given mushroom is edible or inedible/poisonous, based on the analysis of various mushroom characteristics.

**Data-driven decisions:** The model's decisions are based solely on the data it was provided. This means that the model uses the available information to make inferences but lacks knowledge *beyond the scope of the given data.*

**Model limitations:** While the model is highly accurate, it is important to acknowledge its limitations. It cannot account for all possible factors influencing a mushroom's edibility that may be known to experts. The model serves as an additional tool to assist in mushroom identification. However, due to the model's constraints, it is recommended not to treat its predictions as the final determinant but rather as one element in the mushroom identification process.

In our study, we used the standard method of dividing the data into two groups: one for training the model (training set) and the other for testing it (test set).


Histogram for cap_diameter_cm
Boxplot for cap_diameter_cm
Histogram for stem_height_cm
Boxplot for stem_height_cm
Histogram for stem_width_mm
Boxplot for stem_width_mm

| Example | A.I. prediction | Explanation of A.I. prediction |
|---|---|---|

## Example

cap_diameter_cm > 8.54
cap_shape = convex
cap_surface = smooth
cap_color = white
does_bruise_or_bleed = no_bruises_or_bleeding
gills_attachment_to_the_stem = sinuate
gills_spacing = close
gills_color = white
5.96 < stem_height_cm <= 7.74
stem_width_mm > 16.56
stem_root = no_data
stem_surface = no_data
stem_color = white
veil_type = no_data
veil_color = no_data
spore_print_color = no_data
habitat = meadows
season = spring

## A.I. prediction

Eadible

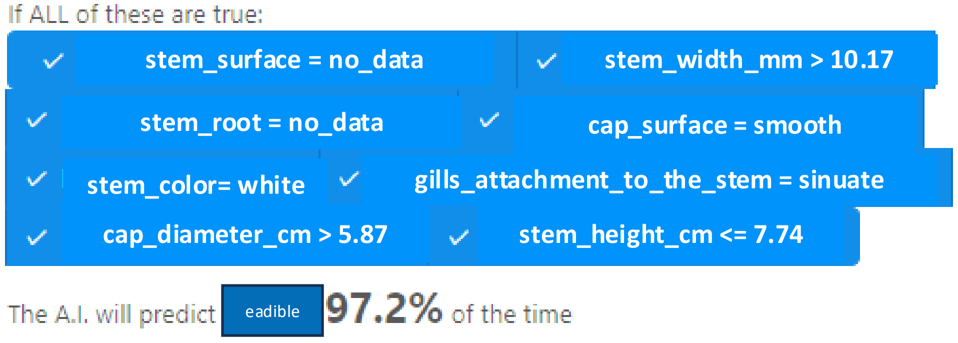
**ANCHOR - method for explaining AI model**

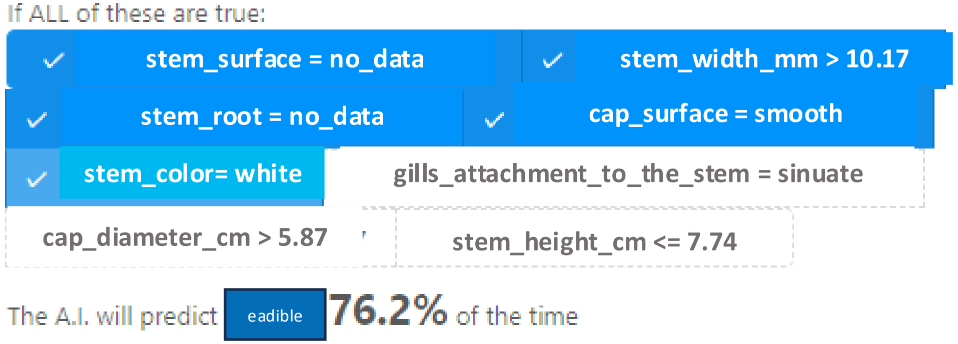
At the top right: anchor, which is a set of features whose combined presence (conjunction) determines how the AI model classifies a given mushroom.

The anchor does not have to reflect a real example from the data.

Below the anchor: the influence of the combined occurrence of the feature set (anchor) on the percentage of cases in which the model predicts a given class (i.e., the so-called classification certainty). When calculating this percentage, the model takes into account the features highlighted in blue.

## Explanation of A.I. prediction



Example

5.77 < capCdiameter_cm <= 8.54
cap_shape = flat
cap_surface = no_data
cap _color = green
does_bruise_or_bleed = no_bruises_or_bleeding
gill_attachment = adnexed
gill_spacing = close
gill_color = brown
5.96 < stem_height_cm <= 7.74
10.17 < stem_width_mm <= 16.56
stem_root = no_data
stem_surface = no_data
stem_color = white
veil_type = no_data
spore_print_color = no_data
habitat = forests
season = autumn

A.I. prediction

Non-eadible / poisonous

**ANCHOR - method for explaining AI model**

At the top right: anchor, which is a set of features whose combined presence (conjunction) determines how the AI model classifies a given mushroom.

The anchor does not have to reflect a real example from the data.

Below the anchor: the influence of the combined occurrence of the feature set (anchor) on the percentage of cases in which the model predicts a given class (i.e., the so-called classification certainty). When calculating this percentage, the model takes into account the features highlighted in blue.

Explanation of A.I. prediction

If ALL of these are true:

✓ gills_attachment_to_the_stem = adnexed
✓ cap_color = green
✓ cap_surface = no_data
✓ cap_shape = flat
✓ cap_diameter_cm > 5.87

The A.I. will predict poisonous 97.5% of the time

If ALL of these are true:

✓ gills_attachment_to_the_stem = adnexed
✓ cap_color = green
cap_surface = no_data
cap_shape = flat
cap_diameter_cm > 5.87

The A.I. will predict poisonous 84.0% of the time