Identifying Wild Mushrooms: What to forage, what to avoid?

A classification project by
Alison Salerno

Process

Step 1

- Gather a dataset of 8,124 mushroom samples from UCI
- Clean data for modeling

Step 2

Use EDA to gain insight into the categorical dataset:

- Visualize categorical features
- Explore the relationships between the feature's different values and class separation

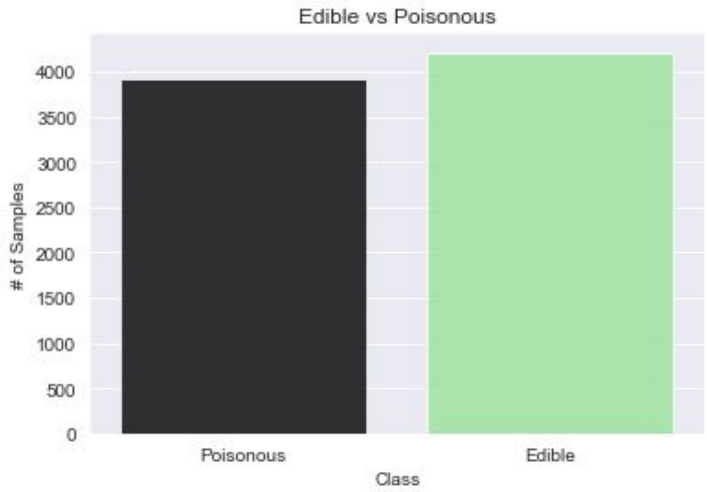
Step 3

Modeling:

- Compare/contrast model iterations
- Tune parameters
- Gather insight from model performance

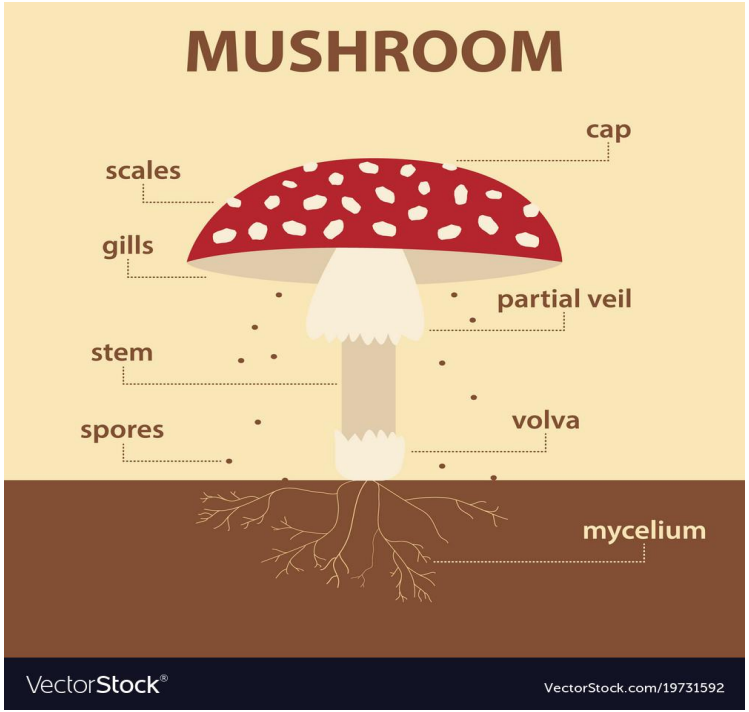
Target Variable

Edible or Poisonous?

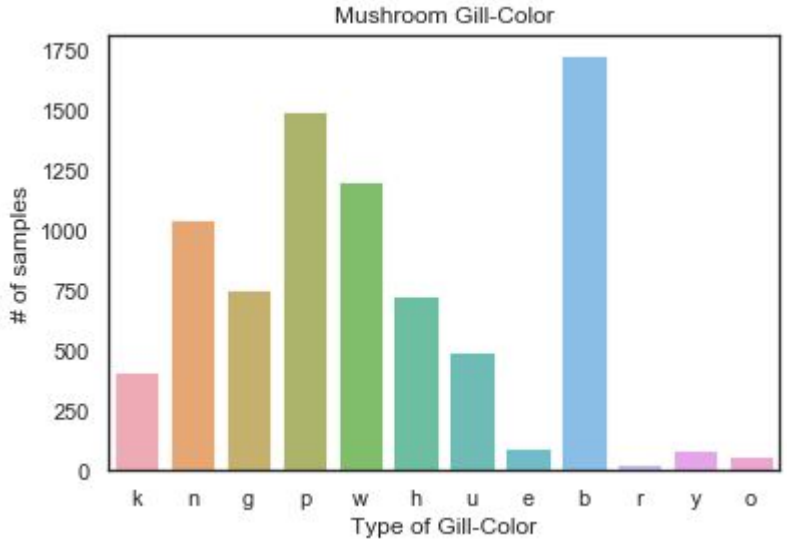


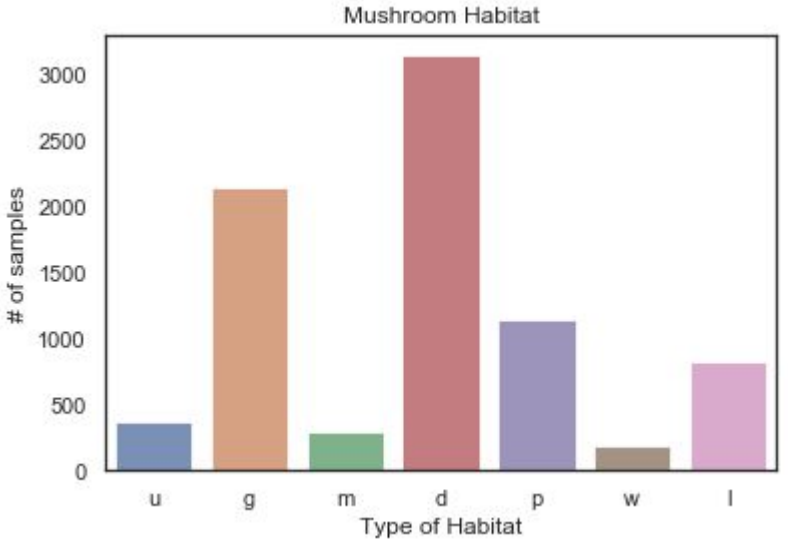
Independent Variables

- Cap Shape
- Cap Surface
- Cap Color
- Bruises
- Odor
- Gill Attachment
- Gill Spacing
- Gill Size
- Gill Color
- Stalk Shape
- Stalk Root
- Stalk Surface above ring
- Stalk surface below ring
- Veil Color
- Ring Number
- Ring Type
- Spore Print Color
- Population
- Habitat

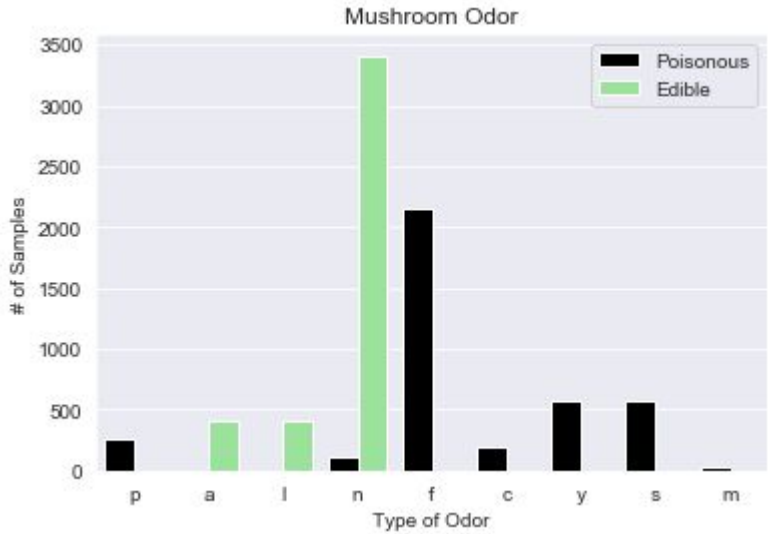


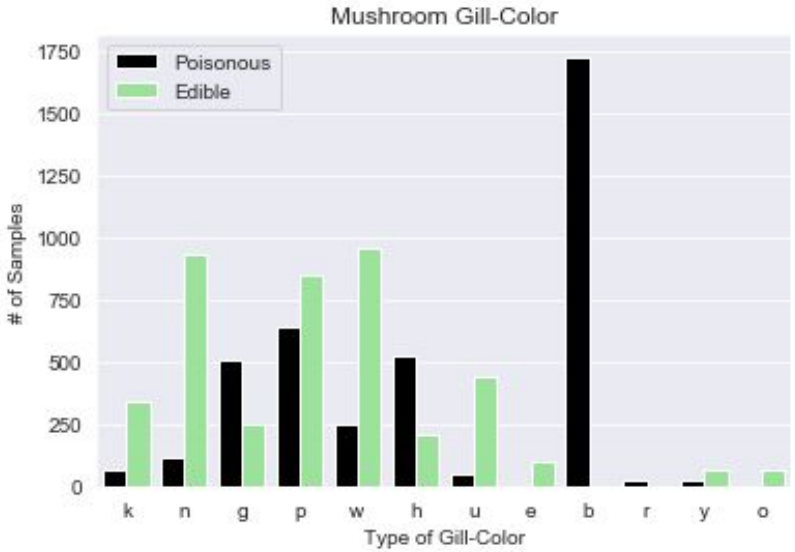
EDA: Features





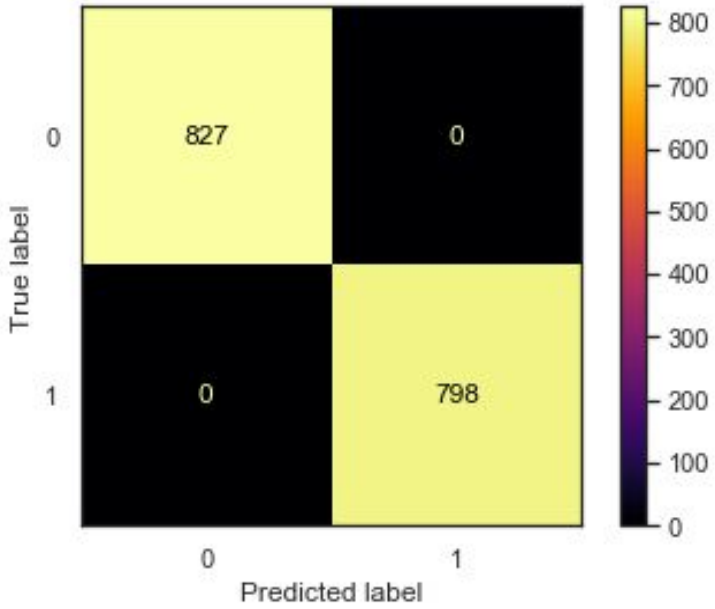
EDA: Features and Class separation



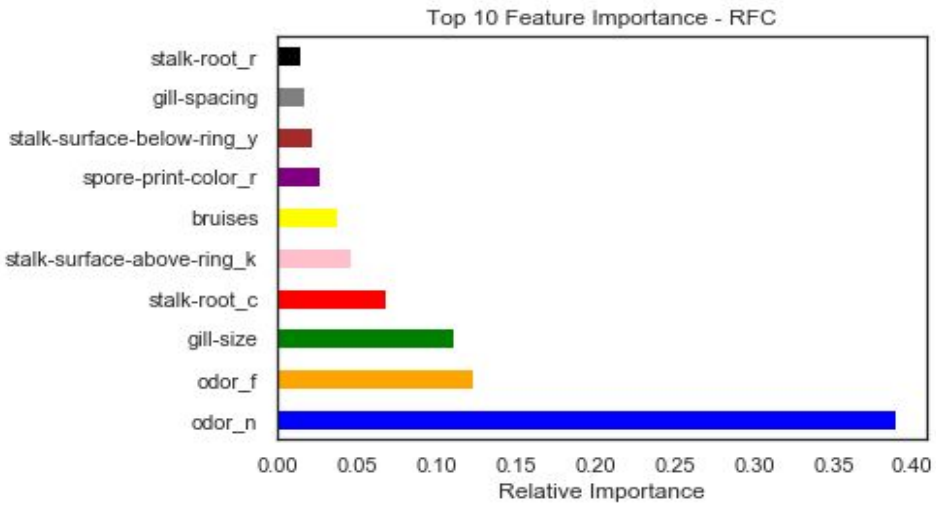


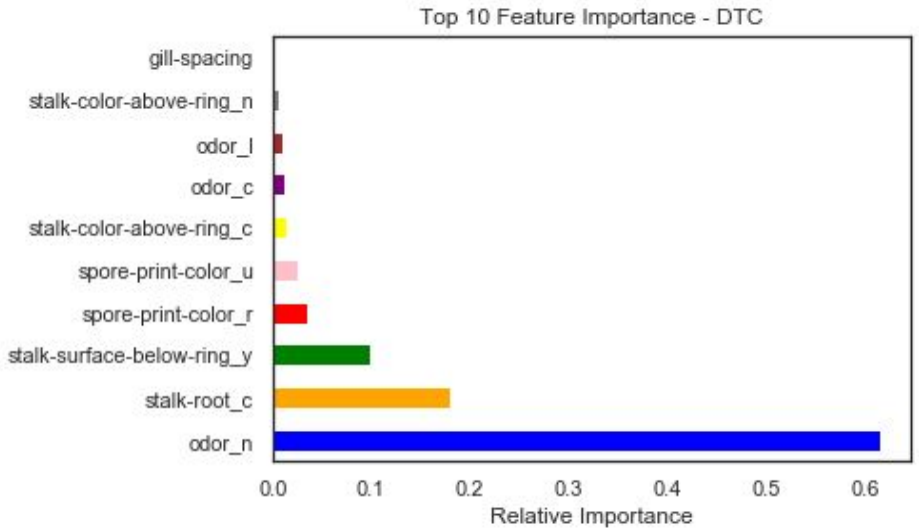
Modeling

	Accuracy	F1
Baseline - Dummy Classifier	.51	0.0
Logistic Regression	.99	.99
KNN	1.0	1.0
Decision Tree Classifier	1.0	1.0
Random Forest Classifier	1.0	1.0



Feature Importance

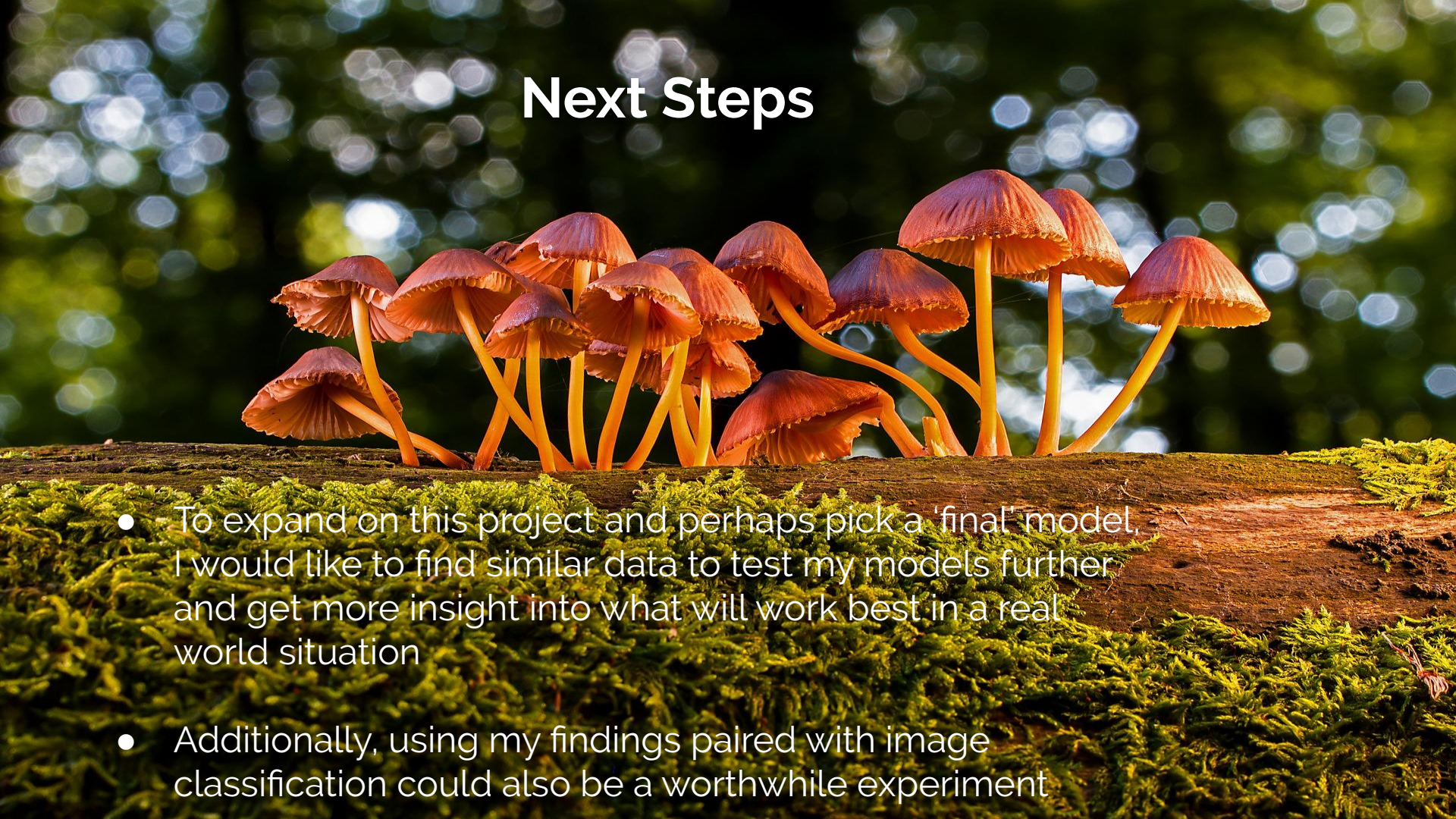




And finally, a few simple rules to identify a poisonous mushroom....

1. ***Odor**: If the Mushroom has an odor, especially if the odor is not pleasant, it is likely to be poisonous.
2. **Gill Size:** If the gills are narrow, it is likely to be poisonous.
3. **Spore Print Color:** White, red, 'chocolate' are likely to be poisonous.
4. **Bruises:** If the mushroom does not have bruises, it is likely to be poisonous.
5. **Stalk Surface Above Ring**: If it appears silky, it is likely to be poisonous.

Next Steps



- To expand on this project and perhaps pick a 'final' model, I would like to find similar data to test my models further and get more insight into what will work best in a real world situation

- Additionally, using my findings paired with image classification could also be a worthwhile experiment