POISONOUS MUSHROOMS

Classification Final Project

Miki Dash

When wild mushrooms are dangerous.

- Survival Analytics wants to provide a reference guide on recognizing poisonous mushrooms in the wild through the use of data science
- Only about 3% of known mushroom varieties are poisonous, and the symptoms of poisoning can vary from gastrointestinal discomfort to liver failure and death, depending on the type of toxin ingested

Background

- Survival Analytics wants to provide a reference guide on recognizing poisonous mushrooms in the wild through the use of data science

- Reference guide must provide indicators that are applicable via:
  - Sight
  - Smell
  - Touch(feel)

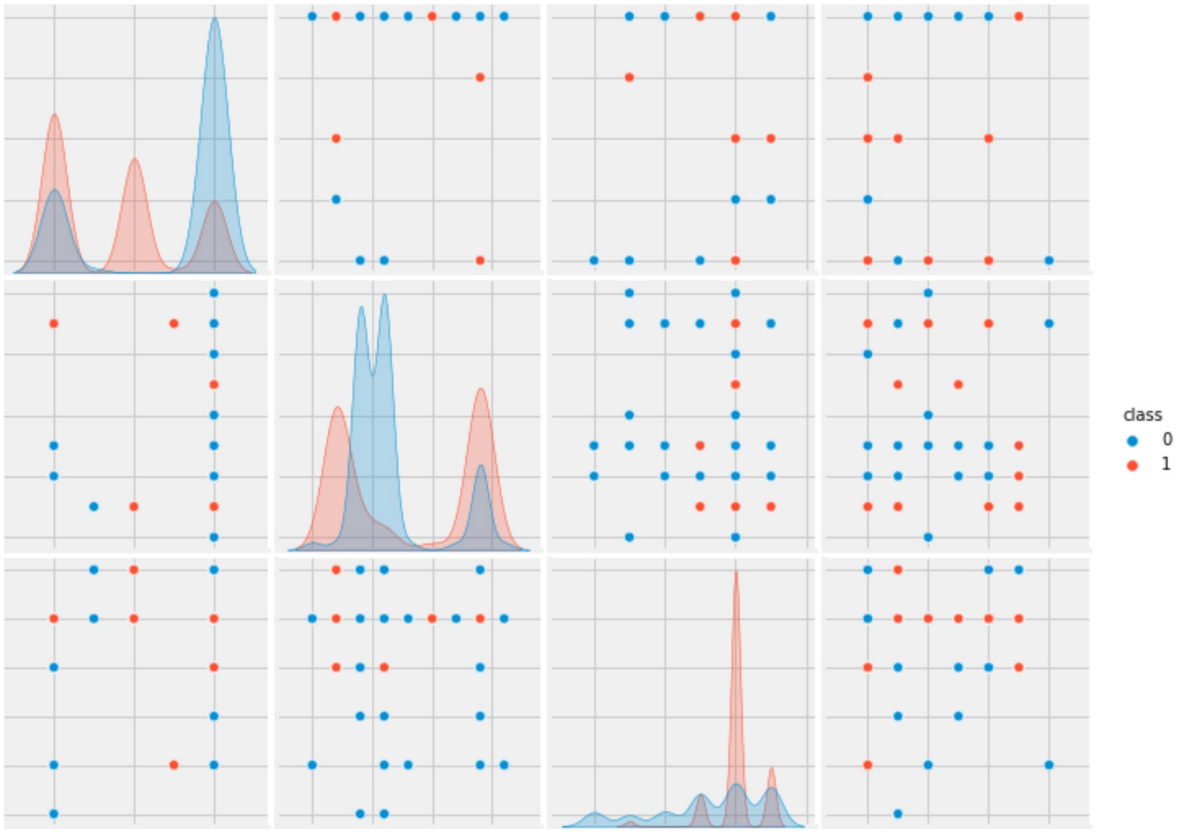
Data Source and Description

- Data Source: UCI Machine Learning Repository
- This data set consists of 23 species of gilled mushrooms in the Agaricus and Lepiota Family
- Each species is identified as definitely edible and poisonous
  - 8124 observations
  - 23 categorical features  Expanded to 93 dummy variables

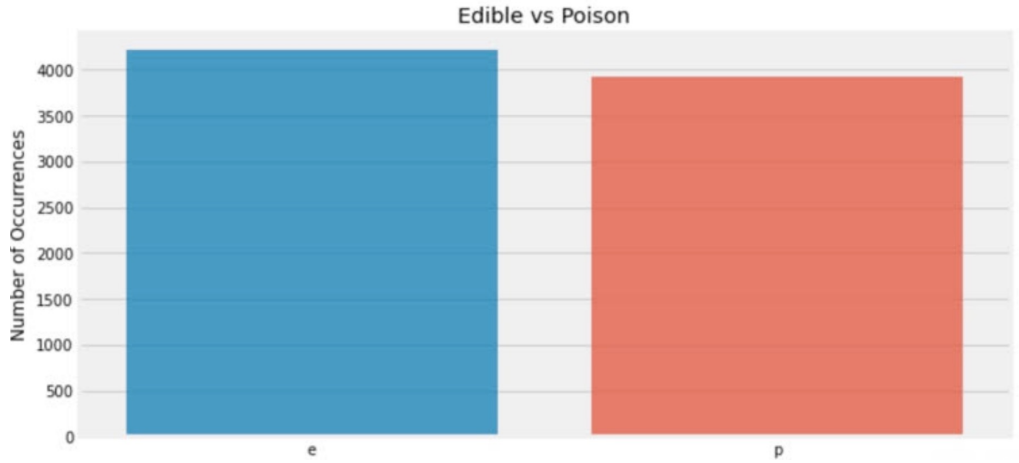
Tools

Exploratory Data Analysis

Snapshot: Difficult to separate classes by feature



Snapshot: Balanced Data



Process

- Thirty five features were initially selected for analysis
  - Using an iterative process the final model consisted of 10 features
- Tested several models
  - Logistic Regression
  - Random Forest
- Metric of evaluation
  - Accuracy
  - Recall

Analysis

- Accuracy
  - Balanced target classes
    - Poison = 1
    - Not Poison (edible) = 0
- Recall
  - Essential to minimize false negatives
    - Avoid classifying a poisonous mushroom as edible

Random Forest Performance

	Precision	Recall	f1-score
Not Poison	0.97	1.00	0.98
Poison	1.00	**0.97**	0.98

Accuracy: **98.0**

Logistic Regression Performance

	Precision	Recall	f1-score
Not Poison	0.97	1.00	0.98
Poison	1.00	**0.96**	0.98

Accuracy: **98.0**

Random Forest: Confusion Matrix

	Predicted Not Poison	Predicted Poison
Actual Not Poison	1365	0
Actual Poison	**44**	1272

Logistic Regression: Confusion Matrix

	Predicted Not Poison	Predicted Poison
Actual Not Poison	1365	0
Actual Poison	47	1269

Key Predictors of Toxicity

- Odor
- Stalk Surface
- Bruises
- Spore Print Color

“You can eat any mushroom … once”

Unknown Mushroom Hunter

Foraging

Schrooming

Be Careful!