Mushroom Analysis

Suganya Balasubramanian, Sam Farber, Katie Darden, Emily Costello & Bianca Calin

Selected Topic

Classifying safe versus poisonous mushrooms

Why did we select this topic?

All five of us are massive mushroom enthusiasts and need to ensure that we can survive the upcoming climate apocalypse in the wild.

Description of Data Source

- CSV file with 8124 instances of mushrooms and their attributes. This data set includes descriptions of hypothetical samples corresponding to 23 species of gilled mushrooms in the Agaricus and Lepiota Family.

- Each species is identified as definitely edible, definitely poisonous, or of unknown edibility and not recommended.

- The data comes from The Audobon Society Field Guide to North American Mushrooms (1981) and was contributed to the UCI Machine Learning Repository in 1987.

Questions we were aiming to answer

1. Which features are most indicative of a poisonous mushroom?
2. Which features of mushrooms are most palatable?
3. Which habitat contains the highest percentage of edible mushrooms?
4. What populations contain the most edible and most poisonous mushrooms?

Data Exploration Phase

- Reviewing different features such as:
  - Cap shape
  - Bruises
  - Odor
  - Gill size
  - Ring number & type

- Determining the appropriate code to use to best suit our dataset

Analysis Phase

- Creating features/target
- Split, train, & test the data
- Make predictions based on the dataset
- Determine accuracy of the data
- The Random Forest Classifier is an ensemble learning model consisting of several decision tree algorithms trained on random subsets of our data.
    - By combining multiple relatively uncorrelated decision trees, the Random Forest Classifier protects against overfitting and improves overall accuracy.
    - The model also allows us to rank the importance of our input variables, and is able to accommodate larger datasets.

Dashboard

- Our dashboard will display the results of our analysis of mushroom features
- Using Tableau, we will determine the features that best indicate an edible v. poisonous mushroom
- Interactive elements will include:
  - Chart to display results*
  - Graph that displays edible v. poisonous elements*