Identifying Features of Poisonous Mushrooms

Background

A new app that helps users scavenge for mushrooms

Poisonous mushrooms: digestive problems, organ failure, death

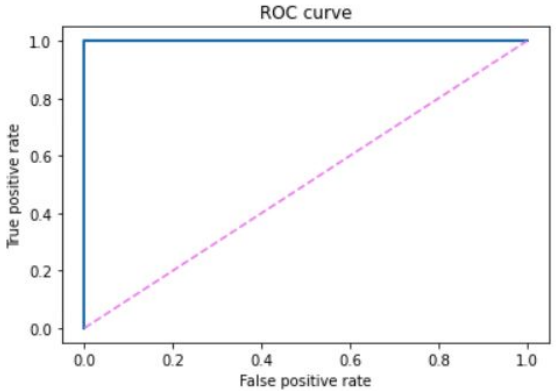
Data & Methodology

Data: UCI Mushroom Dataset
(https://archive.ics.uci.edu/ml/datasets/mushroom)

- Classification: Poisonous vs Edible
- 96 features: odor, roots, gills, etc
- KNN, LR, RF, XGB

Original dataset achieves perfect scores for KNN, LR, and Random Forest

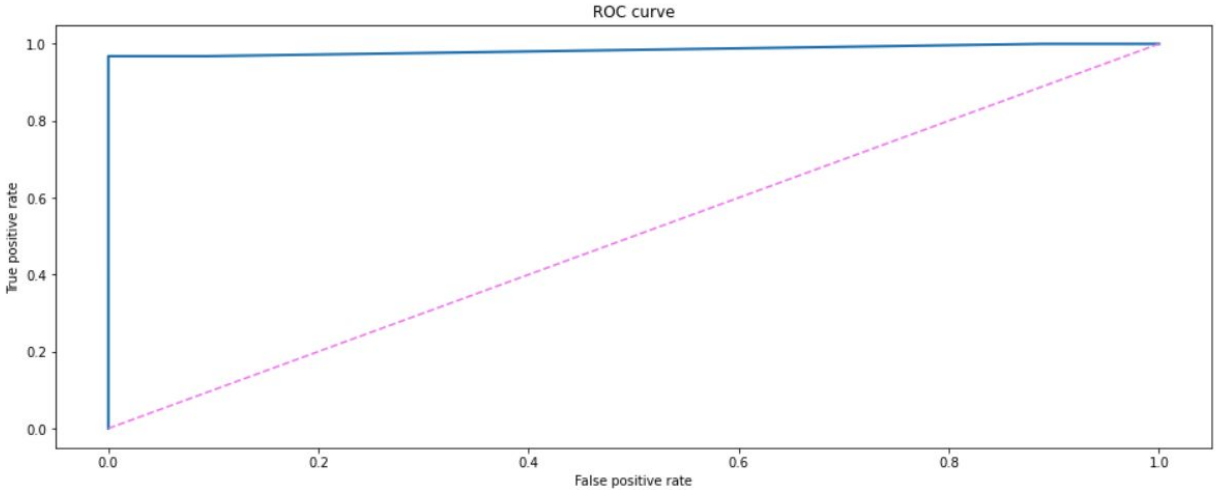
- In the real world, we won't have access to this many features



Identifying Features: Odor

ROC AUC: **98%**

KNN and LR accuracy: **98%**



Stalk-root_c: club-shaped stalk root

Stalk-root_r: rooted stalk root

Odor_n: odorless

Odor_l: anise odor (licorice-like)

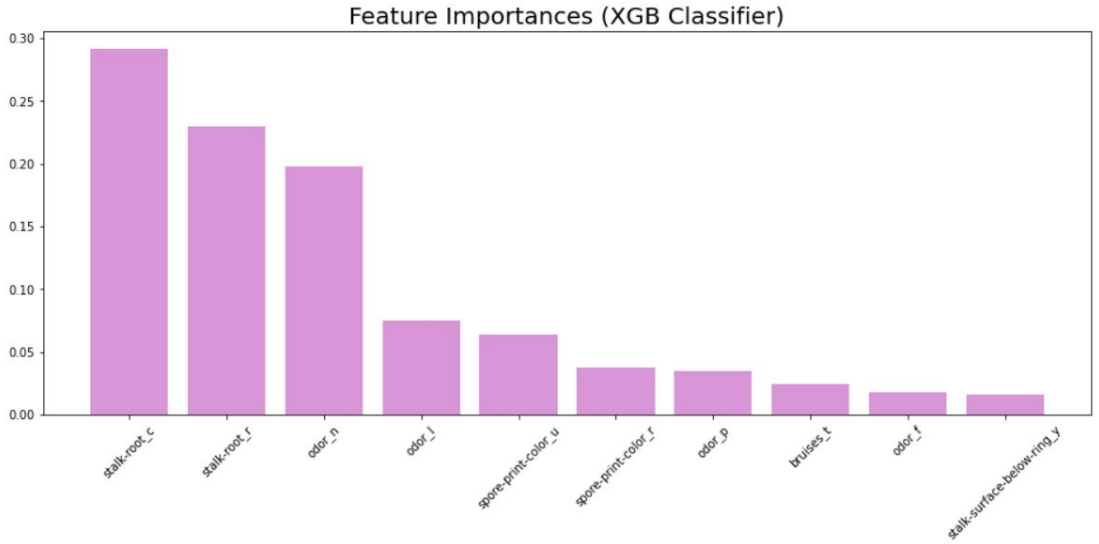
Spore-print-color_u: purple spore print

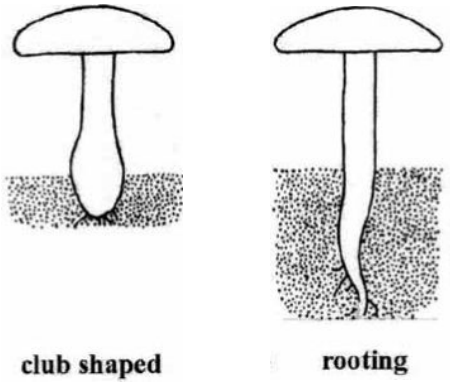
Spore-print-color_r: green spore print

Bruises_t: bruised (the alternative is 'not bruised')

Odor_f: foul odor

Stalk-surface-below-ring_y: scaly surface below ring





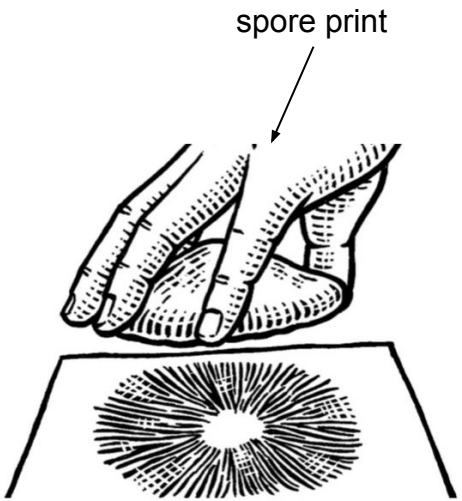
Odor + spore print color + gill features

KNN & LR accuracy: **99.88%**

Recall: **1.0**

Precision: **1.0**

F1: **1.0**



Highly Predictive Features of Poisonous Mushrooms:

- Foul odor
- Narrow gill size
- Chocolate-colored spore print
- White-colored spore print

Conclusion

- You can produce highly accurate models using 1-3 features!

Future work:

- Image Classification