

Deadly V. Edible Gilled Mushrooms

Using supervised machine learning
classification techniques to identify if
a gilled mushroom is edible or
poisonous based on its characteristics

Lee Whieldon



Objective

Goals of This Study

01

Exploration of Gilled Mushroom Features

Examine features of toxic versus edible North American gilled mushrooms

02

Optimal Classification Model

Compare 3 commonly used classification for accuracy: Logistic Regression, K Nearest Neighbor (kNN), and Decision Tree Classification

03

Provide Solid Feedback to Stakeholders

Effectively articulate outcomes of study to key stakeholders in the mushroom foraging communities (in this fictitious case, Oregon)

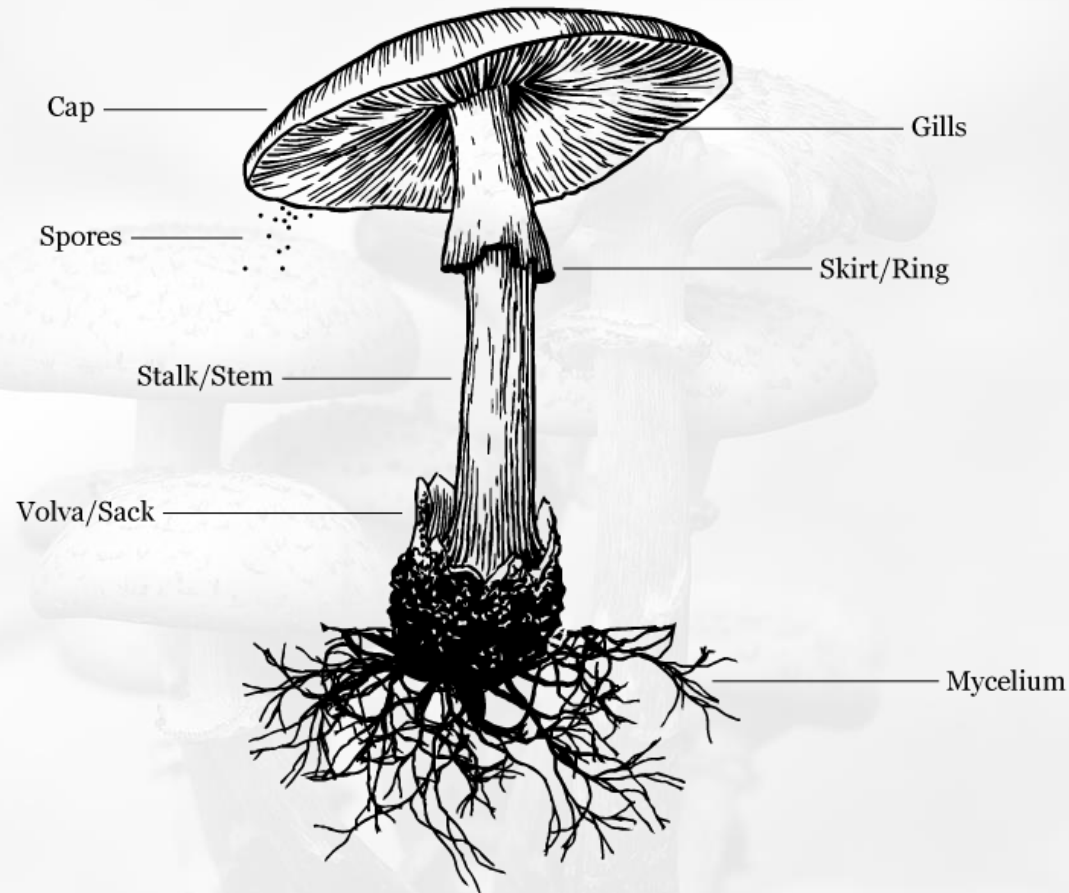


Gilled Mushroom Key Terminology



Below are the high-level aspects of a gilled mushroom's characteristics:

1. **Cap:** Shape, texture, and color
2. **Gills:** Attached/notched/free, proximity, size, color
3. **Stalk:** Shape and root; appearance above & below ring
4. **Veil and Ring:** Intact? Color? Number of rings, shape
5. **Spores:** Color
6. **Population:** How do they grow? Clustered? Solitary?
7. **Habitat:** Grasses, leaves, meadows, waste, woods?



Background



Mushrooms (a.k.a. Fungi) have a multitude of uses for humankind.
Used in medicine, food, and even packaging & biofuel.

Local mushroom foragers are seeing a
big rise in the interest of fungi in big pharma & business.

Foraging of wild mushrooms can result in serious illness and death,
resulting in a 20% mortality rate in adults and 50% in children who
consume them

Local mushroom foragers (in this fictitious case, Oregon) need the
help of data science to identify if a gilled
mushroom is poisonous or edible



Results

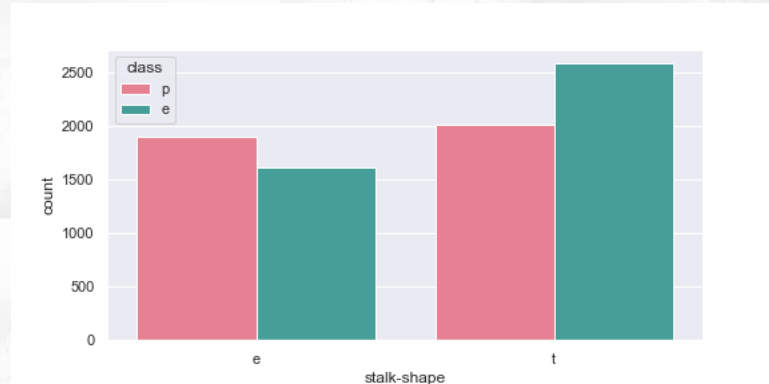


- *Edible & poisonous gilled mushrooms share similar features, making it hard to differentiate*
- *This is a common problem with mushroom foraging; without a skilled eye and significant experience, you could get sick!*
- *Below are some example of characteristics for each category of the mushroom's anatomy (you can see poison or edible categories of labels below, they pretty much share similar characteristics across the board)*

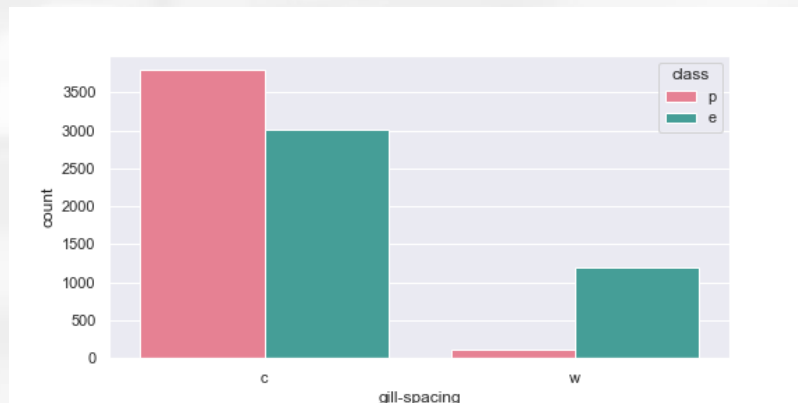
Cap



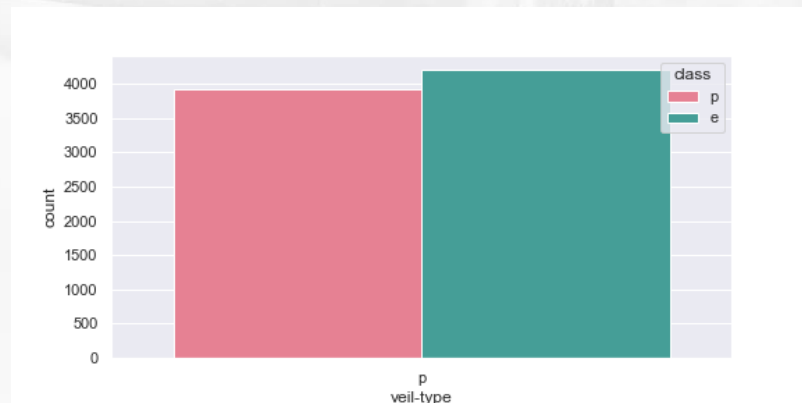
Stalk



Gills



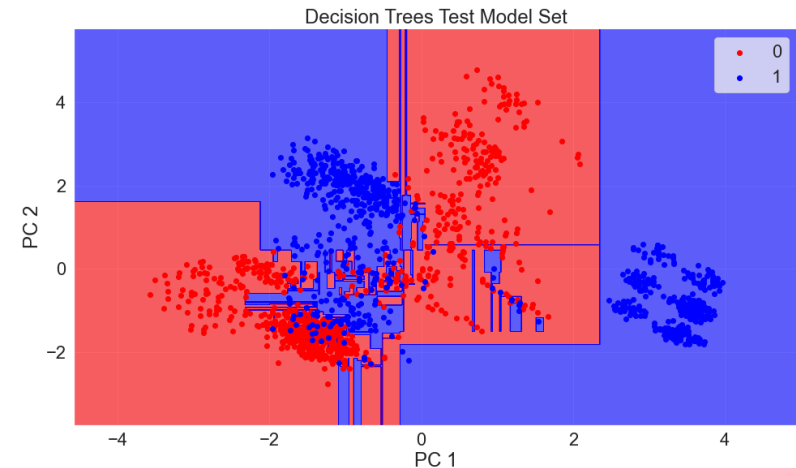
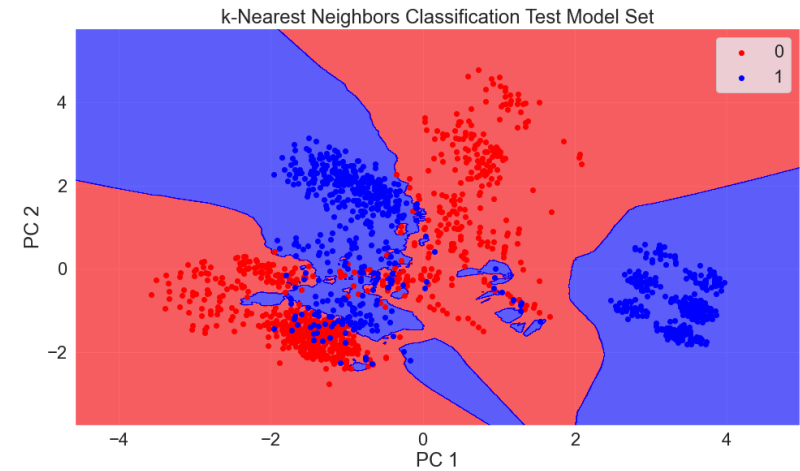
Veil & Ring



Results (Cont'd.)



- Applied PCA to convert features on a 2D plane
- 20% of Data Was Allocated to Test Set (~1600)
- For Both Train & Test Sets:
 - K-Nearest Neighbors Performed Best
 - Logistic Regression Performed Worst
- Mesh grid visual shows model fit
- Scatterplots represent the labels (edible=0; poisonous=1) against the model



Conclusion & What's Next?



Conclusion

- k-Nearest Neighbors has best precision, recall, and f1-score out of the 3 models.
- Logistic regression underfits the data and is not recommended

Results of Model's Weighted Average			
Model Name	Precision	Recall	F1-Score
Logistic Regression	0.64	0.64	0.64
k-NN	0.94	0.94	0.94
Decision Tree	0.93	0.93	0.93

What's Next?

- Expand study to capture many types of mushrooms (gilled mushroom is just one type... There's a TON more!)
- Expand study to a global scale of mushroom collection:
 - Countries like Korea, China and Turkey could benefit from this analysis to reduce fatalities
 - Soil, climate, and topography affect how a mushroom grows and develops.



Mushroom Name: [*Podostroma Cornu-Damae*](#)

One of the Deadliest
Mushrooms in the world
(and not gilled!)

THANK
YOU



Appendix

01

Data Source

02

References



Data Source



[UCI Machine Learning Mushroom Dataset](#)

- *This dataset contains 8124 entries corresponding to 23 species of gilled mushrooms from North America.*
- *Each species is identified as:*
 - *Definitely edible (e),*
 - *Definitely poisonous (p), or*
 - *of unknown edibility and not recommended (also p).*
- *Each entry is a single mushroom and has 22 features related to its physical characteristics. (Data source: The Audubon Society Field Guide to North American mushrooms)*



References



- *Yelken, B., & Mantar, K. K. (2020). Mushroom That Break Hearts: A Case Report. Turk J Intensive Care, 18, p. 43-6.*
- *M. Davis, R. Sommer, and J. Menge, Field guide to mushrooms of western North America, vol. 106. Univ of California Press, 2012. p. 40*
- *Erkmen, B., & Yildirim, T. (2008). Improving classification performance of sonar targets by applying general regression neural network with PCA. Expert Systems with Applications, 35(1-2), p. 472-475.*
- *James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning (Vol. 112, p. 312). New York: springer.*