Mushroom Edibility Analysis

Tofu-FC - Huiwen Wang, Rocky Zhang, Darrick Zhang

2024-12-12

Introduction

Project Motivation / Background:

Mushrooms play a crucial role in maintaining the health and balance of ecosystems worldwide. As nature's decomposers, they break down organic matter and facilitate nutrient cycling in the soil, making essential elements available for other organisms. These fascinating fungi also serve as a nutritious food source for humans and various wildlife, offering valuable proteins, vitamins, and minerals. However, it's important to note that while many mushroom species are beneficial, some can be deadly poisonous, making proper identification and caution essential when encountering wild mushrooms.

The importance of this research has been highlighted in a multitude of studies. There have been over 130,000 reported cases of mushroom exposure from 1999 to 2016, with 704 of these incidents causing "major harm" and 52 being fatal (Brandenburg and Ward 2018). Many of these cases are caused by the misidentification of edible mushroom species, and thus may be preventable through more accurate mushroom classification.

Thus, our research will focus on what physical features and environmental factors of mushrooms humans foragers can use to identify toxic/poisonous mushrooms in the wild. For example, mushrooms in the wild with obvious physical features like white gills, white rings, red caps, or red stems tend to be poisonous. These obvious physical traits are more likely to be spotted by animals, which would provide an evolutionary disadvantage unless they contain certain self-defense mechanisms, such as poison or toxins.

Additionally, the habitat and season in which mushrooms are planted and grow may also affect whether they're poisonous. Different temperatures, humidity, and light can affect the production of toxins, which may also affect the edibility of mushrooms. By conducting a research study on how to distinguish between safe and dangerous species, we can mitigate the incidence of mushroom poisoning and ensure safer foraging practices.

Research Question:

What environmental factors and/or physical features of mushrooms indicate that a wild mushroom is poisonous or edible?

Data Description:

The dataset we used is an expansion of the UCI mushroom dataset, curated and submitted on April 26, 1987 by the National Audubon Society Field Guide. The publisher of the Field Guide, National Audubon Society, conducted extensive field research throughout North America, recording their observations on various aspects of mushrooms. Their research incorporate a wide range of physical characteristics, as well as environmental factors such as the type of habitat and seasonal variations. Importantly, the study also focused on the toxicity of the mushrooms, noting which species were poisonous. In 2020, an effort was made to expand the data, capturing data from 173 species of mushrooms (as opposed to the original 23 species). This comprehensive dataset provides valuable insights into the relationship between mushrooms and their environments, contributing significantly to the understanding of the factors influencing mushroom toxicity.

Our response variable is `class`, which is a qualitative variable labeled "e" for edible or "p" for poisonous.

Because we want our classifier to be easily used by people, and quantitative predictors can be harder to measure, we will focus on only one. That is `cap.diameter`, the diameter of the mushroom cap (cm).

Key qualitative predictor variables include `cap.shape` (the shape of the mushroom cap), `gill.color` (the color of the fungi gills), `stem.color`(the color of the mushroom stem), `habitat` (the habitat that the mushroom is grown/found), and `season` (the season that the mushroom is grown/found). The key for the levels of each categorical variable are described on the following page and in the data dictionary.

The data dictionary can be found here (Dennis Wagner and Hattab 2021). There are 61069 total primary observations, each row represent a mushroom's physical features and also the environment it was found in. For the purpose of this project, the data was split into training (70%) and testing data (30%). Distribution of testing and training distribution was determined by the lack of data of specific combinations of mushroom's features seen later on in this analysis.

Exploratory Data Analysis

We will first look at our response variable:

Table 1: Distribution of Classes

class	n	percentage
e	27181	0.445
p	33888	0.555

Looking at the overall distribution of our response variable `class`, most of the mushrooms in our dataset seem to be poisonous ("p"). 33888 (55.5%) of the observations are labeled poisonous, as opposed to 27181 (44.5%) of them as edible.

Distribution of Cap Diameter
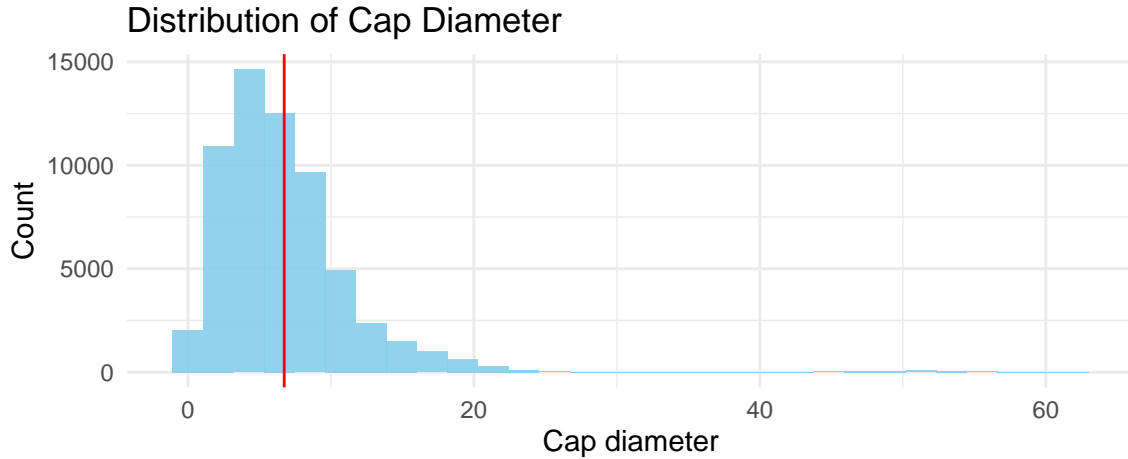
From the visualization, the distribution of cap diameter appears roughly unimodal and skewed right. The mean cap diameter is 6.734 cm, with a standard deviation of 5.265 cm.

For the rest of qualitative predictors, we report some in the tables below:

For `cap_shape`:

Table 2: Distribution of Mushroom Cap Shapes

	bell	conical	flat	other	spherical	sunken	convex
cap.shape	b	c	f	o	p	s	x
percentage	0.093	0.030	0.219	0.057	0.043	0.117	0.441

For `habitat`:

Table 3: Distribution of Mushroom Habitats

	woods	grasses	heaths	leaves	meadows	paths	urban	waste
habitat	d	g	h	l	m	p	u	w
percentage	0.724	0.130	0.033	0.052	0.048	0.006	0.002	0.006

For `season`:

Table 4: Distribution of Mushroom Seasons

	autumn	spring	summer	winter
season	a	s	u	w
percentage	0.494	0.045	0.375	0.086

For `stem_color`:

Table 5: Distribution of Mushroom Stem Colors

|	buff	red	none	gray	black	blue	brown	orange	pink	green	purple	white	yellow
stem.color	b	e	f	g	k	l	n	o	p	r	u	w	y
percentage	0.003	0.034	0.017	0.043	0.014	0.004	0.296	0.036	0.017	0.009	0.024	0.375	0.129

For qualitative variables, there appears to be more common physical and environmental characteristics. For example, for cap shape, flat and convex tends to be the most common; for stem color the most common is white, yellow, and brown; for habitat, woods is the most common. Thus, there are also characteristics which happen to be rarer, yet for some reasons, natural selection has decided to preserve. These characteristics may have evolutionary advantages (such as being poisonous), and we hope that they help us in our logistic regression model.
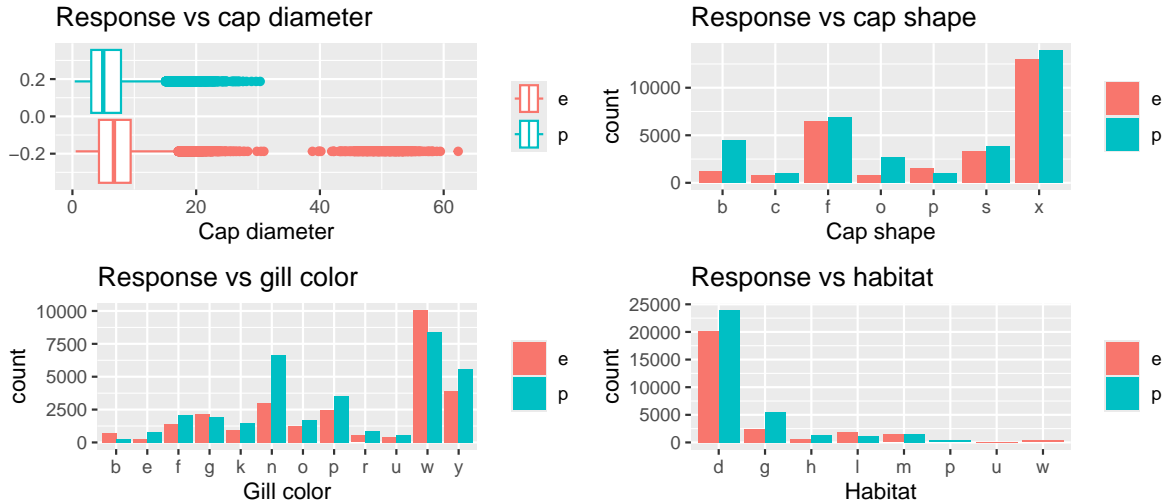

Looking at this bivariate exploratory data analysis, smaller cap diameters seem to be correlated with poisonous mushrooms on average. We also observe some categories with a large disparity between the number of edible and number of poisonous mushrooms, offering potential modeling power. For example, if we randomly select a mushroom with a cap shape of convex, bell, or others, it is more likely to be poisonous/toxic than edible. Similarly, mushrooms with gill color of brown and yellow tends to be poisonous. In many cases, however, it is hard to accurately

predict whether a mushroom is edible or poisonous soley based on a single characteristic. This suggests that our model needs to incorporate multiple predictors and/or interaction terms.

One interaction term we discovered is `habitat*season`. Mushrooms growing in the same habitat may have different toxicity depending on whether poisonous is needed to defend against predators. As the number of predators varies by season, season can alter how habitat affects the log-odds of mushroom toxicity.

Edibility Distribution by Habitat and Season
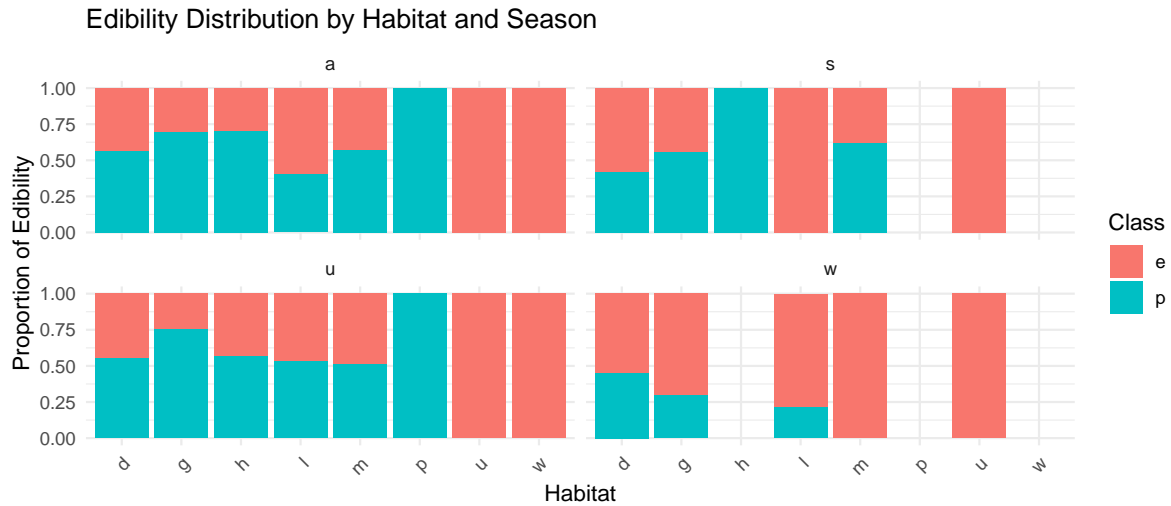
From the graph, mushrooms in certain habitats might only be edible during specific seasons. For example, mushrooms in the meadows are edible exclusively in the winter, but may be poisonous in other. Thus, we should incorporate this interaction term into our final model.

Cap Color and Shape Combinations by Edibility
Using log ratios of edible to poisonous mushrooms
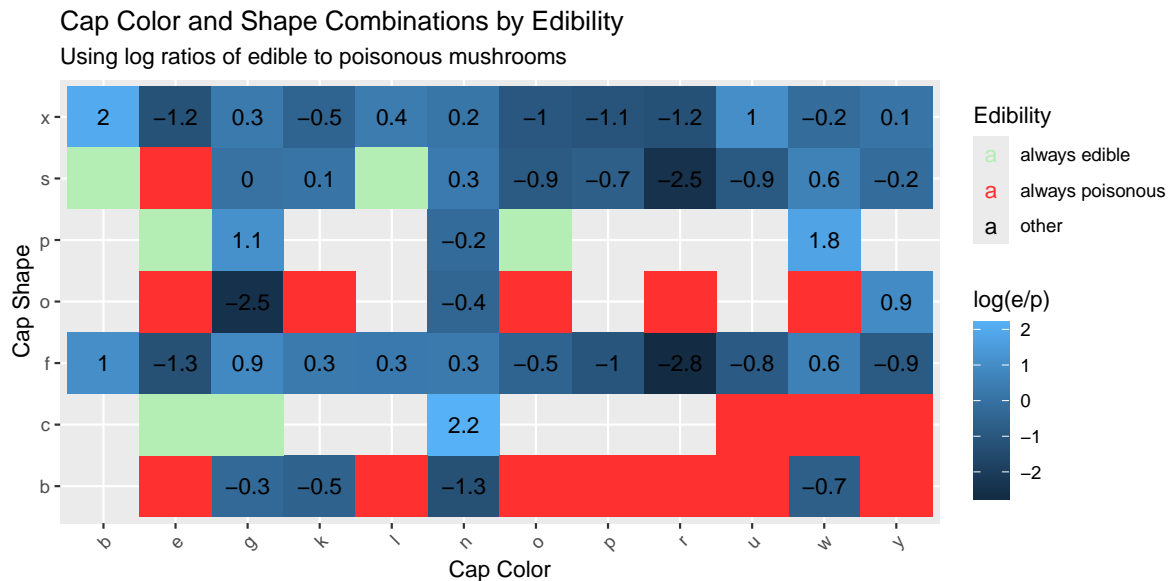
As the heatmap shows, certain combinations of cap color and cap shape are always edible or poisonous. Additionally, the log ratios across combinations of cap colors and cap sizes are varied with no pattern – for a mushroom with a sunken cap shape, it could be always edible (if the color is buff) to always poisonous (if the color is red). Similarly, if a mushroom is brown, it could be high likely it is edible (if the cap shape is conical) or likely it is poisonous (if the cap shape is bell). Thus, we may have to consider this interaction effect in our final model.

Methodology/Analysis

Model Assumption

Since we try to predict a binary outcome (edible or poisonous), we use a logistic regression model. Below are the assumptions necessary for a logistic regression fit.

- Linearity: The log-odds appears to have a linear relationship with the quantitative predictor (see appendix for empirical logit plot).

- Randomness: The data were curated randomly in different places in North America, and thus we assume this condition is met.

- Independence: The observations are collected over a period of time, but in different regions and locations. For our analysis we assumed independence was met as the period did not span several years.

Base Model

We fit a logistic regression model predicting the binary response `class_binary` using predictors `cap.diameter`, `season`, `cap.shape`, `cap.color`, `gill.color`, `stem.color`, `habitat`. We chose predictors based on both the EDA and how easy they're to understand and evaluate. To determine if any predictors may not useful, we looked at model's coefficients with p-values greater than 0.01 (see Table 8 in the appendix).

The Wald's Significance Tests for coefficients of multiple categories of the same predictor variables reveals that certain categories only have limited data (also showen in EDA) and/or limited predictive power. For the simplicity of model, we combine these categories into a general "Other" category. For example, `stem.color` of "w", "y", and "n" were kept while the other observations were assigned to a general "Other" category. For `habitat`, "d" and "g" were kept.

Running a likelihood ratio test to evaluate the overall significance of the coefficients of the new model with modified categorical variables replaced `stem.color` and `habitat` with `stem.color.modified` and `habitat.modified`, we have:

term	residual.deviance	df	deviance	p.value
class_binary ~ 1	83921.51	NA	NA	NA
our model	70569.07	37	13352.44	0

$$H_0 : \beta_1, ..., \beta_p = 0 \qquad\qquad H_a : \beta_j \neq 0 \ for \ at \ least \ 1 \ j$$

Since the p-value is small, and less than $\alpha = 0.05$, we reject H_0. The data provide sufficient evidence of at least one non-zero coefficient in the model. The model coefficients and corresponding Wald test inferential statistics for our main model provide better p-value than the unmodified base model (see Table 10 in the appendix).

Interactive Terms

Our EDA showed evidence of potential interaction terms. To determine the need for them in our model, we performed a drop in deviance test with the added interaction terms of `habitat*season` and `cap.shape*cap.color`.

term	residual.deviance	df	deviance	p.value
reduced model	70569.07	NA	NA	NA
full model (w/ interaction)	65170.76	54	5398.315	0

Since the p-value is low below $\alpha = 0.05$, we decide to include these interaction terms as there is convincing evidence that at least one of these interactive term coefficients are not 0 and thus helpful in the model. Additionally for the base model, the BIC is 7.1×10^4, whereas for the model with interaction effects, the BIC is 6.62×10^4, also suggesting we should include the interaction terms (lower BIC). However, a lot of the coefficients for the interactive terms are N/A's and we run into the same issue as the main effects model (no/lack of data values for particular interaction combinations). We transformed the rest of the categorical variables and the coefficients looked way better (observe Table 11 vs. Table 12 in the appendix).

Model Results

Final Model

Our final model is a logistic regression model predicting the binary response variable `class_binary` as a function of the following predictors: `cap.diameter`, `season.modified`, `cap.shape.modified`, `cap.color.modified`, `gill.color.modified`, `stem.color.modified`, `habitat.modified`, `habitat.modified*season.modified`, and `cap.shape.modified *cap.color.modified`. Full model coefficients can be seen in the appendix (Table 12).

Below highlight some interesting coefficients for foragers to use to help them identify mushrooms that are more likely to be edible. Note the intercept is not useful in context since a mushroom cannot have a cap diameter of 0 cm.

- The coefficient for cap diameter is -0.064, which means that for each 1 cm increase in the cap diameter, we expect the odds of the mushroom being poisonous multiplies by a factor of 0.938, holding all else constant. Broadly speaking, the larger the cap diameter is the less likely the mushroom is poisonous.

- The coefficient for cap shape f is -1.508, suggesting that for mushrooms with a flat cap shape (labled "f"), we expect the odds of the mushroom being poisonous are 0.221 times the odds of being poisonous for mushrooms of bell shaped (baseline), granted all else are held constant (AOR = 0.221. The odds of the mushroom being poisonous decrease significantly when the cap shape is "p" compared to the baseline. As shown in bivariate EDA, the mushroom tends to be poisonous when cap shape is "bell."

- The coefficient for season(w) * habitat.modified(g) is -0.968. This indicates that the combination of specific season (winter) and habitat (grasses) conditions modifies the expected odds of the mushroom being poisonous by a multiplicity factor of 0.380 times the odds of the mushroom being poisonous through an additive model of the same combination, while holding all other variables constant. This trend can be seen in the interactive term EDA as when season is winter and habitat is grasses we are more likely to observe an edible mushroom.

Model Evaluation
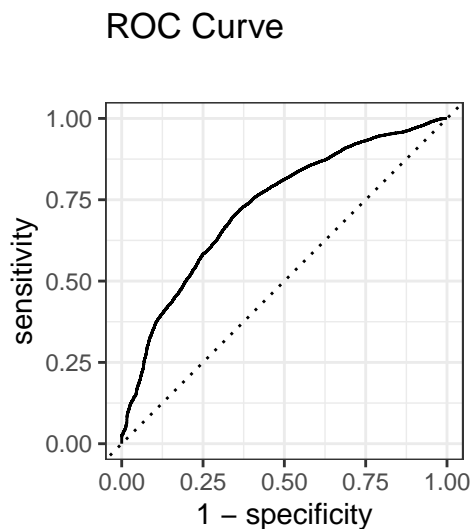
8

The final interactive model was evaluated with the testing data. The model is better than guessing but still not super great as the AUC is 0.728 which is right in between 1 and 0.5. We decided on a threshold of p = 0.294 to achieve a sensitivity of 95%, since we wanted to prioritize minimizing false negatives, which are more expensive – better to be careful than eat a poisonous mushroom classified as "edible".

Using this threshold, we can evaluate our model's performance with a confusion matrix. As desired, for poisonous mushrooms, we are able to successfully classify 95% of them as poisonous. Our model struggles at correctly identifying mushrooms which are actually edible as edible, with a false positive rate of ~81%. We were able to build a model that overall does much better (see appendix). However, this model requires many more predictors and becomes much more complex. For the purposes of this model, we wanted it to be applicable even in situations where humans found themselves having to assess edibility without many special tools or knowledge. Moreover, the initial interactive model also produced slightly better results but becomes harder to use in context for human foragers (see appendix).

Discussion + Conclusion

There are many other significant predictors that helps in the identification of mushroom edibility, but among those that are most useful and interpretable, cap shape, gill color, cap diameter, cap color, and their interactions (e.g., habitat × season, cap shape × cap color) speak the most on the probability of a mushroom being poisonous. We believe keeping a simple, parsimonious model with a high sensitivity was best for the context of our model's usefulness. Foragers should be able to use the model to help identify mushroom edibility. For example, the interaction of winter (season) and grasses (habitat) reduces the odds of toxicity significantly. In general, mushrooms present in the winter seem less likely to be toxic. Further, larger cap diameters are generally associated with lower odds of being poisonous. The presence of any gill color besides "none" increases the odds of the mushroom being poisonous.

The final interactive model performed decently with an AUC of 0.728, achieving 95% sensitivity with a threshold of p = 0.294. However, while the final interactive model seems to be useful, there are many benefits using each type of model. For example, the final main effects model may be more useful in context due to relatively few terms and its interpretability (Table 10). In comparison, on the other end of the spectrum, the interactive with less modified coefficients has the best AUC, but high p-values and N/A's likely because of the scarcity in data of specific combinations, and so some of these predictors may not be use in practice (Table 11). In the middle of the spectrum is the final interaction model that takes account of simplicity, due to the transformation of categorical variables to include only the most common features/environment, and produces a somewhat decent AUC (Table 12).

One limitation of our model is the high false positive rate, due to our decision to limit false negatives and improve safety for foragers. However, if a population of people are reliant on mushrooms as a critical food source, they may wish to be more liberal with classifying
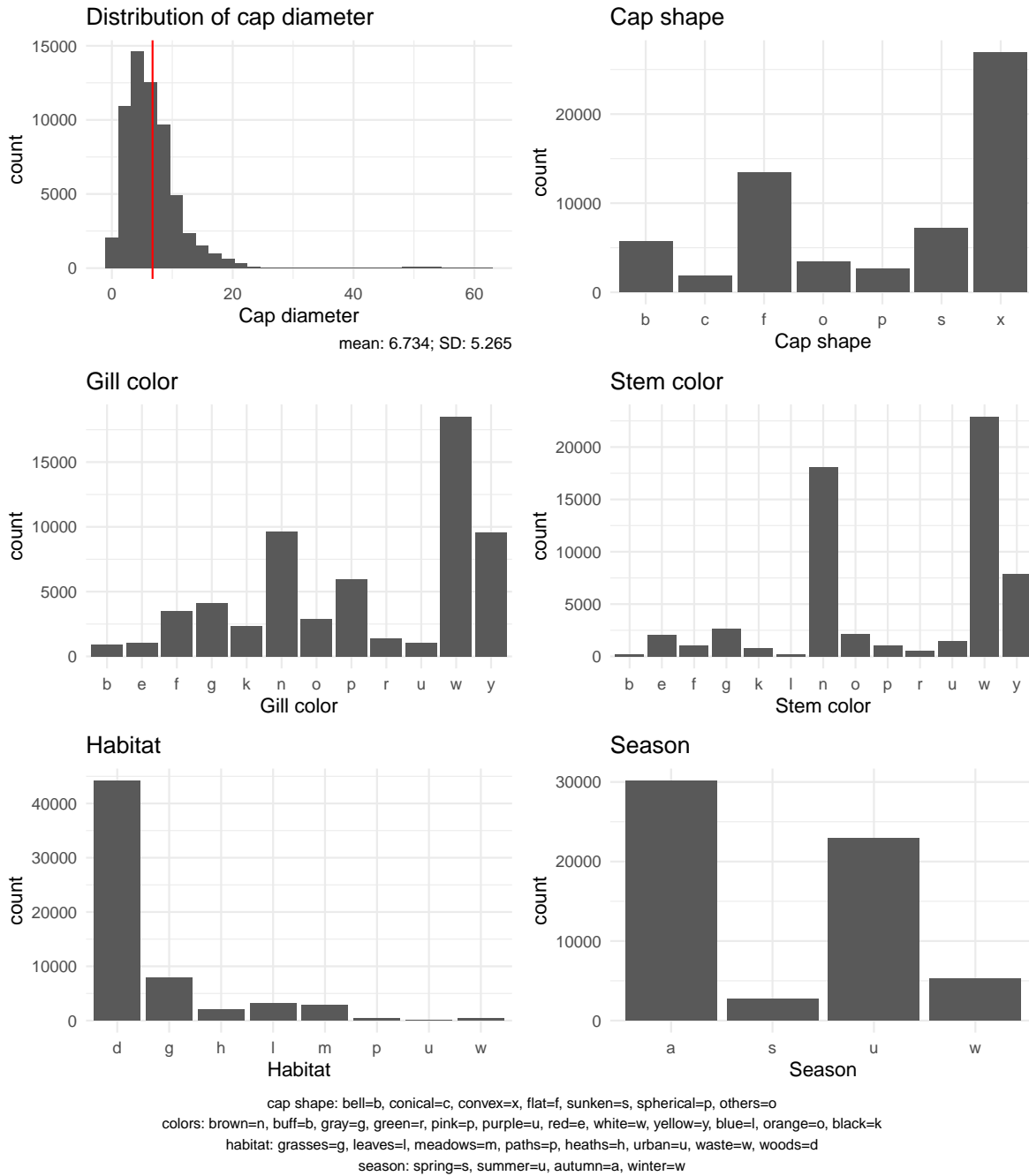
9

mushrooms as edible. Other limitations arise from the data, where some categories (e.g., rare gill, stem colors, and combinations of characteristics and environmental features) have sparse data. For these categories, we sometimes see high multicolinearity after running a VIF test (shown in appendix), which ends up inflating variances, biasing coefficients, and leading to predictors having coefficients with high p-values. Thus, some of these end up being very unreliable predictors. Further, the data is limited in how it represents mushrooms – the data is not able to represent mushrooms with features within the same category that occur simultaneously. For example, a mushroom's cap may be multiple colors, or its cap may be both shiny and smooth, but the dataset would only captures on of these characteristics. Lastly, the independence of observations for model assumption is assumed but not explicitly verified.

Our analysis could be improved by addressing these concerns with sparse data and high multicollinearity. While it's unclear if the high multicollinearity for some of the categories is a direct cause of the sparse data, it would be immensely helpful to expand the dataset to improve data balance for rare categories through targeted sampling. If multicollinearity still exists, it may be simply because dependencies simply exist in the population (ie, certain gill colors are associated with certain cap colors).

Further, our analysis may benefit from exploring other statistical techniques. The original paper the 2020 data was from found that the expanded dataset was not linearly separable, and so it would have been impossible to perfectly classify mushroom edibility through logistic regression (Dennis Wagner and Hattab 2021). They were able to achieve nearly perfect classification using random forests, so it may be worth leveraging similar techniques which can handle non-linear relationships while also remaining interpretable to see if we can achieve better prediction than our current model. For reference, their logistic regression model achieved an AUC of 0.85, but was more complex than ours. Unfortunately, the paper does not show their logistic regression model output (for comparison to ours), likely because the paper was mainly about expanding the dataset.

Future studies should also test the model with mushrooms from other regions or under different environmental conditions to assess robustness. Future applications of our model may also be possible to integrate our model with technology, with mobile applications that allow for real-time classification using the model. This may also allow for more complex models, if mobile devices are able to identify and measure characteristics which might be harder for humans. Finally, future works should focus on collaborating with public health organizations to create user-friendly guides or alerts for foragers, helping to make the model helps keep people safe.

Appendix

Distribution of cap diameter

mean: 6.734; SD: 5.265

Cap shape

Gill color

Stem color

Habitat

Season

cap shape: bell=b, conical=c, convex=x, flat=f, sunken=s, spherical=p, others=o
colors: brown=n, buff=b, gray=g, green=r, pink=p, purple=u, red=e, white=w, yellow=y, blue=l, orange=o, black=k
habitat: grasses=g, leaves=l, meadows=m, paths=p, heaths=h, urban=u, waste=w, woods=d
season: spring=s, summer=u, autumn=a, winter=w

Here is our original univariate EDA with the full visualizations.

cap.diameter	seasons	seasonu
1.367190	1.155352	1.116441
seasonw	cap.shapec	cap.shapef
1.111352	1.527727	3.742673
cap.shapeo	cap.shapep	cap.shapes
2.595727	1.698009	2.915693
cap.shapex	cap.colore	cap.colorg
4.712495	4.490393	6.050493
cap.colork	cap.colorl	cap.colorn
2.373001	2.022609	18.468343
cap.coloro	cap.colorp	cap.colorr
4.644798	2.953790	1.976855
cap.coloru	cap.colorw	cap.colory
3.079011	9.455068	10.657511
gill.colore	gill.colorf	gill.colorg
2.416240	6.826525	7.128370
gill.colork	gill.colorn	gill.coloro
4.581966	12.985624	5.225091
gill.colorp	gill.colorr	gill.coloru
9.977767	3.215672	2.625915
gill.colorw	gill.colory	stem.color.modifiedOther
20.853602	13.923760	1.628936
stem.color.modifiedw	stem.color.modifiedy	habitat.modifiedg
1.805273	1.617748	1.178873
habitat.modifiedOther		
1.135032		

Above is the output from the `vif()` function ran on our main effects model. There are a few categories with high VIF (>10), particularly some of the gill colors and cap colors, suggesting they may be correlated. This may be the result of low data for those specific gill colors and cap colors.

Distribution of Stem Height vs. Stem Width Among Different Edibility Classes
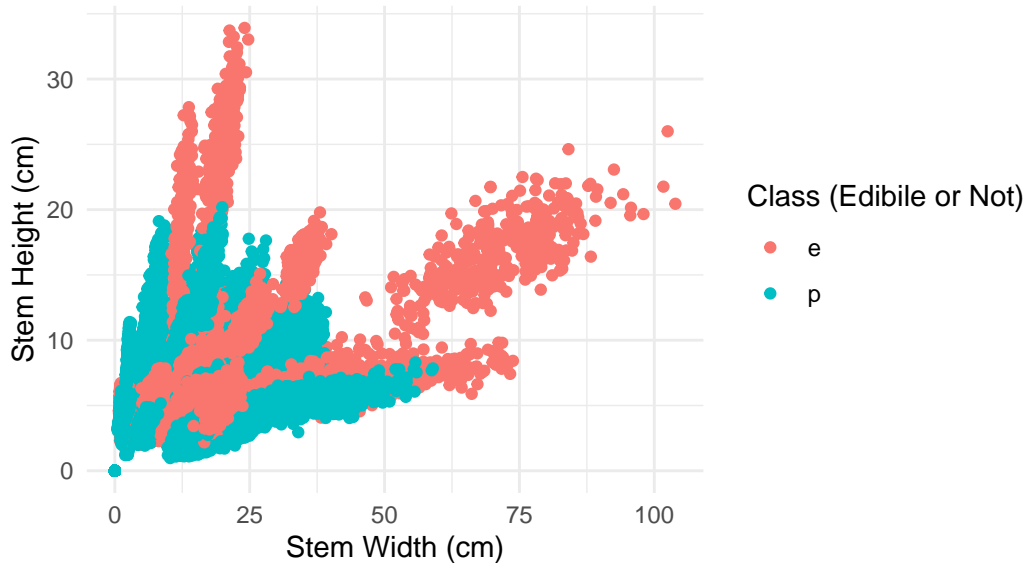
Here, we look at multivariate exploratory data analysis including 2 predictors and our response variable. We visualize the effect of both stem width and stem height on the response variable, class. Interestingly, it seems like mushrooms with either high stem width or stem height seem to be edible. This suggests there may be some potential interaction effects between stem height and stem width – the low value of one alone does not seem to predict if the mushroom is poisonous, but requires the low value of both. However, in our model when we added this interaction effect, the performance did not improve that drastically, and we deemed it more important to keep the model parsimonious as possible. Additionally, quantitative features can be hard to measure, and so may be less practical when serving as a general guideline for foraging mushrooms.
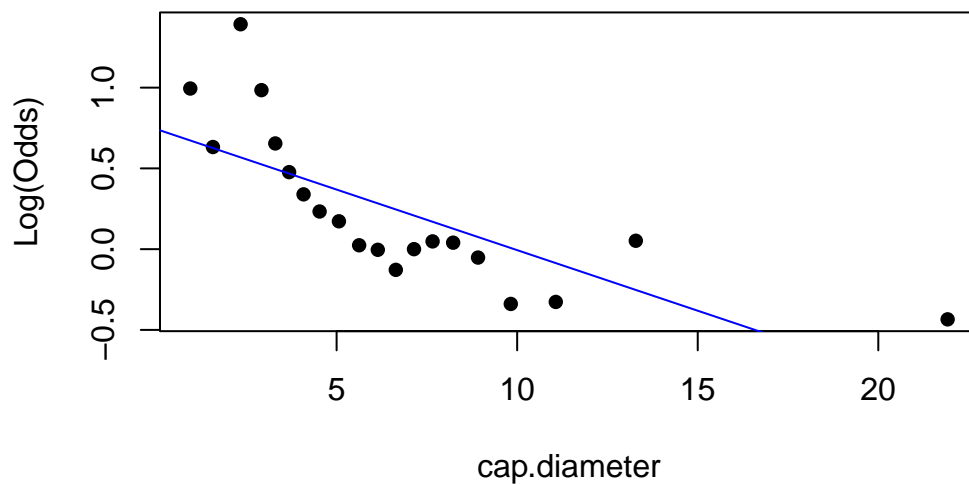
```
# A tibble: 7 x 5
  class_binary cap.shape     n  prop emp_logit
         <dbl> <chr>     <int> <dbl>     <dbl>
1            1 b          3103 0.777    1.25
2            1 c           688 0.561    0.244
3            1 f          4839 0.516    0.0636
4            1 o          1907 0.768    1.20
5            1 p           713 0.396   -0.421
6            1 s          2704 0.539    0.157
7            1 x          9768 0.518    0.0725


# A tibble: 12 x 5
   class_binary gill.color     n  prop emp_logit
          <dbl> <chr>      <int> <dbl>     <dbl>
 1            1 b            159 0.247    -1.12
 2            1 e            549 0.738     1.04
 3            1 f           1550 0.607     0.434
 4            1 g           1357 0.472    -0.113
 5            1 k            998 0.610     0.446
 6            1 n           4694 0.695     0.823
 7            1 o           1204 0.593     0.374
 8            1 p           2422 0.580     0.324
```

```
 9            1 r              590 0.624      0.508
10            1 u              413 0.569      0.277
11            1 w             5842 0.450     -0.202
12            1 y             3944 0.591      0.370
```

```
# A tibble: 4 x 5
  class_binary stem.color.modified      n  prop emp_logit
         <dbl> <chr>                 <int> <dbl>     <dbl>
1            1 Other                  5622 0.652     0.628
2            1 n                      7550 0.597     0.394
3            1 w                      6768 0.424    -0.307
4            1 y                      3782 0.686     0.779
```

```
# A tibble: 3 x 5
  class_binary habitat.modified      n  prop emp_logit
         <dbl> <chr>              <int> <dbl>     <dbl>
1            1 Other              3098 0.490   -0.0396
2            1 d                 16875 0.546    0.185
3            1 g                  3749 0.679    0.747
```

```
# A tibble: 4 x 5
  class_binary season      n  prop emp_logit
         <dbl> <chr>  <int> <dbl>     <dbl>
1            1 a      12143 0.576     0.305
2            1 s        825 0.440    -0.239
3            1 u       9319 0.579     0.317
4            1 w       1435 0.390    -0.448
```

Here is the calculated empirical logit plot for our only quantitative predictor variable and it appears to generally fit the linearity assumption. For the other qualitative predictor variables the empirical logic was calculated for the training dataset.

Table 8: Base Main Model Coefficients Filtered

term	estimate	std.error	statistic	p.value
(Intercept)	-15.328	172.082	-0.089	0.929
seasonu	0.003	0.020	0.156	0.876
cap.shapeo	0.173	0.077	2.253	0.024
gill.colorg	0.226	0.101	2.227	0.026
stem.colore	16.299	172.082	0.095	0.925
stem.colorf	31.777	185.273	0.172	0.864
stem.colorg	14.813	172.082	0.086	0.931
stem.colork	17.654	172.082	0.103	0.918
stem.colorl	14.725	172.082	0.086	0.932
stem.colorn	16.180	172.082	0.094	0.925
stem.coloro	15.797	172.082	0.092	0.927
stem.colorp	17.598	172.082	0.102	0.919
stem.colorr	16.734	172.082	0.097	0.923
stem.coloru	16.325	172.082	0.095	0.924
stem.colorw	15.400	172.082	0.089	0.929
stem.colory	16.390	172.082	0.095	0.924
habitath	0.053	0.054	0.987	0.324
habitatp	15.871	121.020	0.131	0.896
habitatu	-15.782	215.397	-0.073	0.942
habitatw	-16.260	126.274	-0.129	0.898

The base model tidy output with coefficients that have a p.value > 0.01 is shown here. To address these high p-value we collapsed some of the categorical variables into less levels.

Table 9: ANOVA Table for Drop-in-Deviance Test

term	df.residual	residual.deviance	df	deviance	p.value
class_binary ~ cap.diameter + season + cap.shape + cap.color + gill.color + habitat + stem.root * stem.color + veil.type + veil.color + has.ring + ring.type + cap.shape * cap.color + habitat * season	60923	47301.04	NA	NA	NA
class_binary ~ cap.diameter + season + cap.shape + cap.color + gill.color + habitat + stem.root * stem.color + veil.type + veil.color + has.ring * ring.type + gill.attachment * gill.spacing + cap.shape * cap.color + habitat * season	60905	41107.65	18	6193.389	0

```
The AIC for the first model is:  47593.04
 while the AIC for the second model is:  41435.65 .
 The BIC for the first model is:  48909.93
 and the BIC for the second model is:  42914.89 .
```

Here, we played with adding more predictors to our model. We do achieve better ROC curves with these as well, but we decided that a smaller model would still be better, and that many of these predictors that we added here may be hard to identify for the average person.

Table 10: Final Main Model Coefficients

term	estimate	std.error	statistic	p.value
(Intercept)	0.590	0.111	5.333	0.000
cap.diameter	-0.074	0.002	-36.247	0.000
seasons	-0.975	0.049	-20.006	0.000
seasonu	-0.005	0.020	-0.243	0.808
seasonw	-0.766	0.035	-21.870	0.000
cap.shapec	-0.881	0.064	-13.674	0.000
cap.shapef	-1.155	0.042	-27.246	0.000
cap.shapeo	1.027	0.074	13.920	0.000
cap.shapep	-1.003	0.056	-17.825	0.000
cap.shapes	-1.082	0.047	-22.843	0.000
cap.shapex	-1.335	0.040	-33.690	0.000
cap.colore	1.989	0.087	22.762	0.000
cap.colorg	0.557	0.085	6.515	0.000
cap.colork	0.887	0.102	8.674	0.000
cap.colorl	0.650	0.109	5.971	0.000
cap.colorn	0.358	0.079	4.512	0.000
cap.coloro	1.359	0.088	15.402	0.000
cap.colorp	1.614	0.099	16.296	0.000
cap.colorr	2.886	0.109	26.419	0.000
cap.coloru	1.207	0.094	12.784	0.000
cap.colorw	0.821	0.082	9.975	0.000
cap.colory	0.771	0.083	9.311	0.000
gill.colore	1.262	0.117	10.767	0.000
gill.colorf	-0.247	0.104	-2.378	0.017
gill.colorg	0.222	0.097	2.302	0.021
gill.colork	0.936	0.102	9.165	0.000
gill.colorn	1.435	0.093	15.498	0.000
gill.coloro	0.720	0.098	7.350	0.000
gill.colorp	0.983	0.094	10.476	0.000
gill.colorr	0.696	0.110	6.354	0.000
gill.coloru	0.867	0.113	7.701	0.000
gill.colorw	0.548	0.090	6.096	0.000
gill.colory	0.971	0.092	10.562	0.000
stem.color.modifiedOther	-0.128	0.030	-4.339	0.000
stem.color.modifiedw	-0.744	0.025	-29.649	0.000
stem.color.modifiedy	0.224	0.035	6.391	0.000
habitat.modifiedg	0.471	0.031	15.277	0.000
habitat.modifiedOther	-0.341	0.027	-12.391	0.000

Table 11: Interactive Model Coefficients

term	estimate	std.error	statistic	p.value
(Intercept)	15.736	169.081	0.093	0.926
cap.diameter	-0.060	0.002	-24.569	0.000
seasons	-1.155	0.066	-17.548	0.000
seasonu	-0.001	0.024	-0.031	0.975
seasonw	-0.551	0.041	-13.496	0.000
cap.shapec	1.031	219.970	0.005	0.996
cap.shapef	-16.156	169.081	-0.096	0.924
cap.shapeo	-16.219	169.081	-0.096	0.924
cap.shapep	-2.140	0.144	-14.870	0.000
cap.shapes	-32.263	535.647	-0.060	0.952
cap.shapex	-17.544	169.081	-0.104	0.917
cap.colore	1.412	413.344	0.003	0.997
cap.colorg	-15.703	169.081	-0.093	0.926
cap.colork	-16.227	169.081	-0.096	0.924
cap.colorl	1.198	443.343	0.003	0.998
cap.colorn	-14.905	169.081	-0.088	0.930
cap.coloro	0.776	345.966	0.002	0.998
cap.colorp	1.607	399.972	0.004	0.997
cap.colorr	1.206	362.511	0.003	0.997
cap.coloru	1.308	390.015	0.003	0.997
cap.colorw	-15.077	169.081	-0.089	0.929
cap.colory	1.097	0.166	6.621	0.000
gill.colore	1.508	0.126	11.966	0.000
gill.colorf	-0.409	0.118	-3.459	0.001
gill.colorg	0.666	0.108	6.170	0.000
gill.colork	1.382	0.113	12.243	0.000
gill.colorn	1.650	0.104	15.937	0.000
gill.coloro	1.083	0.109	9.924	0.000
gill.colorp	1.244	0.105	11.860	0.000
gill.colorr	0.085	0.126	0.672	0.502
gill.coloru	1.330	0.128	10.427	0.000
gill.colorw	0.899	0.101	8.913	0.000
gill.colory	1.227	0.102	11.979	0.000
stem.color.modifiedOther	-0.062	0.032	-1.972	0.049
stem.color.modifiedw	-0.871	0.026	-32.937	0.000
stem.color.modifiedy	0.284	0.037	7.769	0.000
habitat.modifiedg	0.407	0.045	9.090	0.000
habitat.modifiedOther	-0.259	0.040	-6.507	0.000
seasons:habitat.modifiedg	0.027	0.173	0.156	0.876

term	estimate	std.error	statistic	p.value
seasonu:habitat.modifiedg	0.283	0.063	4.462	0.000
seasonw:habitat.modifiedg	-1.026	0.122	-8.410	0.000
seasons:habitat.modifiedOther	0.363	0.148	2.455	0.014
seasonu:habitat.modifiedOther	-0.143	0.060	-2.390	0.017
seasonw:habitat.modifiedOther	-0.925	0.125	-7.412	0.000
cap.shapec:cap.colore	-36.577	668.979	-0.055	0.956
cap.shapef:cap.colore	0.220	413.344	0.001	1.000
cap.shapeo:cap.colore	18.278	498.350	0.037	0.971
cap.shapep:cap.colore	-31.905	625.853	-0.051	0.959
cap.shapes:cap.colore	32.214	685.439	0.047	0.963
cap.shapex:cap.colore	1.203	413.344	0.003	0.998
cap.shapec:cap.colorg	-16.785	567.768	-0.030	0.976
cap.shapef:cap.colorg	15.398	169.081	0.091	0.927
cap.shapeo:cap.colorg	19.802	169.081	0.117	0.907
cap.shapep:cap.colorg	1.377	0.219	6.288	0.000
cap.shapes:cap.colorg	32.125	535.647	0.060	0.952
cap.shapex:cap.colorg	17.142	169.081	0.101	0.919
cap.shapec:cap.colork	NA	NA	NA	NA
cap.shapef:cap.colork	16.314	169.081	0.096	0.923
cap.shapeo:cap.colork	34.895	342.679	0.102	0.919
cap.shapep:cap.colork	NA	NA	NA	NA
cap.shapes:cap.colork	32.985	535.647	0.062	0.951
cap.shapex:cap.colork	18.507	169.081	0.109	0.913
cap.shapec:cap.colorl	NA	NA	NA	NA
cap.shapef:cap.colorl	-0.922	443.343	-0.002	0.998
cap.shapeo:cap.colorl	NA	NA	NA	NA
cap.shapep:cap.colorl	NA	NA	NA	NA
cap.shapes:cap.colorl	-1.799	902.221	-0.002	0.998
cap.shapex:cap.colorl	0.280	443.343	0.001	0.999
cap.shapec:cap.colorn	-4.771	219.970	-0.022	0.983
cap.shapef:cap.colorn	14.639	169.081	0.087	0.931
cap.shapeo:cap.colorn	16.140	169.081	0.095	0.924
cap.shapep:cap.colorn	1.569	0.163	9.651	0.000
cap.shapes:cap.colorn	30.949	535.647	0.058	0.954
cap.shapex:cap.colorn	16.049	169.081	0.095	0.924
cap.shapec:cap.coloro	NA	NA	NA	NA
cap.shapef:cap.coloro	-0.430	345.966	-0.001	0.999
cap.shapeo:cap.coloro	17.829	444.909	0.040	0.968
cap.shapep:cap.coloro	-31.349	574.066	-0.055	0.956
cap.shapes:cap.coloro	16.579	614.835	0.027	0.978
cap.shapex:cap.coloro	1.365	345.966	0.004	0.997

term	estimate	std.error	statistic	p.value
cap.shapec:cap.colorp	NA	NA	NA	NA
cap.shapef:cap.colorp	-0.371	399.972	-0.001	0.999
cap.shapeo:cap.colorp	NA	NA	NA	NA
cap.shapep:cap.colorp	NA	NA	NA	NA
cap.shapes:cap.colorp	15.244	646.767	0.024	0.981
cap.shapex:cap.colorp	0.870	399.972	0.002	0.998
cap.shapec:cap.colorr	NA	NA	NA	NA
cap.shapef:cap.colorr	2.043	362.511	0.006	0.996
cap.shapeo:cap.colorr	17.198	460.562	0.037	0.970
cap.shapep:cap.colorr	NA	NA	NA	NA
cap.shapes:cap.colorr	17.697	624.294	0.028	0.977
cap.shapex:cap.colorr	2.058	362.511	0.006	0.995
cap.shapec:cap.coloru	-0.903	573.567	-0.002	0.999
cap.shapef:cap.coloru	-0.367	390.015	-0.001	0.999
cap.shapeo:cap.coloru	NA	NA	NA	NA
cap.shapep:cap.coloru	NA	NA	NA	NA
cap.shapes:cap.coloru	15.779	640.657	0.025	0.980
cap.shapex:cap.coloru	-0.761	390.015	-0.002	0.998
cap.shapec:cap.colorw	16.299	317.549	0.051	0.959
cap.shapef:cap.colorw	15.174	169.081	0.090	0.928
cap.shapeo:cap.colorw	34.217	227.081	0.151	0.880
cap.shapep:cap.colorw	NA	NA	NA	NA
cap.shapes:cap.colorw	31.121	535.647	0.058	0.954
cap.shapex:cap.colorw	17.000	169.081	0.101	0.920
cap.shapec:cap.colory	NA	NA	NA	NA
cap.shapef:cap.colory	-0.221	0.199	-1.116	0.265
cap.shapeo:cap.colory	NA	NA	NA	NA
cap.shapep:cap.colory	NA	NA	NA	NA
cap.shapes:cap.colory	15.562	508.261	0.031	0.976
cap.shapex:cap.colory	NA	NA	NA	NA

Interactive effects model's coefficient have a lot that are N/A's and high p-values for Wald test. In comparison the main effects model's coefficient tends to have low p-values for Wald test and enough data for each coefficient (observed in Table 10 above).

Table 12: Final Interactive Model Coefficients

term	estimate	std.error	statistic	p.value
(Intercept)	1.811	0.058	31.195	0.000
cap.diameter	-0.064	0.002	-33.344	0.000
season.modifiedOther	-0.095	0.022	-4.349	0.000
season.modifiedw	-0.543	0.037	-14.529	0.000
cap.shape.modifiedf	-1.508	0.060	-24.996	0.000
cap.shape.modifiedOther	-1.215	0.063	-19.208	0.000
cap.shape.modifieds	-1.466	0.069	-21.384	0.000
cap.shape.modifiedx	-1.475	0.055	-27.060	0.000
cap.color.modifiedOther	0.069	0.077	0.890	0.374
cap.color.modifiedw	-0.204	0.092	-2.220	0.026
gill.color.modifiedf	-0.494	0.047	-10.510	0.000
gill.color.modifiedn	0.742	0.029	25.261	0.000
gill.color.modifiedOther	-0.103	0.026	-3.957	0.000
gill.color.modifiedp	0.325	0.033	9.809	0.000
gill.color.modifiedy	0.286	0.030	9.674	0.000
stem.color.modifiedOther	0.015	0.028	0.549	0.583
stem.color.modifiedw	-0.843	0.024	-34.743	0.000
stem.color.modifiedy	0.123	0.032	3.792	0.000
habitat.modifiedg	0.315	0.042	7.507	0.000
habitat.modifiedOther	-0.273	0.036	-7.542	0.000
season.modifiedOther:habitat.modifiedg	0.301	0.059	5.097	0.000
season.modifiedw:habitat.modifiedg	-0.968	0.114	-8.455	0.000
season.modifiedOther:habitat.modifiedOther	-0.086	0.053	-1.631	0.103
season.modifiedw:habitat.modifiedOther	-1.111	0.118	-9.437	0.000
cap.shape.modifiedf:cap.color.modifiedOther	0.836	0.088	9.524	0.000
cap.shape.modifiedOther:cap.color.modifiedOther	1.457	0.098	14.932	0.000
cap.shape.modifieds:cap.color.modifiedOther	1.040	0.097	10.775	0.000
cap.shape.modifiedx:cap.color.modifiedOther	0.476	0.082	5.804	0.000
cap.shape.modifiedf:cap.color.modifiedw	0.521	0.113	4.626	0.000
cap.shape.modifiedOther:cap.color.modifiedw	1.461	0.116	12.585	0.000
cap.shape.modifieds:cap.color.modifiedw	0.408	0.123	3.311	0.001
cap.shape.modifiedx:cap.color.modifiedw	0.884	0.103	8.607	0.000

To address N/A values in the interactive model, we decided to transformed the other categorical variables where there were a lack of data on specific combinations of predictor variables (this model has every predictor variable modified except for `cap.diameter`). For cap shape "f", "s", "x", and "b" are kept. For stem color "w", "y", and "n" are kept. For cap color "n" and "w" are kept. For gill color "w", "y", "n", "p", and "f" are kept ("f" was releveled to the baseline).

For season "a" and "w" are kept. Features that were kept were collapsed to "Others". For more information on these labels please refer to the data dictionary here.

As seen in the table, there are less coefficients that have high p-values and no coefficients that are N/A's. Additionally, reducing the number of coefficients ultimately helps ease the model use for human foragers. It should be noted that the "Others" category typically has high p-values because they incorporate many different characteristics of mushrooms that can be label as either edible or poisonous (in turn this makes the coefficients for "Others" less relevant in context of applicable use since some characters in others can solely be edible or poisonous). Since most of these have low p-values, we believe the coefficients that have the largest magnitude are the most helpful predictors of edibility.

ROC Curve

Confusion matrix

p=0.247 (0–edible; 1–poisonous)
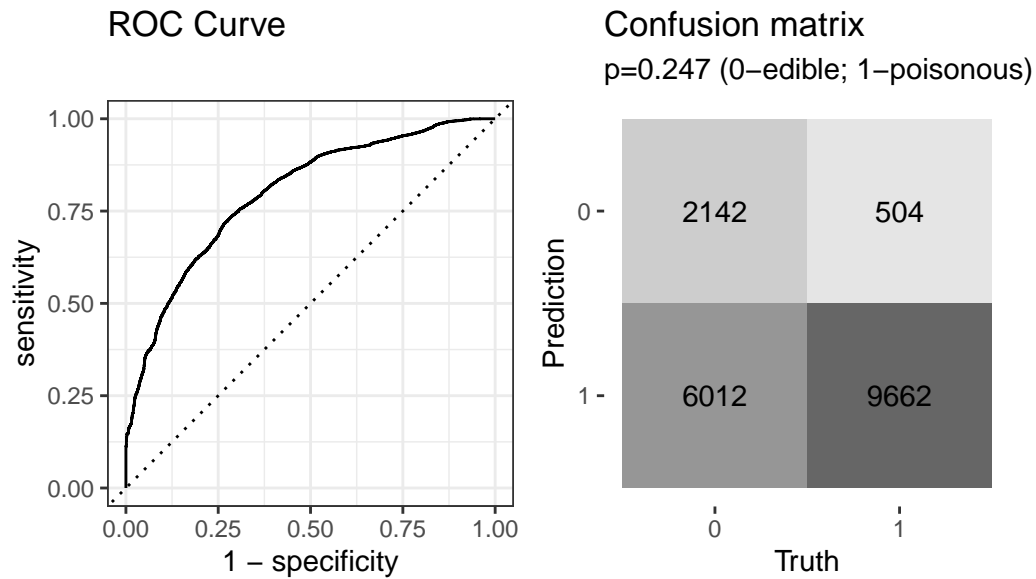
This is the ROC curve and confusion matrix for the initial interactive model. The model here is slightly better than our final interactive model as the AUC is 0.797 which is closer to 1 than 0.5. If we keep the sensitivity at 95% the threshold would be p = 0.247. Our confusion matrix here indicates that of poisonous mushrooms, we are able to successfully classify 95% of them as poisonous. It also suggests that our model struggles at correctly identifying mushrooms which are actually edible as edible, with a false positive rate of 73.7%, but still lower than the false positive rate of the final interactive model.

References

Brandenburg, William E., and Karlee J. Ward. 2018. "Mushroom Poisoning Epidemiology in the United States." *Mycologia* 110 (4): 637–41. https://doi.org/10.1080/00275514.2018.1479561.

Dennis Wagner, Dominik Heider, and Georges Hattab. 2021. "Mushroom Data Creation, Curation, and Simulation to Support Classification Tasks." *Scientific Reports* 11 (1). https://doi.org/10.1038/s41598-021-87602-3.