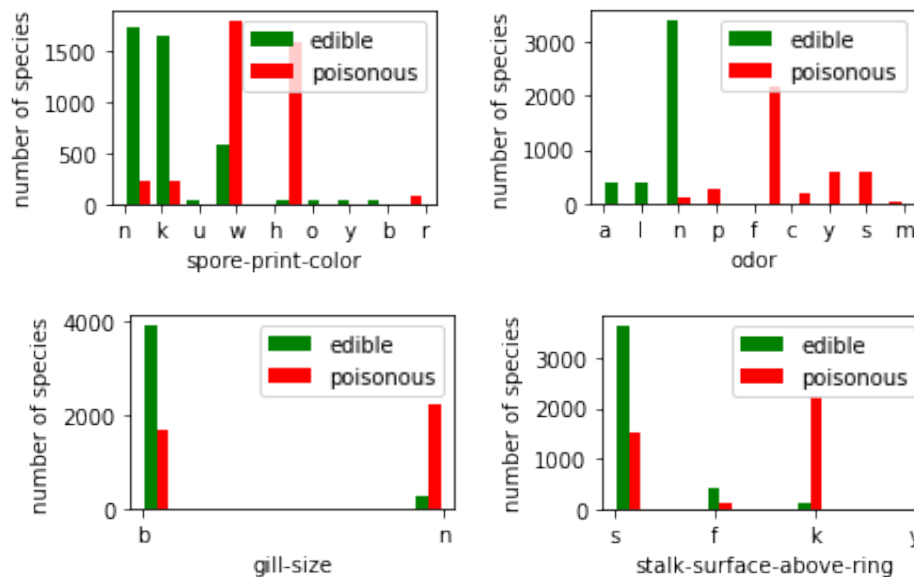


What characteristics make a mushroom most likely to be edible? What characteristics make a mushroom most likely to be poisonous? To what extent can we train machine learning models to predict whether a mushroom is edible or poisonous?

Data wrangling shows us that the target feature to predict is class (edible or poisonous), and exploratory data analysis shows us that there are various pairs of categorical features that are highly correlated or anti-correlated. After one-hot encoding these categorical features, we apply and compare five machine learning techniques. We find that support vector classification (SVC) is 100% accurate and has the best explanatory power, although logistic regression and random forest are also informative and perform just as well with 100% accuracy.

We are able to determine the four traits most strongly associated with poisonous mushrooms. They are, in order of importance, green spore prints (**r**), creosote odors (**c**), narrow gills (**n**), and silky stalk surfaces above the ring (**k**), with SVC coefficients 1.59, 1.05, 0.80, and 0.78 respectively. This agrees with the histograms in our exploratory data analysis:



We are also able to determine the four traits most strongly associated with edible mushrooms. They are, in order of importance, almond odors (**a**), anise odors (**l**), no odors (**n**), and crowded gill spacing (**w**), with SVC coefficients -1.02 , -1.02 , -0.84 , and -0.62 respectively. This agrees with the histograms in our exploratory data analysis:

