CSE514, Fall 2022, HW 4 Name: Student ID:
Note: This homework is worth a total of 15 points

Please refer to the
petNB classifier.xlsx
file posted on Canvas to find the probability values needed for the first three questions.

Q1 (4pts): Use the Naïve Bayes Classifier to label the following pet:

Color	Size	Temp	Baths	Pet type:
Yellow	2lb	Warm-blooded	Everyday	-

P(Cat | sample) [is proportional to]
P(Cat) * P(Yellow | Cat) * P(2lb | Cat) * P(Warm-blooded | Cat) * P(Everyday baths | Cat)
= 0.343 * 0.122 * 0.250 * 0.921 * 0.154 = 0.00148

P(Dog | sample) [is proportional to]
= 0.451 * 0.269 * 0.255 * 0.898 * 0.340 = 0.00945

P(Small mammal | sample) [is proportional to]
= 0.020 * 0.250 * 0.286 * 0.600 * 0.167 = 0.000143

P(Reptile | sample) [is proportional to]
= 0.078 * 0.214 * 0.077 * 0.182 * 0.167 = 0.0000391

P(Fish or amphibian | sample) [is proportional to]
= 0.049 * 0.091 * 0.100 * 0.250 * 0.222 = 0.0000247

P(Bird | sample) [is proportional to]
= 0.029 * 0.111 * 0.250 * 0.667 * 0.286 = 0.000154

P(Other | sample) [is proportional to]
= 0.029 * 0.111 * 0.125 * 0.333 * 0.286 = 0.0000383

Since the highest probability class is Dog, that's the predicted label

Q2 (4pts): A quick google search reveals that the number of pets in America is about:

Cats	= 58.4 mil
Dogs	= 76.8 mil
Small mammal	= 6.2 mil
Reptile	= 6.0 mil
Fish/amphibian	= 76.3 mil
Birds	= 22.9 mil
Other	= 4.7 mil
Total	= 251.3mil

P(Cats) = 58.4/251.3 = 0.232
P(Dogs) = 76.8/251.3 = 0.306
… = 6.2/251.3 = 0.025
… = 6.0/251.3 = 0.024
… = 76.3/251.3 = 0.304
… = 22.9/251.3 = 0.091
… = 4.7/251.3 = 0.019

Explain how this would change your classifier, and then re-classify the pet.

I would calculate new probabilities for each class label as above, and replace the "prior" probability values in my classifier

P(Cat | sample) [is proportional to]
P(Cat) * P(Yellow | Cat) * P(2lb | Cat) * P(Warm-blooded | Cat) * P(Everyday baths | Cat)
= 0.232 * 0.122 * 0.250 * 0.921 * 0.154 = 0.00100

P(Dog | sample) [is proportional to]
= 0.306 * 0.269 * 0.255 * 0.898 * 0.340 = 0.00641

P(Small mammal | sample) [is proportional to]
= 0.025 * 0.250 * 0.286 * 0.600 * 0.167 = 0.000179

P(Reptile | sample) [is proportional to]
= 0.024 * 0.214 * 0.077 * 0.182 * 0.167 = 0.0000120

P(Fish or amphibian | sample) [is proportional to]
= 0.304 * 0.091 * 0.100 * 0.250 * 0.222 = 0.000154

P(Bird | sample) [is proportional to]
= 0.091 * 0.111 * 0.250 * 0.667 * 0.286 = 0.000482

P(Other | sample) [is proportional to]
= 0.019 * 0.111 * 0.125 * 0.333 * 0.286 = 0.0000251

Since the highest probability class is still Dog, that's the predicted label

Q3 (4pts): Imagine that the size of the test pet was actually unit-less. The data collector forgot to record whether the pet was 2oz, 2g, 2lb, 2kg, or even 2tons. As such, your supervisor tells you to treat this value as missing. Explain how this would change your classification approach, and then re-classify the pet.
Ignore the information from Q2 for this problem.

I would simply drop the probability values for "size" during the calculations

P(Cat | sample) [is proportional to]
P(Cat) * P(Yellow | Cat) * P(Warm-blooded | Cat) * P(Everyday baths | Cat)
= 0.343 * 0.122 * 0.921 * 0.154 = 0.00594

P(Dog | sample) [is proportional to]
= 0.451 * 0.269 * 0.898 * 0.340 = 0.0370

P(Small mammal | sample) [is proportional to]
= 0.020 * 0.250 * 0.600 * 0.167 = 0.000501

P(Reptile | sample) [is proportional to]
= 0.078 * 0.214 * 0.182 * 0.167 = 0.000507

P(Fish or amphibian | sample) [is proportional to]
= 0.049 * 0.091 * 0.250 * 0.222 = 0.000247

P(Bird | sample) [is proportional to]
= 0.029 * 0.111 * 0.667 * 0.286 = 0.000614

P(Other | sample) [is proportional to]
= 0.029 * 0.111 * 0.333 * 0.286 = 0.000307

Since the highest probability class is still Dog, that's the predicted label

Q4 (3pts): Describe the difference between hard vs. soft margin classifiers, and give one advantage for each
(ie. Why would you pick a hard margin classifier over a soft margin classifier, and vice-versa)

A hard margin classifier requires that all the samples of each class are perfectly separated so that the positive class is on one side and the negative class is on the other

This classifier has the obvious advantage of perfect accuracy on the training data, as well as a simpler optimization goal of just finding the maximum margin, given this constraint of perfect accuracy. I might choose a hard margin classifier if my data is linearly separable and has little noise.

A soft margin classifier allows mis-classifications, balancing the two goals of large margins at the boundary and high accuracy of separating the two classes

This classifier is more robust to noisy data, and also allows a classifier to be fit even when data isn't linearly separable.