Israel Nolazco
Machine Learning
DATA 55000
July 22, 2020

SHORT ANSWER QUESTIONS

1. **Briefly describe why generalization bounds are important.**

Generalization bounds are important as it attempts to use a model as a general format to train and work with a dataset. In essence, generalization bounds are just attempting to create a hypothetical model that fits any data set vs the true model. The closer generalization bounds are, more closer the hypothetical to actual models are.

https://mostafa-samir.github.io/ml-theory-pt2/
https://towardsdatascience.com/generalization-bounds-rely-on-your-deep-learning-models-4842ed4bcb2a

2. **In a two-class, two-action problem if the loss function is λ11 = λ22 = 0, λ12 = 12, and λ21 = 5, write the optimal decision rule. (Follow the convention in Alpaydin exercise 3.4 and in the slides.) Show your work.**

#2

loss function $\lambda_{11} = \lambda_{22} = 0$

$\lambda_{12} = 12$, $\lambda_{21} = 5$

	C_1	C_2
a_1	0	12
a_2	5	0

$R(\alpha_1|x) = P(C_1|x)(\lambda_{11}) + P(C_2|x)(\lambda_{12})$

$R(\alpha_1|x) = P(C_1|x)(0) + P(C_2|x)(12)$

$R(\alpha_1|x) = 12(1 - P(C_1|x))$

$R(\alpha_2|x) = P(C_1|x)(5) + P(C_2|x)(0)$

$= (5)(P(C_1|x))$

$R(\alpha_1|x) < R(\alpha_2|x)$

$12(1 - P(C_1|x)) < 5(P(C_1|x))$

$12 - 12P < 5P$

$12 < 17P$

Choose α_1 if: $P(C_1|x) > \frac{17}{12}$

Otherwise choose α_2

Israel Nolazco
Machine Learning
DATA 55000
July 22, 2020

3. **How does the rule change if we add the third action of reject with λr = 1? Show your work.**

#3

	C_1	C_2
α_1	0	12
α_2	5	0
α_r	1	1

$R(\alpha_1|x) = 12 \cdot (1 - P(C_1|x))$

$R(\alpha_2|x) = 5\, P(C_1|x)$

$R(\alpha_r|x) = 1$

Choose α_1 if

$R(\alpha_1|x) < R(\alpha_r|x)$

$12 - 12 \cdot P(C_1|x) < 1$

$-12\, P(C_1|x) < -11$

$P(C_1|x) > 11/12$

Choose α_2 if

$R(\alpha_2|x) < R(\alpha_r|x)$

$P(C_1|x) < 1/5$

$P(C_2|x) > 4/5$

Reject otherwise

4. **In a two-class, two-action problem, use the decision rule for equal loss to solve for a posterior probabilities for each class for a new sample x = —0.1. The prior distributions are defined by point estimates with probability P(C1) = 0.45 and P(C2) = 0.55. The likelihoods are univariate normal: P(x|C1) having mean (μ) of 0.35 and standard deviation (σ) of 1.5; and P(x|C2) has a mean of 1.5 and standard deviation**

of 1. What are the posterior probabilities for the classes? Which action should be chosen, α1 or α2? (Hint: You can use the class Jupiter notebook to solve for the posterior probabilities for this problem, there is no need to calculate by hand.)

Problem 4.

```
[3]: %matplotlib inline
import matplotlib.pyplot as plt
import pandas as pd
import numpy as np
from scipy.stats import norm

prior_1 = 0.45
prior_2 = 1 - prior_1
liklihood_1_mean = 0.35
liklihood_1_std = 1.5
liklihood_2_mean = 1.5
liklihood_2_std = 1

x_i = -0.1

# normal density function: loc = mean, scale = standard deviation
print ("Posterior 1: ", norm.pdf(x_i, loc=liklihood_1_mean, scale=liklihood_1_std)*prior_1 /  \
    (norm.pdf(x_i, loc=liklihood_1_mean, scale=liklihood_1_std)*prior_1+  \
    norm.pdf(x_i, loc=liklihood_2_mean, scale=liklihood_2_std)*prior_2))
print ("Posterior 2: ", norm.pdf(x_i, loc=liklihood_2_mean, scale=liklihood_2_std)*prior_2 /  \
    (norm.pdf(x_i, loc=liklihood_1_mean, scale=liklihood_1_std)*prior_1+  \
    norm.pdf(x_i, loc=liklihood_2_mean, scale=liklihood_2_std)*prior_2))

Posterior 1:  0.6522318775961008
Posterior 2:  0.3477681224038992
```

Based on the results we have got. It is logical to use Alpha 1.

5. **Use the decision rule you found in problem 2 to interpret the posterior probability from problem 4. Which action should we choose, α1 or α2?**

The results on my optimal decision rule came with the conclusion to choose alpha 1 if $P(C1|x) > 17/12$). Since the posterior results shows to be less than that value, then we must choose Alpha 2.

6. **Use the decision rule with reject you found in problem 3 to interpret the posterior probability from problem.**

The results of the decision rule state that we choose Alpha 1 if $P(C1|x) < 11/12$ or Choose Alpha 2 if $P(C2|x) > 4/5$ Since, neither posteriors fall in that range than we much reject alphas.

4. **Which action should we choose, α1, α2, or αr?**

Use Alpha r.

7. **Briefly describe a hypothetical situation where a decision rule with reject would be helpful in a real-world problem.**

In the real world, test hypotheses are constantly run for a variety of reasons one of them could potentially be a vaccine. Say for example the vaccine is meant to cure a disease and to get it approve there needs to be a significant level of 0.05, but after testing and gathering data a p-value is calculated to be 0.20. Therefore, the vaccine seems to be unreliable and not up to the standard required to be considered a "cure."
https://examples.yourdictionary.com/examples-of-hypothesis-testing.html

8. **For the table below, calculate the support and confidence of the following rules (show the fractional values, do not calculate with decimals):**

broccoli -> rice
rice -> broccoli
tea -> beets
beets -> tea
Customer 1: broccoli, rice, tea
Customer 2: broccoli, tea, beets
Customer 3: broccoli
Customer 4: rice, tea, beets
Customer 5: broccoli
Customer 6: broccoli, rice
Customer 7: beets, tea

customer	Items
1	broccoli, rice, tea
2	
3	broccoli, tea, beets
4	broccoli
5	rice, tea, beets
	broccoli
6	broccoli, rice
7	broccoli, tea

$$\text{support}(X \to Y) = \frac{\#\text{ transactions with } X \,\&\, Y}{\text{Total } \#\text{ of transactions}}$$

support of

(broccoli, rice) $= \frac{2}{7}$

(rice, broccoli) $= \frac{0}{7} = 0$

(tea, beets) $= \frac{2}{7} = \frac{2}{7}$

(beets, tea) $= \frac{0}{7} = 0$

confidence of

(broccoli, rice) $= \frac{2}{6} = \frac{1}{3}$

(rice, broccoli) $= \frac{0}{3} = 0$

(tea, beets) $= \frac{2}{4} = \frac{1}{2}$

(beets, tea) $= \frac{0}{2} = 0$

9. **Interpret the results of problem 8.**

Support measures how frequent an itemset is in all the transactions. Therefore, given the results from Problem 8, we see that (broccoli, rice) and (beets, tea) are equally bought together, 2 out of 7 times to be precise. However, buying (rice, broccoli) are not bought together. This, in essence, means a customer is more than likely buying rice when purchasing broccoli than buying broccoli when purchasing rice.

Confidence measures how often the itemset is found to be true. Therefore, given the results from Problem 8, we found the confidence of broccoli and rice being purchase together with a total of

⅓. Additionally, tea and beets are purchase together with a total of ½. Therefore, we are more confident that tea and beets are purchase together than broccoli and rice.