CS540 Summer 2021

- Home
- Week 1
- Week 2
- Week 3
- Week 5
- Week 6
- Week 7
- Midterm
- Final

• Enter your ID (the wisc email ID without @wisc.edu) here: yw and click Confirm (or hit enter key)

$w = w - \lambda \frac{\partial C}{\partial w}$

$b = b - \lambda \frac{\partial C}{\partial b}$

$C = 1 - a_i \cdot y_i$

Question 1

• [4 points] Consider a linear model $a_i = w^T x_i + b$, with the hinge cost function $\max\left(0, 1 - a_i \cdot y_i\right)$. The initial weight is $\begin{bmatrix} w \\ b \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. What is the updated weight and bias after one stochastic gradient (sub)gradient descent step if the chosen training data is x = -3, y = 1? The learning rate is 2.

• Answer (comma separated vector): 1-2*(3),0-2*(-1) Calculate

$y_i = 1$

$a_i = w^T x_i + b = (1)(-3) + 0 = -3$

$\frac{\partial C}{\partial a_i} \cdot \frac{\partial a_i}{\partial w}$

$-1 \qquad (-3)$

Question 2

• [4 points] What is the gradient magnitude of the center element (pixel) of the image $\begin{bmatrix} 3 & 0 & -5 \\ 3 & -9 & 7 \\ -5 & -3 & -10 \end{bmatrix}$. Use the x gradient filter: $\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$ and the y gradient filter: $\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$. Remember to flip the filters.

• Answer: sqrt((3+6-9+5-14+10)^2+(3-5 Calculate

$\nabla_x \qquad D_y$

$G = \sqrt{(3+6-9+5-14+10)^2 + (3+0-5+9+6+10)^2}$

Question 3

• [4 points] Given the following transition matrix for a bigram model with words "I" (label 0), "am" (label 1) and "Groot" (label 2): $\begin{array}{c} \text{I} \\ \text{am} \\ \text{Groot} \end{array} \begin{bmatrix} 0.43 & 0.35 & 0.22 \\ 0.39 & 0.11 & 0.5 \\ 0.21 & 0.59 & 0.2 \end{bmatrix}$. Row i column j is $\mathbb{P}\left\{w_t = j | w_{t-1} = i\right\}$. Two uniform random numbers between 0 and 1 are generated to simulate the words after "I", say u_1 = 0.32 and u_2 = 0.63. Using the CDF inversion method, which two words are generated? Enter two integer labels (0, 1, or 2), not strings.

• Answer (comma separated vector): 0,1 Calculate

I, I, am

$u_1 \quad u_2$

0.43 0.35 0.22

I am Groot

Question 4

• [4 points] Consider a kernel $K(x, y) = 7 \cdot x \cdot y + 5 \cdot 1 + 3 \cdot 2^{x+y}$, where both x and y are positive real numbers. What is the feature vector $\varphi\left(x\right)$ induced by this kernel evaluated at x = 1?

• Answer (comma separated vector): sqrt(7),sqrt(5),sqrt(3)*2 Calculate

$(\sqrt{7}, \sqrt{5}, \sqrt{3} \cdot 2^1)$

$K(x,y) = \phi^T(x)\phi(y) = \begin{bmatrix} \sqrt{7}x \\ \sqrt{5} \\ \sqrt{3} \cdot 2^x \end{bmatrix}^T \begin{bmatrix} \sqrt{7}y \\ \sqrt{5} \\ \sqrt{3}\, 2^y \end{bmatrix}$

Question 5

• [4 points] You have a data set with 32 positive items and 26 negative items. You perform a "leave-one-out" procedure: for each item i, learn a separate kNN (k Nearest Neighbor) classifier on all items except item i, and compute that kNN's accuracy in predicting item i. The leave-one-out accuracy is defined to be the average of the accuracy for each item. What is the leave-one-out accuracy when k = 57?

• Answer: 32/(32+26) Calculate

58 → 1 57 → positive $\frac{32}{32+26}$

Question 6

• [4 points] John tells his professor that he forgot to submit his homework assignment. From experience, the professor knows that students who finish their homework on time forget to turn it in with probability 0.63. She also knows that 0.49 of the students who have not finished their homework will tell her they forgot to turn it in. She thinks that 0.25 of the students in this class completed their homework on time. What is the probability that John is telling the truth (i.e. he finished it given that he forgot to submit it)?

• Answer: (0.63*0.25)/(0.25*0.63+0.75* Calculate

$P_r\{F_{in} | F_{or}\} = \frac{P_r\{F_{or} | F_{in}\} \cdot P_r\{F_{in}\}}{P_r\{F_{or} | F_{in}\} \cdot P_r\{F_{in}\} + P_r\{F_{or} | \neg F_{in}\} \cdot P_r\{\neg F_{in}\}}$

0.63 0.25

0.63 · 0.25

0.49 (1 − 0.25)

$P_r\{F_{or} | F_{in}\} = 0.63$

$P_r\{F_{or} | \neg F_{in}\} = 0.49$

$P_r\{F_{in}\} = 0.25$

Question 7

• [4 points] Given the number of instances in each class summarized in the following table, how many instances are used to train an one-vs-one SVM (Support Vector Machine) for class 3 vs 2?

one vs all

y_i	0	1	2	3	4
Count	25	65	10	86	6

• Answer: 10+86 Calculate

Question 8

• [4 points] Say we have a training set consisting of 25 positive examples and 32 negative examples where each example is a point in a two-dimensional, real-valued feature space. What will the classification accuracy be on the training set with 1NN (Nearest Neighbor).

• Answer: 1 Calculate

57 → majority negative

if k = 57, $\frac{32}{25+32}$.

Question 9

• [4 points] What is the conditional entropy $H\left(B|A\right)$ for the following set of training examples.

item	A	B
1	T	F
2	F	F
3	F	T

$H(B | A=T) \cdot P_r\{A=T\} \longrightarrow \frac{3}{8} \cdot \left(-\frac{1}{3}\log_2\frac{1}{3} - \frac{2}{3}\log_2\frac{2}{3}\right) +$

$+H(B \mid A=F)\cdot \Pr\{A=F\} \longrightarrow \frac{5}{8}\left(-\frac{2}{5}\log_2\frac{2}{5} - \frac{3}{5}\log_2\frac{3}{5}\right)$

4	T	F
5	F	F
6	T	T
7	F	F
8	F	T

• Answer: -3/8*(1/3*log2(1/3)+2/3*log2(. Calculate

Question 10

• [4 points] There are 50 parrots. They have either a red beak or a black beak. They can either talk or not. Complete the two cells in the following table so that the mutual information (i.e. information gain) between "Beak" and "Talk" is 0.07: 0

Number of parrots	Beak	Talk
16	Red	Yes
?	Red	No
??	Black	Yes
16	Black	No

• Answer (comma separated vector): 12,6 . Calculate

Question 11

• [4 points] In a convolutional neural network, suppose the activation map of a convolution layer is

$$\begin{bmatrix} 3 & 0 & -5 & 3 \\ -9 & 7 & -9 & -3 \\ -10 & -4 & 3 & -6 \\ -2 & 1 & -8 & 5 \end{bmatrix}$$

. What is the activation map after a non-overlapping (stride 2) 2 by 2 max-pooling layer?

• Answer (matrix with multiple lines, each line is a comma separated vector):
```
7,3
1,5
```
. Calculate

Question 12

• [4 points] Some Na'vi's don't wear underwear, but they are too embarrassed to admit that. A surveyor wants to estimate that fraction and comes up with the following less-embarrassing scheme: Upon being asked "do you wear your underwear", a Na'vi would flip a fair coin outside the sight of the surveyor. If the coin ends up head, the Na'vi agrees to say "Yes"; otherwise the Na'vi agrees to answer the question truthfully. On a very large population, the surveyor hears the answer "Yes" for 0.82. What is the estimated fraction of Na'vi's that don't wear underwear? Enter a fraction like 0.01 instead of a percentage 1%.

• Answer: 1-(0.82-0.5)/0.5 . Calculate

$p = \frac{0.82-0.5}{0.5}$

p wear

$\underbrace{0.5 + 0.5\cdot p}_{\%\ \text{Yes.}} = 0.82$

Question 13

• [4 points] Fill in the missing values in the following joint probability table so that A and B are independent.

-	A = 0	A = 1
B = 0	$\frac{4}{14}$	$\frac{3}{14}$
B = 1	?? p	?? $\frac{1}{2} - p$

$\Pr\{A \mid B\} = \Pr\{A\}$

$\Pr\{A, B\} = \Pr\{A\}\cdot\Pr\{B\}$

• Answer (comma separated vector): (4/14)/(1/2)-4/14,1/2-((4/14)/(. Calculate

$\Pr\{A=0\}\cdot\Pr\{B=0\} = \left(\frac{4}{14} + p\right)\left(\frac{4}{14} + \frac{3}{14}\right) = \frac{4}{14} = \Pr\{A=0, B=0\}$

$p = \frac{\frac{4}{14}}{1/2} - \frac{4}{14}$

Question 14

• [4 points] Say we use Naive Bayes in an application where there are 4 features represented by 4 variables, each having 3 possible values, and there are 5 classes. How many total probabilities must be stored in the CPTs (Conditional Probability Table) in the Bayesian network for this problem? Do not include probabilities that can be computed from other numbers?

• Answer: 5*2*4+4 . Calculate

Y

$X_1 \quad X_2 \quad X_3 \quad X_4$

Question 15

• [1 points] Please enter any comments including possible mistakes and bugs with the questions or your answers. If you have no comments, please enter "None": do not leave it blank.

• Answer: None. .

$1 - \Pr\{X_1 = 1 \mid Y=y\} - \Pr\{X_1 = 2 \mid Y=y\} = \Pr\{X_1 = 3 \mid Y=y\}$

$\underbrace{\Pr\{X_i \mid Y\}}_{2\times 5} \quad \underbrace{i=1,2,3,4}_{\times 4}, \quad \underbrace{\Pr\{Y\}}_{+\ 4}$

Grade

Grade

***** ***** ***** ***** *****

ID: yw

Question 1 is correct. (4/4)

Question 2 is correct. (4/4)

Question 3 is correct. (4/4)

Question 4 is correct. (4/4)

Question 5 is correct. (4/4)

Question 6 is correct. (4/4)

Question 7 is correct. (4/4)

Question 8 is correct. (4/4)
Question 9 is correct. (4/4)
Question 10 is correct. (4/4)
Question 11 is correct. (4/4)
Question 12 is correct. (4/4)
Question 13 is correct. (4/4)
Question 14 is correct. (4/4)
Question 15 is correct. (1/1)
Grade: 57 out of 57.
Code: tz79"9&9"9&9"9&9"9&9"9&9"9&9"9&9"9&9"9&9"9&9"9&9"9&9"9&9"9&<"<&08:"8:
***** ***** ***** ***** *****

```
##m: 13
##id: yw
##1: 1-2*(3),0-2*(-1)
##2: sqrt((3+6-9+5-14+10)^2+(3-
```

• You could save the text in the above text box to a file using the button Download or copy and paste it into a file yourself Copy.
• You could load your answers from the text (or txt file) in the text box below using the button Load. The first two lines should be "##m: 13" and "##id: your id", and the format of the remaining lines should be "##1: your answer to question 1" newline "##2: your answer to question 2", etc. Please make sure that your answers are loaded correctly before submitting them.
Choose File M13Y.txt

```
##m: 13
##id: yw
```

Last Updated: July 15, 2021 at 4:21 PM