

• Answer: $5/(3+7+1+4)$ Calculate

Question 6

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

3	0	-5	3
-9	7	0	-3
-10	-4	3	-6
-2	1	-8	5

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

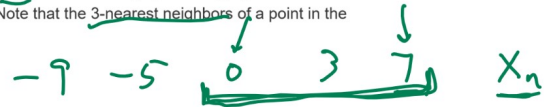
$$\frac{3+0-7+7}{4}, \dots$$

• Answer (matrix with multiple lines, each line is a comma separated vector): $(3,-9,7)/4, (-5+3-9-2)/4$ Calculate

Question 7

• [3 points] Let a dataset consists of $n = 6$ points in \mathbb{R} , specifically, the first $n - 1$ points are $[-9 \ -5 \ 0 \ 3 \ 7]$ and the last point x_n is unknown. What is the smallest value of x_n above which x_{n-1} is among x_n 's 3-nearest neighbors, but x_n is NOT among x_{n-1} 's 3-nearest neighbor? Note that the 3-nearest neighbors of a point in the training set include the point itself.

• Answer: $7+(7-0)$ Calculate



Question 8

• [4 points] What is the gradient magnitude of the center element (pixel) of the image $\begin{bmatrix} 3 & 0 & -5 \\ 3 & -9 & 7 \\ -9 & -3 & -10 \end{bmatrix}$ Use the

x gradient filter: $\begin{bmatrix} -1 & 0 & 1 \end{bmatrix}$ and the y gradient filter: $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$. Remember to flip the filters.

$$|x_n - 7| > |7 - 0|$$

$$x_n > 14$$

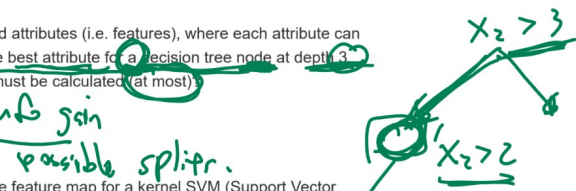
• Answer: $\sqrt{(-4)^2 + (-3)^2}$ Calculate

$$\sqrt{(3-7)^2 + (0+3)^2}$$

Question 9

• [4 points] In a problem where each example has 9 real-valued attributes (i.e. features), where each attribute can be split at 3 possible thresholds (i.e. binary splits), to select the best attribute for a decision tree node at depth 3 where the root is at depth 0, how many conditional entropies must be calculated (at most)?

• Answer: $9 \cdot 3 \cdot 3$ Calculate



Question 10

• [4 points] Given two instances $x_1 = 3$ and $x_2 = 0$, suppose the feature map for a kernel SVM (Support Vector Machine) is $\varphi(x) = \begin{bmatrix} x \\ x^3 \\ 1 \end{bmatrix}$, what is the kernel (Gram) matrix?

• Answer (matrix with multiple lines, each line is a comma separated vector): $1, 1$ Calculate

$$K_{ij} = \phi^T(x_i) \phi(x_j)$$

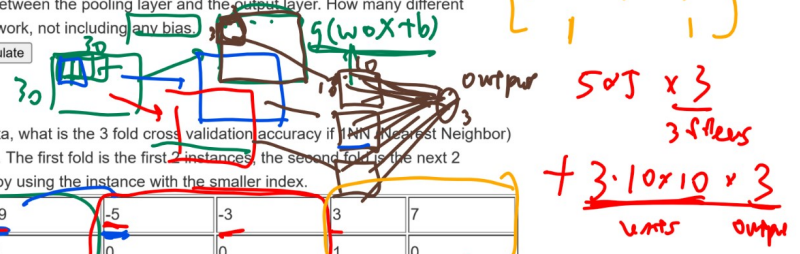
$$\phi_1 = \phi(x_1) = \begin{pmatrix} 3 \\ 27 \\ 1 \end{pmatrix} \quad \phi_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$K = \begin{bmatrix} \phi_1^T \phi_1 & \phi_1^T \phi_2 \\ \phi_2^T \phi_1 & \phi_2^T \phi_2 \end{bmatrix}$$

Question 11

• [4 points] A convolutional neural network has input image of size 30×30 that is connected to a convolutional layer that uses a 5×5 filter, zero padding of the image, and a stride of 1. There are 3 activation maps. The convolutional layer is then connected to a pooling layer that uses 3×3 max pooling, a stride of 3 (non-overlapping), and no padding of the convolutional layer. The pooling layer is then fully connected to an output layer that contains 3 output units. There are no hidden layers between the pooling layer and the output layer. How many different weights must be learned in this whole network, not including any bias?

• Answer: $5 \cdot 5 \cdot 3 + (30/3) \cdot (30/3) \cdot 3 \cdot 3$ Calculate



Question 12

• [4 points] Given the following training data, what is the 3 fold cross validation accuracy if 1NN (Nearest Neighbor) classifier with Manhattan distance is used. The first fold is the first 2 instances, the second fold is the next 2 instances, etc. Break the tie (in distance) by using the instance with the smaller index.

x_i	-10	-9	-5	-3	3	7
y_i	0	1	0	0	1	0

• Answer: $(1+0+1)/3$ Calculate

Question 13

• [4 points] What is the convolution between the image $\begin{bmatrix} 3 & 0 & -5 \\ 3 & -9 & 7 \\ -9 & -3 & -10 \end{bmatrix}$ and the filter $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ using zero padding? Remember to flip the filter first.

• Answer (matrix with multiple lines, each line is a comma separated vector): $3, -9, 7, -9, -3, -10$ Calculate

Question 14

• [4 points] Given the following transition matrix for a bigram model with words "I", "am" and "Groot": $\begin{bmatrix} 0.43 & 0.35 & 0.22 \\ 0.39 & 0.11 & 0.5 \\ 0.21 & 0.59 & 0.2 \end{bmatrix}$. Row i column j is $\mathbb{P}\{w_t = j | w_{t-1} = i\}$. What is the probability that the third word is "I" given the first word is "am"?

• Answer: $0.39 \cdot 0.43 + 0.11 \cdot 0.39 + 0.5 \cdot 0.2$ Calculate

$$P_r\{X_3 = 0 | X_1 = 1\} = 0.39 \cdot 0.43 + 0.11 \cdot 0.39 + 0.5 \cdot 0.2$$

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:

Grade

ID: yw

Question 1 is correct. (4/4)

Question 2 is correct. (5/5)

Question 3 is correct. (3/3)

Question 4 is correct. (4/4)

Question 5 is correct. (4/4)

Question 6 is correct. (4/4)

Question 7 is correct. (3/3)

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: 56 out of 56.

Code: tz79"9&&"8&>">&9"9&&9"9&9"9&>">&9"9&9"9&&9"9&&9"9&&9"9&9"9&<"<&08;"8;

```
##m: 14
##id: yw
##1: 0.49/5
##2: 0.63*
```

• You could save the text in the above text box to a file using the button or copy and paste it into a file yourself .

• You could load your answers from the text (or txt file) in the text box below using the button . The first two lines should be "##m: 14" 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.

M14Y.txt
##m: 14
##id: yw

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



UNIVERSITY OF WISCONSIN-MADISON