CSE 493 G1/ 599 G1
Deep Learning
Autumn 2024 Quiz 1

SOLUTIONS

2024

Full Name: _____

UW Net ID: _____

Question		Score
True/False	(4 pts)	
Multiple Choice	(8 pts)	
Short Answer	(8 pts)	
Total	(20 pts)	

Welcome to the CSE 493 G1 Quiz 1!

- The exam is 30 min and is **double-sided**.
- No electronic devices are allowed.

I understand and agree to uphold the University of Washington Honor Code during this exam.

Signature: _____ Date: _____

Good luck!

This page is left blank for scratch work only. DO NOT write your answers here.

1 True / False (4 points) - Recommended 4 Minutes

Fill in the circle next to True or False, or fill in neither. Fill it in completely like this: ●. No explanations are required.

Scoring: Correct answer is worth 1 points.

1.1 A KNN classifier always has 100% accuracy on the *train* set for all values of k.

◯ True

◯ False

SOLUTION:
False, This is only true for k=1

1.2 Relu is a linear function but softmax is not.

◯ True

◯ False

SOLUTION:
False, Relu is non-linear (see lecture 4)

1.3 Consider the following loss function

$$L(w) = \frac{1}{N}\sum_{i=1}^{N} L_i(f(x_i, w), y_i) + \lambda R(w)$$

where $R(w)$ is L2 regularization and L is a standard training loss. If we set $\lambda = -1$, then for every w, there exists another set of weights w' with lower loss, but an accuracy less than or equal to that of w.

◯ True

◯ False

SOLUTION:
True. Setting $\lambda = -1$ now means that larger weights have a lower loss, so you could use a multiple of w

1.4 For KNN classifiers, increasing the value of k generally leads to a more complex decision boundary.

◯ True

◯ False

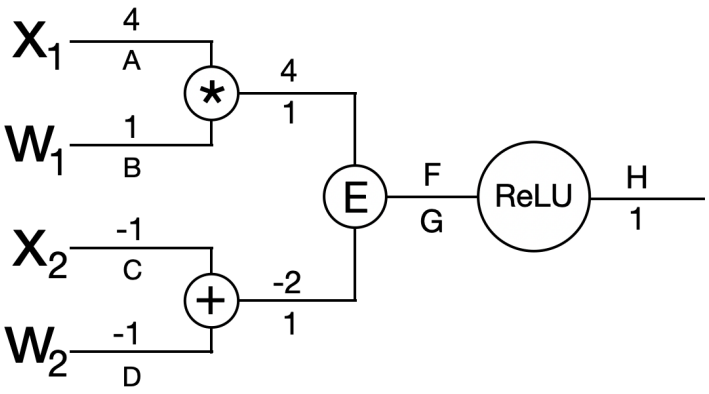
SOLUTION:
False, it is the opposite

2 Multiple Choices (8 points) - Recommended 8 Minutes

Fill in the circle next to the letter(s) of your choice (like this: ●). No explanations are required. **Choose ALL options that apply.**

Each question is worth 4 points and the answer may contain one or more options. Selecting all of the correct options and none of the incorrect options will get full credits. For questions with multiple correct options, each incorrect or missing selection gets a 2-point deduction (up to 4 points).

2.1 Below is a network made entirely of multiply, add, and ReLU gates. The values computed during the forward pass are above each line and the values computed during the backwards pass are below each line. Some values and operations have been replaced with letters.

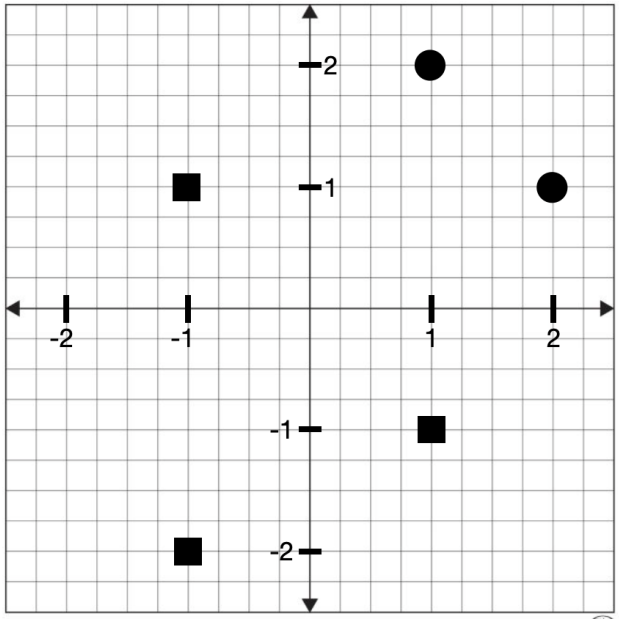


- ◯ A: E is an add operation
- ◯ B: H is equal to 2
- ◯ C: A is equal to 4
- ◯ D: A, C and D are all equal

SOLUTION:

- ◯ A: True, see that the gradient is distributed so E is an add operation
- ◯ B: True, as per Part a, E is an add operation. So F is 2. Relu does not affect this so H is also 2.
- ◯ C: False, it is 1.
- ◯ D: True, they are all 1.

2.2 Consider the below KNN classifier. We visualize all train points and their corresponding categories (circle and square). For all questions, we use the L2 distance metric.



◯ A: (2, 2) is classified as a circle for k = 3

◯ B: (0, 1) is classified as a circle for k = 3

◯ C: Adding a square to the train set at (0, 0) does not change the classifier for any value of k.

◯ D: All test points are classified as a square for $k = 5$

SOLUTION:

◯ A: True, the two circles are closest to it so 2 out of the 3 closest are circles

◯ B: True, the three points above the x axis are the closest.

◯ C: False, consider the point (0, 1) and k=3. This is classifies as a circle unless there is a square present at (0,0)

◯ D: True, then you consider the whole train set, and most of these are squares

3 Short Answers (8 points) - Recommended 8 Minutes

Please make sure to write your answer only in the provided space.

3.1 Geometric View of Linear Classifier

Consider the below linear classifier, which classifies an object into 3 different classes based on 2 different features – its temperature and its size.

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \text{temp} \\ \text{size} \end{bmatrix} = \begin{bmatrix} \text{score A} \\ \text{score B} \\ \text{score C} \end{bmatrix}$$

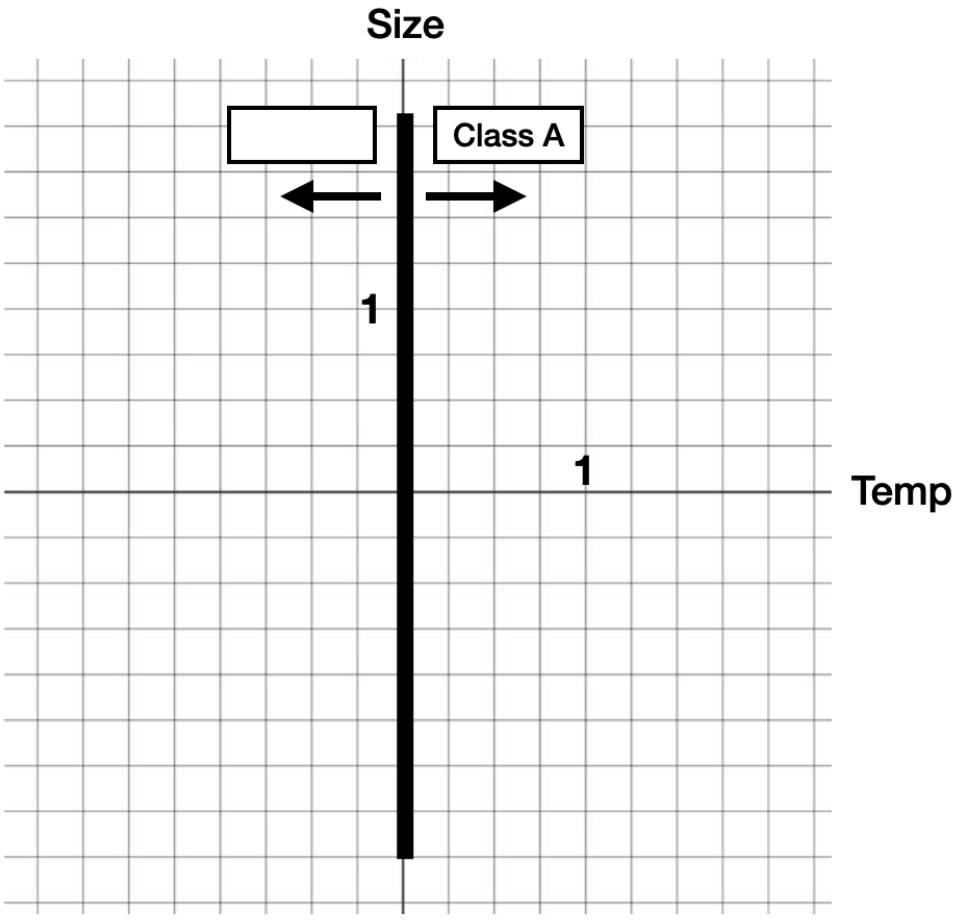
1. (1 pt) If an object has temp 1 and size -1, what does this linear classifier classify the object as? Put a star on the grid at this location and label it as the correct label (A, B, or C). The x-axis represents temp and the y-axis represents size.

2. (1 pt) Previously in lecture, we have been plotting the line which represents the score for a given class being positive or negative. This means that for 3 classes, we have 3 lines. However, we could instead plot a line which divides every pair of classes. So we choose 2 classes, and plot the line so that on one side of the line the score for one of the classes in higher and on the other side, the score for the other class is higher. We will call these binary decision boundaries. If we have 3 classes total, how many binary decision boundaries will exist?

3. (1 pt) On the grid below, we have plotted the line temp = 0. The represents the binary decision boundaries which divides Class A and another Class. Which class is it? Fill in "Class B" or "Class C" in the provided box on the grid.

4. (3 pts) Plot all the other binary decision boundaries for the above classifier. Label each side of the line with the class that is associated with that size of the line, as you did for "Class B" or "Class C" in the previous part of the question.

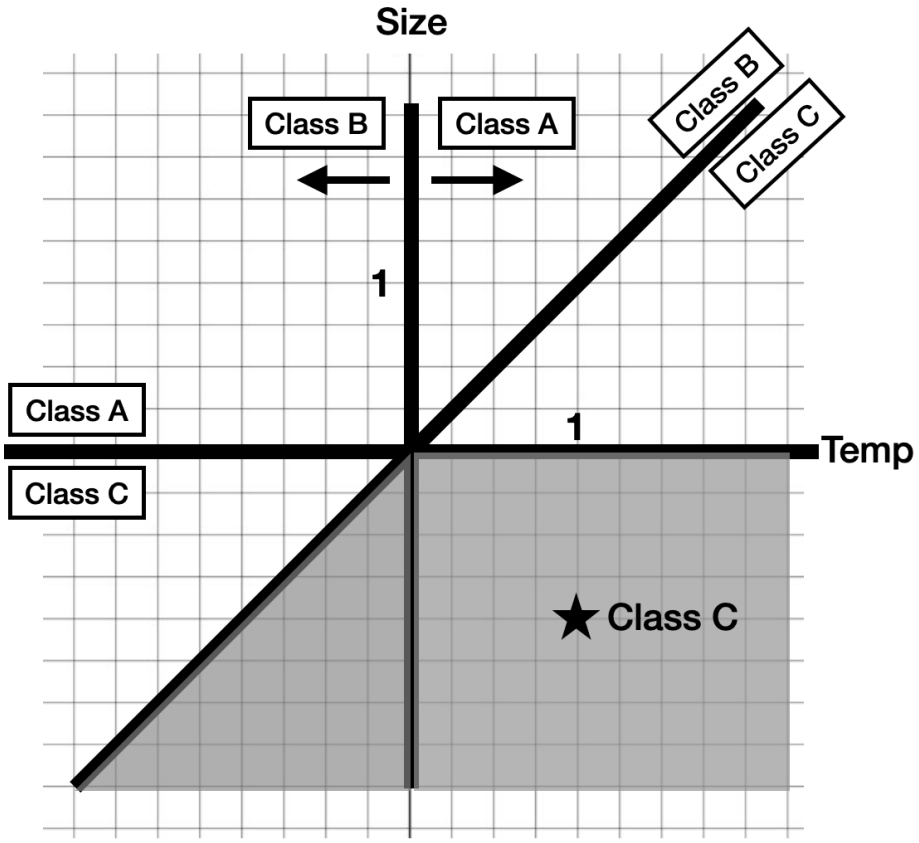
5. (2 pts) Lightly shade the region of the plot where points will be classified as Class C.



SOLUTION:

1. See image
2. We have 3 binary decision boundaries, (A and C) (A and B) (B and C)
3. See image
4. See image

5.


Size
Class B
Class A
Class B
Class C
1
Class A
1
Temp
Class C
Class C