1. Consider k-nearest neighbor classification in Fig. 1. Assume $k = 3$. What is the estimated class of point x?

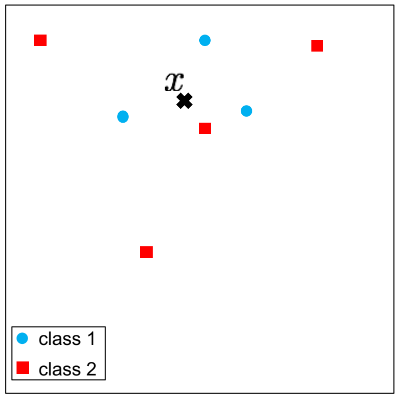

Figure 1: k-NN classification

SOL: class 1

2. Suppose we would like to classify input x given by

$$x = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

Consider parameter W is given by

$$W = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

and bias b given by

$$b = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

We would like to use linear score $s = Wx + b$, where the 1st, 2nd and 3rd element of s represents the score for "cat" , "dog" and "ship", respectively.

(a) (4 pts) Calculate the score for "cat" category

(b) (4 pts) Calculate the score for "dog" category

(c) (4 pts) Calculate the score for "ship" category

SOL:

(a) 3
(b) 6
(c) 7

3. Suppose we want to calculate Multiclass SVM loss (hinge loss) for the following score vectors. The score s_1, s_2, and s_3 for the 1st, 2nd and 3rd input data samples are given by

$$s_1 = \begin{bmatrix} 3 \\ 5 \\ -1 \end{bmatrix}, \quad s_2 = \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix}, \quad s_3 = \begin{bmatrix} 2 \\ 2 \\ -3 \end{bmatrix}$$

The 1st, 2nd and 3rd element of the score vector represent the following classes: "cat", "dog" and "ship". The ground truth labels for the 1st, 2nd and 3rd data sample are given by "cat", "dog" and "ship", respectively. Find the hinge loss averaged over the three data samples.
SOL: (3+0+12)/3=15/3=5

4. Find accuracy, precision, recall and F-score.

	Predicted class: Positive	Predicted class: Negative
actual class: Positive	85	15
actual class: Negative	890	10

SOL: accuracy: 95/1000, precision: 85/975, recall: 85/100, F-score: 2*precision*recall/(precision+recall)≈ 0.158

5. Suppose we have binary classifier which uses $\sigma(x)$ for 1-dimensional input x and sigmoid function $\sigma(\cdot)$. The decision rule is such that, given threshold s, we decide input x is positive if $\sigma(x) > s$, and negative otherwise. Suppose the following 4 data samples are given in format (x_i, y_i) such that input x_i and output y_i where $y_i = 1$ represents that x_i is positive, and $y_i = 0$ means x_i is negative:

$$(-2, 0), (-1, 1), (1, 0), (2, 1)$$

Draw the ROC curve. What is AUC?
SOL: For $s = 0$, we achieve (FP rate, TP rate)=$(1, 1)$.

For $s = \sigma(-2)$, we achieve (FP rate, TP rate)=$(0.5, 1)$.

For $s = \sigma(-1)$, we achieve (FP rate, TP rate)=$(0.5, 0.5)$.

For $s = \sigma(1)$, we achieve (FP rate, TP rate)=$(0, 0.5)$.

For $s = 1.1$, we achieve (FP rate, TP rate)=$(0, 0)$.

AUC is 3/4.

6. Suppose we have the score

$$s = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$$

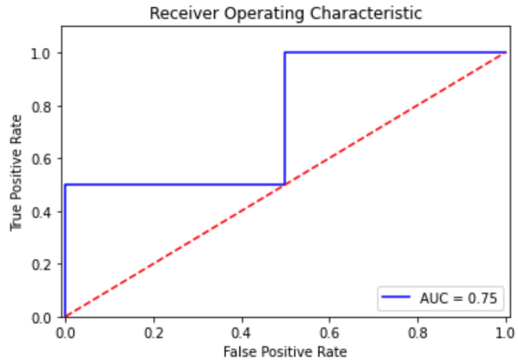

Figure 2: ROC curve

What is the softmax applied to s?
SOL:

$$\begin{bmatrix} \frac{\exp(2)}{\exp(2)+\exp(1)+\exp(0)} \\ \frac{\exp(1)}{\exp(2)+\exp(1)+\exp(0)} \\ \frac{\exp(0)}{\exp(2)+\exp(1)+\exp(0)} \end{bmatrix}$$

7. Write down the cross-entropy loss L for the linear score for class 1, 2 and 3

$$s = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$$

where the ground truth label for this data sample was class 2.
SOL:

$$L = -\log\left(\frac{\exp(1)}{\exp(2)+\exp(1)+\exp(0)}\right)$$

8. Consider the neural network in Fig. 3. The hidden layer uses ReLU activation. The output layer is a linear layer, and outputs the softmax of the linear score. Suppose the input is $(x_1, x_2) = (1, 3)$. What is the output (y_1, y_2)? Assume all the biases are 0.

SOL: The output of the first neuron of the hidden layer is

$$\max(1 \times (-2) + 3 \times 1, 0) = 1$$

The output of the second neuron of the hidden layer is

$$\max(1 \times (3) + 3 \times (-4), 0) = 0$$

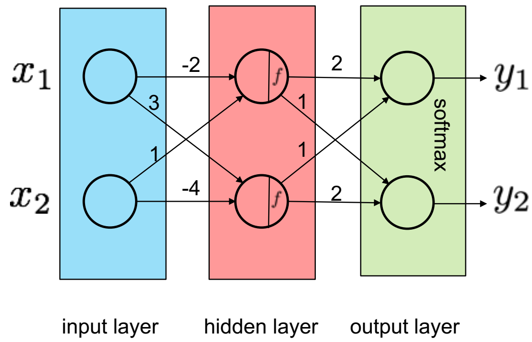

Figure 3: Neural Network

So the output of the hidden layer is $(1, 0)$. Then the output layer is

$$\text{softmax}(1 \times 2 + 0 \times 1, 1 \times 1 + 0 \times 2) = \text{softmax}(2, 1) = \left(\frac{\exp(2)}{\exp(2) + \exp(1)}, \frac{\exp(1)}{\exp(2) + \exp(1)} \right)$$

9. Consider loss function $L(x, y) = 2x + 3xy$ with learning rate α. We would like to minimize the loss using gradient descent. What is the step for gradient descent at point (x, y)?
SOL:

$$-\alpha \nabla L = -\alpha \begin{bmatrix} \frac{\partial L}{\partial x} \\ \frac{\partial L}{\partial y} \end{bmatrix} = -\alpha \begin{bmatrix} 2 + 3y \\ 3x \end{bmatrix}$$