CS6910: Tutorial 1

VERSION I

1. Given a single perceptron, can the Perceptron Learning Algorithm (PLA) find a separating boundary for the following data points: {(1, 0), (0, -1)} $\in$ P and {(0, 1) and (-1, 0)} $\in$ N. P denotes the positive class and N denotes the negative class. State True or False. If True, then write the equation of the separating boundary for each case.

   (a) False
   (b) True, $x_1 + x_2 = 1$
   (c) True, $x_1 - x_2 = 0$
   (d) True, $x_1 + x_2 = 0$

2. Consider a McCulloch Pitts (MP) neuron with the following standard notations:

$$g(x_1, x_2, x_3, ..., x_n) = g(x) = \sum_{i=1}^{n} x_i$$

$$y = f(g(x)) = \begin{cases} 1, & \text{if } g(x) \geq \theta \\ 0, & \text{if } g(x) < \theta \end{cases}$$

   Suppose, we provide 2 boolean inputs x_1 and x_2 to the model than for what value of θ can model represent the NOR function?

   (a) 1 (b) 0 (c) 2 (d) -1

3. What is the number of perceptron units that are sufficient to model the XOR function with n variables with only a single hidden layer and an output layer?

   (a) $2^{n-1} + 1$
   (b) $3^{n-1} + 1$
   (c) n + 1
   (d) $\sqrt{n} + 1$

4. What is the minimum number of perceptron units that are sufficient to model the XOR function with 6 variables with a deeper network (more than 1 hidden layer)?

   (a) 13
   (b) 16
   (c) 15
   (d) 32

5. Suppose we have some 1 dimensional input and a binary output. Suppose the input points belonging to the 2 classes are generated using two Gaussian functions: $G_1 = \mathcal{N}(2.5,\ 1.5)$ for class 1 (C1) and $G_2 = \mathcal{N}(6,\ 0.75)$ for class 2 (C2). We observe that the data is not fully linearly separable due to some points. In order to make it linearly separable, we need to discard several points. Which points should we minimally discard from the data to make it fully linearly separable?

   (a) x $\leq$ 4.6 present in C1 and x $\geq$ 4.6 present in C2.
   (b) x $\leq$ 2.6 present in C1 and x $\leq$ 2.6 present in C2.
   (c) x $\leq$ 3.6 present in C1 and x $\geq$ 3.6 present in C2.
   (d) x $\geq$ 4.6 present in C1 and x $\leq$ 4.6 present in C2.

6. Choose the statements which are True.

   (a) Perceptron Learning Algorithm (PLA) always results in a unique separating boundary between the positive and negative class points.
   (b) The separating boundary found by PLA is sensitive to the order of data points and may change if the data points are shuffled.
   (c) The separating boundary found by PLA is independent to the order of the data points and remains unique even when the data points are shuffled.
   (d) None of the above.

7. What is the minimum number of perceptron units that are sufficient to model the XOR function with n variables with a deeper network (more than 1 hidden layer)?

   (a) $2^{n-1} + 1$
   (b) 3(n-1)
   (c) 2(n-1)
   (d) log(n)

8. We want to model the problem of whether a planet is habitable or not using a McCulloch Pitts (MP) neuron. The following factors that are important for the habitability: (i) presence of a conducive atmosphere, (ii) existence of water on the surface, (iii) a suitable mass, and (iv) the right distance from its star. For simplicity, assume that these are the only factors involved and we can convert them into boolean features (isAtmosphereGood, isWaterPresent etc). Additionally, we know that the right distance of the planet from its star is 3 times more important than the other factors. Can we model this using an MP neuron? State either True or False with appropriate reasoning.

   (a) True
   (b) False

9. Consider the following set of points with the last column (t) indicating the class label:

x_1	x_2	x_3	**t**
0	0	0	0
0	1	0	1
1	0	0	1
1	1	1	0

   We are trying to model this function using a single perceptron.

   1. Are these set of points linearly separable?
   2. If yes, then find the values of w_1, w_2, w_3 and b so that the plane separates the positive/negative class points.

   (a) Yes; w_1=1, w_2=1, w_3=0, b=1
   (b) Yes; w_1=1, w_2=1, w_3=-2, b=-0.5
   (c) Yes; w_1=2, w_2=-1, w_3=0.5, b=1
   (d) No

10. Suppose the separating line learnt by a single perceptron is given by the following equation: $2x_1$ - x_2 + 1 = 0. Which of the following points belong to the positive halfspace of this line?

   (a) (0,0) (b) (-1, -1) (c) (-3, 8) (d) (3,4)

11. Suppose the data is linearly separable with two classes and you are using a single perceptron to model the relationship between the input and output. Further, suppose that you are learnint the weights of the percetron using the Pereptron Learning Algorithm (PLA). In PLA, p_i represents the i-th data point belonging to the positive class. Let $\mathbf{w}^*$ be the optimal solution such that $\forall i\ \mathbf{w}^* \cdot \mathbf{p}_i > 0$ and $\delta = min_i(\mathbf{w}^*.\mathbf{p}_i)$. What would happens if $\delta = 0$?

    (a) There is no guarantee that PLA will converge in a finite number of steps.
    (b) It is guaranteed that PLA will still converge in a finite steps.

12. What will happen if we change the update rule in the perceptron learning algorithm to $\mathbf{w} = \mathbf{w} + 2\mathbf{x}$ and $\mathbf{w} = \mathbf{w} - 2\mathbf{x}$, instead of $\mathbf{w} = \mathbf{w} + \mathbf{x}$ and $\mathbf{w} = \mathbf{w} - \mathbf{x}$ respectively?

    (a) This change will affect the correctness of the algorithm.
    (b) Correctness of the algorithm is still guaranteed.
    (c) This is similar to setting a high learning rate, hence may take long to converge.
    (d) None of the above.

13. Choose the statements which are True:

    (a) A McCulloch Pitts (MP) neuron has many parameters that can be learnt with the help of a learning algorithm.
    (b) A McCulloch Pitts (MP) neuron has only one parameter which is the firing threshold.
    (c) The Perceptron model has many parameters which can be learned using a learning algorithm.
    (d) The Perceptron model has only one parameter which is the firing threshold.
    (e) None of the above

14. Consider the following statements about the Perceptron Learning Algorithm for learning the weights of a single perceptron. Choose the ones which are True.

    (a) If a point $\mathbf{x}$ belongs to the negative class, and $\mathbf{w} \cdot \mathbf{x} < 0$, then the update is $\mathbf{w} = \mathbf{w} + \mathbf{x}$, where $\mathbf{w}$ is the vector of weights.
    (b) If a point $\mathbf{x}$ belongs to the negative class, and $w \geq 0$, then the update is $\mathbf{w} = \mathbf{w} - \mathbf{x}$, where $\mathbf{w}$ is the vector of weights.
    (c) It can work with 100% accuracy on any type of data (irrespective of whether it is linearly separable or not) if allowed to train for many epochs.

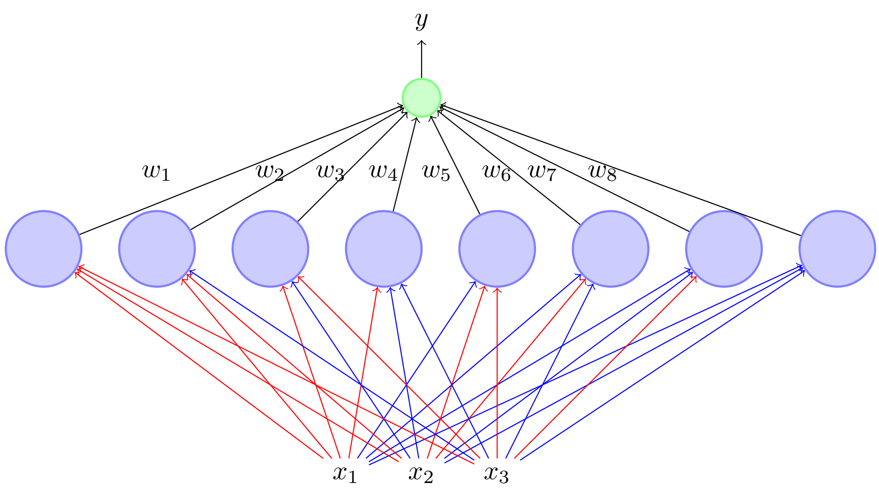
15. What is the number of perceptron units that are sufficient to model the XOR function with 6 variables with only a single hidden layer and an output layer?

   (a) 33
   (b) 32
   (c) 65
   (d) 64

16. You can think of the perceptron learning algorithm as a special case of the gradient descent algorithm with a specific loss function. Can you think of an appropriate loss function that results in the perceptron weight update equation using the framework of gradient descent (N is the total number of data points)?

   (a) $\sum_{n=1}^{N} \max(y_n \hat{y}_n \ , 0)$
   (b) $\sum_{n=1}^{N} \frac{1}{2} \ (y_n \text{ - } \hat{y}_n)^2$
   (c) $\sum_{n=1}^{N} max(-y_n \hat{y}_n \ , 0)$
   (d) $\sum_{n=1}^{N} y_n log \hat{y}_n$

17. Consider a network of perceptrons with 1 hidden and 1 output layer. Find the weights of this network to model the function $O = x_1 \oplus x_2 \oplus x_3$.
w_1 = ______, w_2 = ______, w_3 = ______, w_4 = ______, w_5 = ______, w_6 = ______, w_7 = ______, w_8 = ______



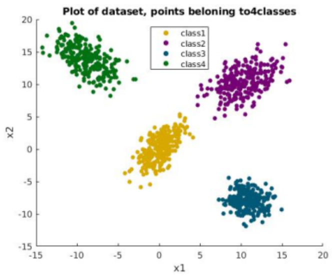
18. Given a single perceptron, can the Perceptron Learning Algorithm (PLA) find a separating boundary for the following data points: {(1, 0), (0, 1)} $\in$ P and {(-1, 0) and (0, -1)} $\in$ N. P denotes the positive class and N denotes the negative class. State True or False. If True, then write the equation of the separating boundary for each case.

    (a) True, $x_1 + x_2 = 0$
    (b) True, $x_1 - x_2 = 0$
    (c) True, $x_1 + x_2 = 1$
    (d) False

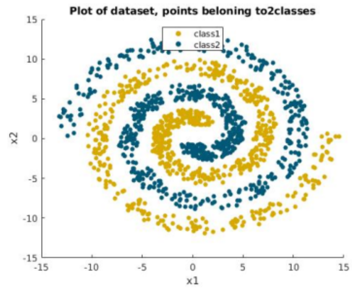
19. Given a single perceptron, can the Perceptron Learning Algorithm (PLA) find a separating boundary for the following data points: {(0, 1), (0, -1)} $\in$ P and {(1, 0) and (-1, 0)} $\in$ N. P denotes the positive class and N denotes the negative class. State True or False. If True, then write the equation of the separating boundary for each case.

    (a) True, $x_1 + x_2 = 0$
    (b) True, $x_1 - x_2 = 0$
    (c) True, $x_1 + x_2 = 1$
    (d) False

20. For which of the following types of data, can you use a multilayer network of perceptrons to fit the data well (give 100% accuracy)? Note that you are allowed to use a network of perceptrons as opposed to a single perceptron.



(a)



(b)

21. A student implemented PLA using the following train data and computed the separating boundary as: $x_1 + x_2$ - 1 = 0.

x_1	x_2	**t**
1	2	1
2	3	1
4	4.5	1
0	0	0
0.5	0	0
0.5	1	0
-1	0	0

When he applied this model on the validation samples, he noticed that the accuracy was low. He debugged the code and found that the separating boundary is not computed accurately. Your task is to find:
(i) Which point(s) are misclassified in the above samples?

(a) (1, 2, 1)
(b) (0, 0, 0)
(c) (0.5, 1, 0)
(d) (-1, 0, 0)

(ii) Suggest appropriate corrections to the parameters of the separating boundary. Write the new equation: ______________

END