Prac 2
Neural Networks and Deep Learning, 2020

August 17, 2020

1. Which of the following sets of points are linearly separable?
   (a) Class1: $p_1 = (0, 0, 0), p_2 = (1, 1, 1), p_3 = (2, 2, 2)$
   Class2: $n_1 = (3, 3, 3), n_2 = (4, 4, 4), n_3 = (5, 5, 5)$
   (b) Class1: $p_1 = (0, 0, 0), p_2 = (1, 1, 1), p_3 = (4, 4, 4)$
   Class2: $n_1 = (2, 2, 2), n_2 = (3, 3, 3)$
   (c) Class1: $p_1 = (0, 0, 0, 0), p_2 = (1, 0, 1, 0), p_3 = (0, 1, 0, 1)$
   Class2: $n_1 = (1, 1, 1, 1), n_2 = (1, 1, 1, 2), n_3 = (1, 2, 1, 1)$
   (d) Class1: $p_1 = (0, 0, 0, 0, 0), p_2 = (0, 1, 0, 1, 0), p_3 = (1, 0, 1, 0, 1)$
   Class2: $n_1 = (1, 1, 1, 1, 1), n_2 = (1, 1, 1, 1, 2), n_3 = (0, 1, 2, 1, 1)$

Solution: The general idea here is that in 2D/3D, we can identify separability visually. For n-dimensional vectors, we need to show that there exist either $\boldsymbol{w} \in \mathbb{R}^n$ such that $\boldsymbol{w}^\top \boldsymbol{x} \geq 0$ for all the samples of one class and $\boldsymbol{w}^\top \boldsymbol{x} < 0$ for the samples of the other class. If we can prove that such a hyperplane does not exist, then we have shown that the samples are not separable.

For (a), choose the hyperplane $\mathcal{H} : x[1]-3$ (here we denote a vector in $\mathbb{R}^n$ by $(x[1], x[2], \cdots, x[n])$. This hyperplane perfectly cut the two classes so sets in (a) are separable.

For (b), from the hyperplane theorem, we know that there exists a separating hyperplane if the two sets are convex and disjoint (and one should be compact but since we are dealing with finite sets, we can ignore this condition). Now consider the middle point between p_2 and p_3 which is $(2.5, 2.5, 2.5)$. Similarly, consider the middle point between n_1 and n_2 which is $(2.5, 2.5, 2.5)$. These two points collide so the two sets are not disjoint convex sets and hence the sets are not separable.
For (c), note that with $x[1]$, you can differentiate p_1, p_3 from n_1, n_2, n_3. We need to find a way to differentiate p_3 using the rest of elements. This gives us among others $\mathcal{H} : x[1] + x[4] - 2$ as a separating hyperplane.

For (d), simply note that $\mathcal{H} : x[1] + x[2] + x[3] + x[4] - 4$ separates the two sets

2. For a specific problem in $\mathbb{R}^2$, assume that we know that available data can be completely separated into two classes using an ellipsoid instead of a straight line. Suggest a modification of the perceptron algorithm that determines the equation of the appropriate circle.

Solution: An ellipsoid in $\mathbb{R}^2$ is identified by

$$f(x_1, x_2) = ax_1^2 + bx_2^2 + cx_1x_2 + dx_1 + ex_2 + f = 0 \ .$$

This suggests that a perceptron with the following feature augmentation can realize an ellipsoid $(x_1, x_2) \to (x_1^2, x_2^2, x_1x_2, x_1, x_2)$

3. we want to learn an "OR" gate for a 2 dimensional input using a perceptron. Obtain the perceptron learning rule.

Solution: Consider the OR gate for 1 2 dimensional inputs. In essence, we want to design a machine to respond to inputs according to the following table Note that in perceptron the output is either -1 or 1. This table basically tells us that we have four data points, 3 in class 1 and 1 in class -1 (negative). It is easy to see that $\boldsymbol{w} = (w_0, w_1, w_2) = (-1, 1, 1)$ is one solution (*i.e.*, $w_0 + w_1x[1] + w_2x[2]$ is less than 0 for the first point and greater or equal than 0 for the rest of points and hence is one solution for the perceptron.