CS540-1, HW3 Oct 24 Jianyi Liu

Question 1: Perceptron [5]
Using a linear threshold unit perceptron, implement the NOR function, shown below. That is, write down weights w_0, w_A, w_B such that a LTU will produce the NOR outputs given inputs A and B. (Choose the weights such that $|w_A| = |w_B| = 1$)

A	B	NOR
0	0	1
0	1	0
1	0	0
1	1	0

Solution: g(h) = 0, if h < 0;
g(h) = 1, if h ≥ 0.
So, let $w_0 = 0.5$, $w_A = -1$, $w_B = -1$(w_0 can be any number between 0 and 1.

Question 2: Probabilities [15]
(a) [7] Jack, Queen and King are called face cards. What is the conditional probability that a card drawn at random from a pack of 52 cards is a face card, given that the drawn card is a diamond.
Solution: According to the conditional probability,

$$\text{P (face card| diamond)} = \frac{\text{P(face card,diamond)}}{\text{P(diamond)}} = \frac{3/52}{13/52} = \frac{3}{13}$$

(b) [8] In a bag there are 3 blue, 4 green and 2 red marbles. What is the probability of picking up two marbles of the same color from the bag one after another without replacement?
Solution:
The probability to pick 2 blue marbles is $\text{P(blue)} = \frac{3}{9} * \frac{2}{8}$, The probability to pick 2 green marbles is $\text{P(green)} = \frac{4}{9} * \frac{3}{8}$, The probability to pick 2 red marbles is $\text{P(red)} = \frac{2}{9} * \frac{1}{8}$.
The probability to pick 2 marbles with same colors without replacement is:

$$\text{P(blue)} + \text{P(green)} + \text{P(red)} = \frac{3}{9} * \frac{2}{8} + \frac{4}{9} * \frac{3}{8} + \frac{2}{9} * \frac{1}{8} = \frac{1}{12} + \frac{1}{6} + \frac{1}{36} = \frac{5}{18}.$$

Question 3: Support Vector Machines [20]
In this question, you are provided with several points of two classes in a two-dimensional space, as showed in Fig. 1. Points of the positive class are represented as blue ● while points of the negative are represented as green ×. It's apparently a linear inseparable question. Now you are required to find a linear SVM with the help following kernel:

$$K(x, y) = 2||x||||y||$$

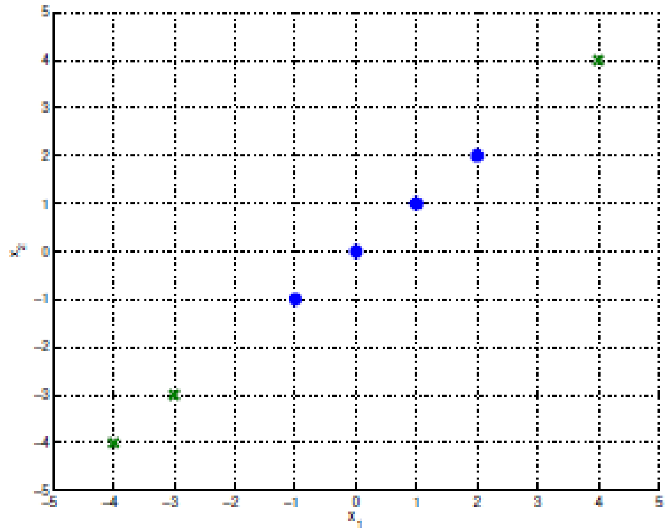


Figure 1: Points of two classes

(a) [8] Find the non-linear mapping ∅ (·) for the given kernel K that satisfies K(x, y) = ∅ (x) · ∅ (y) and then convert all original points into the new space using the mapping∅ (·). (The original coordinates of points are all integers)

Solution: $\emptyset\ (x)\ = \sqrt{2 * \sum_i x_i^2}$ or $\emptyset\ (x) = \sqrt{2}||x||$

Then, (-4, -4) → 8, (-3, -3) → 6, (-1, -1) → 2, (0, 0) → 0, (1, 1) → 2, (2, 2) → 4, (4, 4) → 8.

(b) [8] Find the linear SVM for the transformed space and identify the support vectors on the margin.

Solution: The linear SVM for the transformed one-dimension space is x = 5, and the support vectors on the margin are x = 4 for the positive class, and x = 6 for the negative class.

(c) [6] Write the function form of the corresponding decision boundary in the original space.

Solution: $x^2 + y^2 = 5^2, \mathrm{y} = \pm\sqrt{5^2 - x^2}$

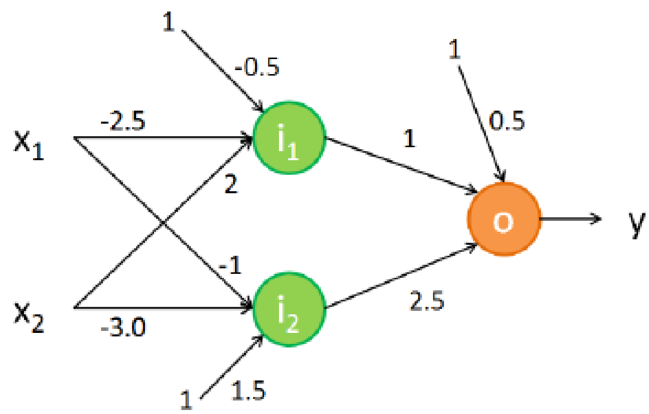
$\mathrm{y} = 5 - \sqrt{2}||x||$

Question 4: Neural Networks [20]

The following is a representation of a neural network with hidden layer nodes colored green and the output node colored orange. X_1 *and* X_2 the input variables. For the following questions, assume that the learning rate α = 0.1. Each node also has a bias input of value 1. Also, assume that there is a sigmoid activation function at the hidden layer nodes and at the output layer node. A sigmoid activation function takes the form:

$$\sigma(\mathbf{p}) = \frac{1}{1 + e^{-(\mathbf{w}^T\mathbf{p})}}$$

where p is the input vector and w is the weight vector for that particular node.



(a) [10] Calculate the output values of nodes i_1, i_2 and o of this network for the input $\{x_1 = 0, x_2 = 1\}$. Show all the step of your calculation.

Solution: $i_1 = \frac{1}{1+e^{-(-0.5+2)}} = \frac{1}{1+e^{-1.5}} \approx 0.81757$

$i_2 = \frac{1}{1+e^{-(1.5-3)}} = \frac{1}{1+e^{1.5}} \approx 0.18243$

$o = \frac{1}{1+e^{-(0.5+i_1+i_2)}} = \frac{1}{1+e^{-(0.5+\frac{1}{1+e^{-1.5}}+\frac{1}{1+e^{1.5}})}} \approx 0.85491$

(b) [10] Now you are going to compute one step of the back propagation algorithm. The input for the training instance is $\{x_1 = 0, x_2 = 1\}$.and the output of this training instance is y = 1. Please compute the updated weights for the output layer (the three incoming weights to the orange node) by performing ONE step of gradient descent. Show all steps of your calculation.

Solution: $w_1 = 1$, $w_2 = 2.5$, o = 0.85491, $\alpha = 0.1$

$w_0 \leftarrow 0.5 - 0.1\ (0.85491 - 1)\ 0.85491\ (1-0.85491)\ 1 \approx 0.50180$

$w_1 \leftarrow 1 - 0.1\ (0.85491 - 1)\ 0.85491\ (1-0.85491)\ 0.81757 \approx 1.00147$

$w_2 \leftarrow 2.5 - 0.1\ (0.85491 - 1)\ 0.85491\ (1-0.85491)\ 0.18243 \approx 2.50033$