



Q1. Avengers End Game Perceptron

Let us consider a movie example. Suppose we want to predict whether a movie freak would like to watch a movie or not. Note that each movie is represented by a vector $X = [x_1, x_2, x_3, x_4]$ and the description of each input is - x_1 : popularity(between 1 to 10), x_2 : Is Genre Superhero(boolean), x_3 : Is Director Anthony Russo(boolean), x_4 : IMDB Rating(between 0 to 1). Also the weight assigned to each of these inputs is given by $W = [w_1, w_2, w_3, w_4]$ and the bias is represented by the parameter θ . Now consider the movie Avengers ENDGAME has the feature vector $X = [8, 1, 1, 0.86]$. Now suppose we assign the following weights to the these inputs as $W = [0.14, 1, 0.9, 0.6]$ and suppose bias is 2. Based on the information will we watch the movie or not



YES

YES



NO

NO



Result

1

2

3

4

5

6

7

8

9

10



Q2. The Green Lantern Perceptron

Let us consider a movie example. Suppose we want to predict whether a movie freak would like to watch a movie or not. Note that each movie is represented by a vector $X = [x_1, x_2, x_3, x_4]$ and the description of each input is - x_1 : popularity(between 1 to 10), x_2 : Is Genre Superhero(boolean), x_3 : Is Director Anthony Russo(boolean), x_4 : IMDB Rating(between 0 to 1). Also the weight assigned to each of these inputs is given by $W = [w_1, w_2, w_3, w_4]$ and the bias is represented by the parameter θ . Now consider the movie The Green Lantern has the feature vector $X = [5, 1, 0, 0.53]$. Now suppose we assign the following weights to the these inputs as $W = [0.8, 1, 0.4, 0.8]$ and suppose bias is 7. Based on the information will we watch the movie or not



YES

YES



NO

NO



[Previous](#) | [Next](#)

Result

1

2

3

4

5

6

7

8

9

10



Q3. Convergence In Perceptron Algorithm

Index	Points $[x_0, x, y, z]$	Class
n_1	$[1, 0, 0, 0]$	Class 0
p_1	$[1, 0, 0, 1]$	Class 1
p_2	$[1, 0, 1, 0]$	Class 1
p_3	$[1, 0, 1, 1]$	Class 1
p_4	$[1, 1, 0, 0]$	Class 1
p_5	$[1, 1, 0, 1]$	Class 1
p_6	$[1, 1, 1, 0]$	Class 1
p_7	$[1, 1, 1, 1]$	Class 1

Consider the following training set:

We are interested in finding the plane which divides the input space into two classes. Starting with the weight vector $W = [0, -1, 2]$ apply the perceptron algorithm by going over the points in the following order $[n_1, p_1, p_2, p_3, p_4, p_5, p_6, p_7]$. If needed repeat the same order until convergence. After convergence what is the value of the weight vector



$[1, 1, 2, 3]$

$[1, 1, 2, 3]$



$[-1, 1, 1, 2]$

$[-1, 1, 1, 2]$



$[-3, -2, -1, -1]$

$[-3, -2, -1, -1]$



Result





n_1	$[1,0,0,0]$	Class 0
p_1	$[1,0,0,1]$	Class 1
p_2	$[1,0,1,0]$	Class 1
p_3	$[1,0,1,1]$	Class 1
p_4	$[1,1,0,0]$	Class 1
p_5	$[1,1,0,1]$	Class 1
p_6	$[1,1,1,0]$	Class 1
p_7	$[1,1,1,1]$	Class 1

Consider the following training set:

We are interested in finding the plane which divides the input space into two classes. Starting with the weight vector $W = [0, -1, 2]$ apply the perceptron algorithm by going over the points in the following order $[n_1, p_1, p_2, p_3, p_4, p_5, p_6, p_7]$. If needed repeat the same order until convergence. After convergence what is the value of the weight vector



$[1, 1, 2, 3]$

$[1, 1, 2, 3]$



$[-1, 1, 1, 2]$

$[-1, 1, 1, 2]$



$[-3, -2, -1, -1]$

$[-3, -2, -1, -1]$



$[-2, -1, -1, 1]$

$[-2, -1, -1, 1]$



Result

1

2

3

4

5

6

7

8

9

10



Q4. Partial Derivative

Let $f(x)$ be the logistic function where $f(x) = 1/(1+e^{-(Wx+B)})$. The loss function L is taken as $L = (y-f(x))^2/2$. Here x, y are re constants and W, B are parameters that will be updated. In other words L is a function of W and B . Let the partial derivative of L with respect to W be 'M' and with respect to B be 'N'. Calculate the value of 'm' and 'n'.



$$N = (y-f(x))f(x)(1-f(x))x, M = (y-f(x))(1-f(x))x$$

$$N = (y-f(x))f(x)(1-f(x))x, M = (y-f(x))(1-f(x))x$$



$$N = -(y-f(x))f(x)(1-f(x)), M = (y-f(x))f(x)(1-f(x))x$$

$$N = -(y-f(x))f(x)(1-f(x)), M = (y-f(x))f(x)(1-f(x))x$$



$$N = (y-f(x))f(x)(1-f(x)), M = (y-f(x))f(x)(1-f(x))x$$

$$N = (y-f(x))f(x)(1-f(x)), M = (y-f(x))f(x)(1-f(x))x$$



Result

1

2

3

4

5

6

7

8

9

10



Q5. Boolean Perceptron

Consider two perceptrons one denoting boolean AND and one denoting boolean OR. What are two ways to go from an AND perceptron to an OR perceptron?

- ☐ Increase the magnitude of the bias
Increase the magnitude of the bias
- ☐ Decrease the value of weights
Decrease the value of weights
- ☐ Increase the value of weights
Increase the value of weights
- ☒ Decrease the magnitude of the bias
Decrease the magnitude of the bias



Previous | Next

Result

1	2	3
4	5	6
7	8	9
10		



Q6. Scaling 1

Let the learned weights after the perceptron algorithm finishes training be weight vector W . Suppose that the bias term is 0. If we scale W by a positive constant factor (multiply each element of W with c), then the new set of weights will



produces the exact same classification for all the data points

produces the exact same classification for all the data points



may output different classification results for some data points

may output different classification results for some data points



Cannot be decided

Cannot be decided



Result

1

2

3

4

5

6

7

8

9

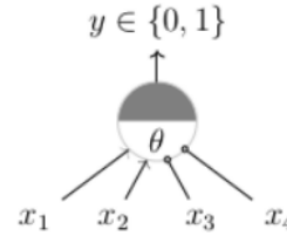
10

[Previous](#) | [Next](#)



Q7. Boolean Perceptron 2

Let us assume neuron which aggregates the input and makes a decision on the basis of this aggregation. If sum of all inputs is greater than the given threshold θ then output of this neuron is 1 else it is 0. We can implement a boolean function with this neuron if output of this neuron is consistent with the truth table of the boolean function. In other words if the given input configuration for the boolean function outputs 1 then neuron should also give 1 to the same configuration.



Consider the following boolean function $(x_1 \text{ And } x_2) \text{ AND } (!x_3 \text{ And } !x_4)$ The neuron is depicted in the following figure . What should be the value of the threshold θ such that neuron implements this boolean function. (Note that circle at the end of x_3 and x_4 input in the diagram depicts it is an inhibitory input i.e. if this input is 1 then the output will be 0 for sure irrespective of other inputs)



1

1



2

2



3

3



4

4



Result

1

2

3

4

5

6

7

8

9

10



Q8. Dimension Modification 1

We have a training set with 4 data points, as follows : $X = (0, 0) \Rightarrow Y = 1$: $X = (0, 1) \Rightarrow Y = -1$: $X = (1, 0) \Rightarrow Y = -1$: $X = (1, 1) \Rightarrow Y = 1$. Notice that the data above is not linearly separable, hence the perceptron algorithm will not be able to learn a classifier that gives the correct prediction for all four above data points. Add a 3rd dimension to each of the extra input dimension so that the data becomes linearly separable:



Third value is equal to first value for each data

Third value is equal to first value for each data



Third value is 1 for one data point, and 0 for other three

Third value is 1 for one data point, and 0 for other three



Third value is opposite of second value for each data point

Third value is opposite of second value for each data point



None of the above

None of the above

Result

1	2	3
4	5	6
7	8	9
10		



Q9. Dimension Modification 2

We have a training set with 4 data points, as follows : $X = (0, 0) \Rightarrow Y = 1$: $X = (0, 1) \Rightarrow Y = -1$: $X = (1, 0) \Rightarrow Y = -1$: $X = (1, 1) \Rightarrow Y = 1$. Notice that the data above is not linearly separable, hence the perceptron algorithm will not be able to learn a classifier that gives the correct prediction for all four above data points. Add a 3rd dimension to each of the extra input dimension so that the data becomes linearly separable:



Third value is equal to target value for each data point

Third value is equal to target value for each data point



Third value is opposite of target value for each data point

Third value is opposite of target value for each data point



Both of the above

Both of the above

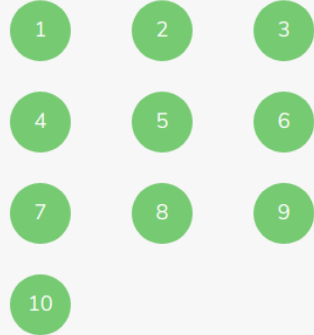


None of the above

None of the above



Result





Q10. Scaling 2

Let the learned weights after the perceptron algorithm finishes training be weight vector W . Suppose that the bias term is 0. If we translate W by a positive constant factor c (add c to each element of W), then the new set of weights will



produces the exact same classification for all the data points

produces the exact same classification for all the data points



may output different classification results for some data points

may output different classification results for some data points



Cannot be decided

Cannot be decided

[Previous](#) | [Next](#)

Result

1

2

3

4

5

6

7

8

9

10