ICTS 4156: Introduction to ML

Razvan Bunescu

Lecture notes, March 10, 2021

1 Notes on lecture slides material

1.1 The Perceptron

We have the logical OR training dataset:

1. $x^{(1)} = [1, 0, 0]$, with label $t_1 = -1$
2. $x^{(2)} = [1, 0, 1]$, with label $t_2 = +1$
3. $x^{(3)} = [1, 1, 0]$, with label $t_3 = +1$
4. $x^{(4)} = [1, 1, 1]$, with label $t_4 = +1$

Want to train a weight vector $w = [w_0, w_1, w_2]$ such that $w^T x \geq 0$ if and only if $t(x) = +1$.

Let's run the Perceptron algorithm: Initialize $\mathbf{w} = [\mathbf{0}, \mathbf{0}, \mathbf{0}]$.

1. Epoch 1:
   - For example $x^{(1)}$, prediction is $h_1 = w^T x^{(1)} = 0$. This means $h_1 t_1 = 0 \leq 0$ so we made a mistake (line 4). Change the weight vector $w = w + t_1 x^{(1)} = w - x^{(1)} = [-1, 0, 0]$. So, $\mathbf{w} = [-\mathbf{1}, \mathbf{0}, \mathbf{0}]$.
   - For example $x^{(2)}$, prediction is $h_2 = w^T x^{(2)} = -1$. This means $h_2 t_2 = -1 \leq 0$ so we made a mistake (line 4). Change the weight vector $w = w + t_2 x^{(2)} = w + x^{(2)} = [-1, 0, 0] + [1, 0, 1]$. So, $\mathbf{w} = [\mathbf{0}, \mathbf{0}, \mathbf{1}]$.
   - For example $x^{(3)}$, prediction is $h_3 = w^T x^{(3)} = 0$. This means $h_3 t_3 = 0 \leq 0$ so we made a mistake (line 4). Change the weight vector $w = w + t_3 x^{(3)} = w + x^{(3)} = [0, 0, 1] + [1, 1, 0]$. So, $\mathbf{w} = [\mathbf{1}, \mathbf{1}, \mathbf{1}]$.
   - For example $x^{(4)}$, prediction is $h_4 = w^T x^{(4)} = 3$. This means $h_4 t_4 = 3 > 0$ so no mistake. The weight vector stays unchanged, i.e. $\mathbf{w} = [\mathbf{1}, \mathbf{1}, \mathbf{1}]$.
2. Epoch 2:
   - For example $x^{(1)}$, prediction is $h_1 = w^T x^{(1)} = 1$. This means $h_1 t_1 = 1 \leq 0$ so we made a mistake (line 4). Change the weight vector $w = w + t_1 x^{(1)} = w - x^{(1)} = [1, 1, 1] - [1, 0, 0]$. So, $\mathbf{w} = [\mathbf{0}, \mathbf{1}, \mathbf{1}]$.
   - For examples $x^{(2)}, x^{(3)}, x^{(4)}$, w is good, no mistake.
3. Epoch 3:

- For example $x^{(1)}$, prediction is $h_1 = w^T x^{(1)} = 0$. This means $h_1 t_1 = 0 \leq 0$ so we made a mistake (line 4). Change the weight vector $w = w + t_1 x^{(1)} = w - x^{(1)} = [0, 1, 1] - [1, 0, 0]$. So, $\mathbf{w} = [\mathbf{-1}, \mathbf{1}, \mathbf{1}]$.
- For examples $x^{(2)}, x^{(3)}, x^{(4)}$, w is good, no mistake.

4. Epoch 4:

   - $\mathbf{w} = [\mathbf{-1}, \mathbf{2}, \mathbf{2}]$. ?