Assignment7 – Learning

Given: June 9 Due: July 18

Problem 7.1 (Weight Updates) 0 pt

Our hypothesis space contains the functions $h_{\mathbf{w}}(\mathbf{x}) = F(\mathbf{w} \cdot \mathbf{x})$ for 2+1-dimensional vectors $\mathbf{w}, \mathbf{x}$ (using the trick $\mathbf{x}_0 = 1$ to allow for the constant term $\mathbf{w}_0$) and some fixed function F. Our examples are the set

Example number	$\mathbf{x}_1$	$\mathbf{x}_2$	y
1	2	0	2
2	3	1	2

As the initial weights, we use $\mathbf{w}_0 = \mathbf{w}_1 = \mathbf{w}_2 = 0$.

For each of the following cases, iterate the weight update rule once for each example (using the examples in the order listed). Use learning rate $\alpha = 1$.

1. Using the threshold function $F(z) = \mathcal{T}(z)$, i.e., $F(z) = 1$ if $z > 0$ and $F(z) = 0$ otherwise.
2. Using the logistic function $F(z) = 1/(1 + e^{-x})$.

Solution:

1. The update rule is $\mathbf{w}_i \leftarrow \mathbf{w}_i + \alpha(y - h_{\mathbf{w}}(\mathbf{x}))\mathbf{x}_i$. Using the examples, we obtain:
   - Example 1: $y - h_{\mathbf{w}}(\mathbf{x}) = 2 - \mathcal{T}((0,0,0) \cdot (1,2,0)) = 2$, i.e., $\mathbf{w}_i \leftarrow \mathbf{w}_i + 1\mathbf{x}_i$. Thus, $\mathbf{w} \leftarrow (2,4,0)$.
   - Example 2: $y - h_{\mathbf{w}}(\mathbf{x}) = 2 - \mathcal{T}((2,4,0) \cdot (1,3,1)) = 1$, i.e., $\mathbf{w}_i \leftarrow \mathbf{w}_i + 1\mathbf{x}_i$. Thus, $\mathbf{w} \leftarrow (3,7,1)$.
2. The update rule is $\mathbf{w}_i \leftarrow \mathbf{w}_i + \alpha(y - h_{\mathbf{w}}(\mathbf{x}))h_{\mathbf{w}}(\mathbf{x})(1 - h_{\mathbf{w}}(\mathbf{x}))\mathbf{x}_i$. Using the examples, we obtain:
   - Example 1: $h_{\mathbf{w}}(\mathbf{x}) = 1/(1 + e^0) = 1/2$, i.e., $\mathbf{w}_i \leftarrow \mathbf{w}_i + 1/4(2 - 1/2)\mathbf{x}_i$. Thus, $\mathbf{w} \leftarrow (3/8, 3/4, 0)$.
   - Example 2: $h_{\mathbf{w}}(\mathbf{x}) = 1/(1 + e^{-((3/8,3/4,0)\cdot(1,3,1))}) = 1/(1 + e^{-21/8})$, i.e., (after rounding) $\mathbf{w}_i \leftarrow \mathbf{w}_i + 0.07\mathbf{x}_i$. Thus, $\mathbf{w} \leftarrow (0.44, 0.95, 0.07)$.

Problem 7.2 (Decision Tree Learning in Python) 40 pt

Implement the *Decision Tree Learning* algorithm (`DTL`) in Python using the files at `https://kwarc.info/teaching/AI/resources/AI2/dtl`.

Solution:

Problem 7.3 (Decision List) 30 pt

We want to construct a decision list to classify the data below where result values V depend on 4 attributes A, B, C, D. The tests should be conjunctions of literals.

1. Assume your literals must be of the form $attribute = number$. For which k-can we give the shortest possible decision list in k-DL (i.e., using at most k literals per test)? Give the list.

2. Now assume your literals may also be of the form $attribute = attribute$. Answer the same question as above.

Example	A	B	C	D	V
#1	1	0	0	0	1
#2	1	0	1	1	1
#3	0	1	0	0	1
#4	1	1	0	1	1
#5	0	0	1	1	1
#6	0	1	1	0	0
#7	0	1	0	1	0
#8	0	0	1	0	0

Solution:

1. For $k = 3$, we can build a list of length 2:
   if $A = 0 \wedge C = 1 \wedge D = 0$ then 0 else
   if $A = 0 \wedge C = 0 \wedge D = 1$ then 0 else
   1.

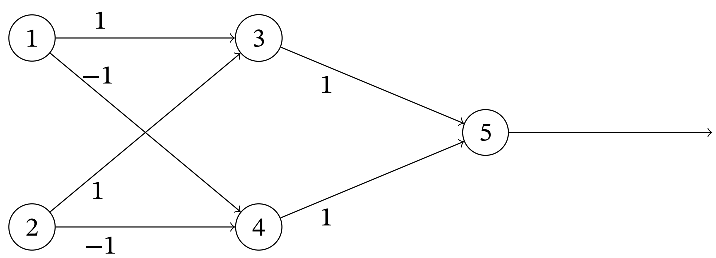
   We also accepted the following solutions for $k = 2$ with a list of length 3:
   if $A = 1$ then 1 else
   if $C = 0 \wedge D = 0$ then 1 else
   if $C = 1 \wedge D = 1$ then 1 else
   0

2. For $k = 1$, we can build a list of length 2:
   if $A = 1$ then 1 else
   if $C = D$ then 1 else
   0

Problem 7.4 (XOR Neural Network) 30 pt

Consider the following neural network with

- inputs a_1 and a_2
- units 3, 4, 5 with activation functions such that $a_i \leftarrow \begin{cases} 1 & \text{if } \Sigma_j w_{ji} a_j > b_i \\ 0 & \text{otherwise} \end{cases}$
- weights w_{ij} as given by the labels on the edges



1. Assume $b_3 = b_4 = b_5 = 0$ and inputs $a_1 = a_2 = 1$. What are the resulting activations a_3, a_4, and a_5?
2. Choose appropriate values for b_3, b_4, and b_5 such that the network implements the XOR function.

Solution:

1. $a_3 = 1$, $a_4 = 0$, $a_5 = 1$
2. E.g., $b_3 = 0.5$, $b_4 = -1.5$, $b_5 = 1.5$. More generally, any values work that satisfy $0 \leq b3 < 1$, $-2 \leq b4 < -1$, and $1 \leq b5 < 2$.
