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 $F(z) = 1/(1 + e^{-x})$.

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`.

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.

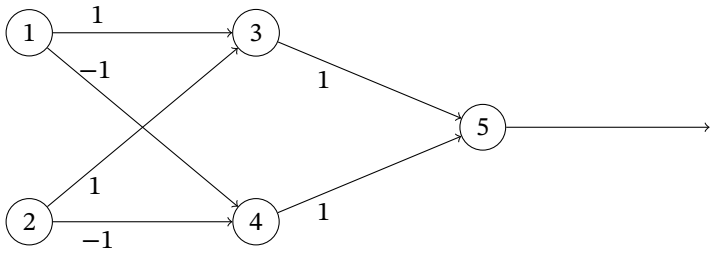
1. Assume your literals must be of the form $attribute = number$. For which k-can we give the shortest possible decision list in k-DL (i.e., using at most k literals per test)? Give the list.
2. Now assume your literals may also be of the form $attribute = attribute$. Answer the same question as above.

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

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.