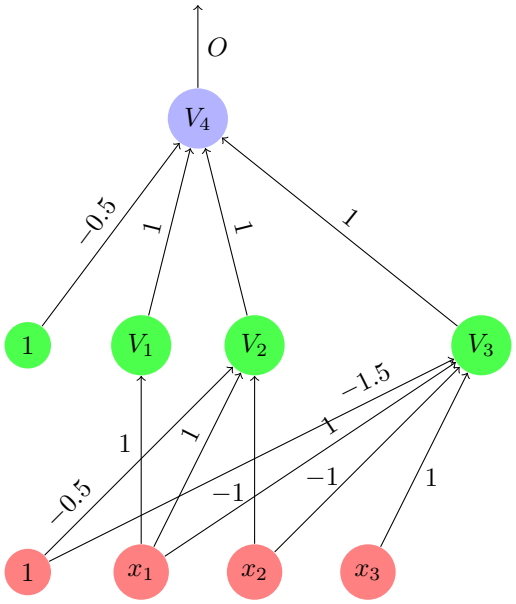
Homework-3 Solutions

Question 1

Give a drawing with weights of a neural net with a single hidden layer that computes the following Boolean formula:

$$f(x_1, x_2, x_3) = x_1 \vee (x_1 \wedge \bar{x}_2) \vee (x_1 \wedge \bar{x}_2 \wedge x_3)$$

Assume that each node computes a threshold function.



Question 2

Part 1

$$e_A = 0.5$$

$$e_B = 0.5$$

Part 2

What are the computed errors in each case:

Permutations 1: $e_A = 5/6$ $e_B = 5/6$
Permutations 2: $e_A = 4/6$ $e_B = 0.5$
Permutations 3: $e_A = 5/6$ $e_B = 5/6$

What are the estimates to the error and standard deviation for each algorithm:

Algorithm A: $e = 0.778$ $\sigma = 0.096$
Algorithm B: $e = 0.722$ $\sigma = 0.192$

Question 3

Apply the linear discriminant algorithm to the following training data:

x	y
0 -1	P
0 0	N
0 1	P
0 2	N
0 3	N
1 0	P
2 0	P
2 1	P

1. Compute the matrix X and the vector y.

$$X = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 2 \\ 1 & 0 & 3 \\ 1 & 1 & 0 \\ 1 & 2 & 0 \\ 1 & 2 & 1 \end{pmatrix}, \quad y = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \\ -1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

2. Compute the matrix $B = X^T X$ and the vector $h = X^T y$.

$$B = \begin{pmatrix} 8 & 5 & 6 \\ 5 & 9 & 2 \\ 6 & 2 & 16 \end{pmatrix}, \quad h = \begin{pmatrix} 2 \\ 5 \\ -4 \end{pmatrix}$$

3. Compute the coefficients vector a, the solution of the linear system $Ba = h$.

$$a = \begin{pmatrix} 0.224806 \\ 0.51938 \\ -0.399225 \end{pmatrix}$$

4. For each training example (x_i, y_i) compute $\tilde{y}_i = a^T x_i$, the estimate of y_i. (Observe that the product Xa gives the desired esimates in a vector form.)

$$\tilde{y} = \begin{pmatrix} 0.624031 \\ 0.224806 \\ -0.174419 \\ -0.573643 \\ -0.972868 \\ 0.744186 \\ 1.26357 \\ 0.864341 \end{pmatrix}$$

5. How many of the training examples are wrongly classified if the classification rule is:

Classify x_i as P if $\tilde{y}_i > 0$. Otherwise classify x_i as N.

Two examples are wrongly classified.

6. Determine a threshold value t to minimize the number of errors on the training data of the following rule:

Classify x_i as P if $\tilde{y}_i > t$. Otherwise classify x_i as N.

Any value of t satisfying $-0.573643 \geq t < -0.174419$. There are reasons to prefer the middle value: $t = (-0.573643 + (-0.174419)/2 = -0.374031$.