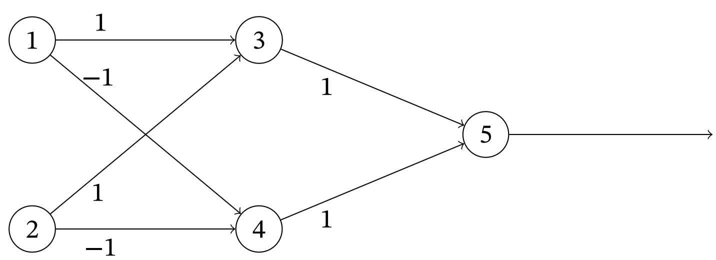
Assignment9 – Learning

Given: July 3 Due: July 8

Problem 9.1 (XOR Neural Network)

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?

   Solution: $a_3 = 1$, $a_4 = 0$, $a_5 = 1$

   ---

2. Choose appropriate values for b_3, b_4, and b_5 such that the network *implements* the XOR function.

   Solution: 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 $1 \leq b5 < 2$.

Problem 9.2 (Neural Networks in Python)

Implement neural networkss in Python by completing the implementation of `network.py` at `https://kwarc.info/teaching/AI/resources/AI2/network/`.

Hint: You can test your *implementation* with `test.py`. Note that `test_train_xor_gate` may occasionally fail for a correct solution because it is randomized.

Solution: See `https://kwarc.info/teaching/AI/resources/AI2/network/`.

Problem 9.3 (Statistical Learning)

We use two observations to determine if it has rained on our property: whether the ground is wet, and whether a bucket we left outside is full.

1. Model this situation as a naive Bayesian network with a boolean class and two boolean attributes.

   Solution: We use three boolean variables R (rain), G (ground wet), and B (bucket full) with edges $R \to G$ and $R \to B$.

2. Explain why this network requires 5 parameters ($2n + 1$ where $n = 2$ is the number of attributes). Choose 5 names for the parameters and use them to give the conditional probability table of the network.

   Solution: We need

   - one parameter θ for the probability of the class variable R: $P(R = true) = \theta$ and $P(R = false) = 1 - \theta$
   - for each attribute, one parameter for each possible value of the class variable, i.e., 2 if the class variable is boolean:
     - θ_{G1} and θ_{G2} for attribute G: $P(G = true \mid R = true) = \theta_{G1}$ and $P(G = false \mid R = true) = 1 - \theta_{G1}$ as well as $P(G = true \mid R = false) = \theta_{G2}$ and $P(G = false \mid R = false) = 1 - \theta_{G2}$
     - θ_{B1} and θ_{B2} for attribute B: $P(B = true \mid R = true) = \theta_{B1}$ and $P(B = false \mid R = true) = 1 - \theta_{B1}$ as well as $P(B = true \mid B = false) = \theta_{B2}$ and $P(B = false \mid R = false) = 1 - \theta_{B2}$

3. Now assume we have observed for 50 days with the following results:

rain	ground wet	bucket full	number of days
yes	yes	yes	10
yes	yes	no	5
yes	no	yes	6
yes	no	no	4
no	yes	yes	2
no	yes	no	9
no	no	yes	3
no	no	no	11

   State the formula for the likelihood of this list of 50 observations in terms of the 5 parameters.

Solution: Let $\mathbf{d}$ be the list of observations. The likelihood is

$$P(\mathbf{d} \mid h_{\theta,\theta_{G1},\theta_{G2},\theta_{B1},\theta_{B2}}) =$$

$$\theta^{10+5+6+4}(1-\theta)^{2+9+3+11}\theta_{G1}^{10+5}(1-\theta_{G1})^{6+4}\theta_{G2}^{2+9}(1-\theta_{G2})^{3+11}\theta_{B1}^{10+6}(1-\theta_{B1})^{5+4}\theta_{B2}^{2+3}(1-\theta_{B2})^{9+11}$$

---

4. Give the Maximum Likelihood approximations for the 5 parameters given these 50 observations. (You just need to compute them, not derive the formula for computing them.)

Solution: We can derive the formulas by taking logarithms and maximizing. The resulting formulas are the fraction of positive over total examples. So we get

$$\theta = (10+5+6+4)/(10+5+6+4+2+9+3+11)$$

$$\theta_{G1} = (10+5)/(10+5+6+4) \qquad \theta_{G2} = (2+9)/(2+9+3+11)$$

$$\theta_{B1} = (10+6)/(10+5+6+4) \qquad \theta_{B2} = (2+3)/(2+9+3+11)$$

Problem 9.4 (Statistical Learning)

You observe the values below for 20 games of a sports team. You want to predict the result based on weather and opponent.

		Number of	
Weather	Opponent	wins	losses
Rainy	Weak	3	1
Cloudy	Weak	0	1
Sunny	Weak	4	2
Rainy	Strong	0	2
Cloudy	Strong	2	3
Sunny	Strong	0	2

1. What is the hypothesis space for this situation, seen as an *inductive learning problem*?

Solution: The set of functions $\{Rainy, Cloudy, Sunny\} \times \{Weak, Strong\} \to \{Win, Loss\}$.

---

2. Explain whether we can learn the function by building a decision tree.

Solution: We cannot. Even all attributes together, i.e., $Weather$ and $Opponent$, do not determine the result. So no decision tree exists.

---

3. To apply Bayesian learning, we model this situation as a Bayesian network $W \rightarrow R \leftarrow O$ using random variables W (weather), O (opponent), and R (game result). What are the resulting entries of the conditional probability table for the cases
   1. $P(W = rainy) =$ 3/10
   2. $P(R = win \mid O = weak) =$ 7/11

Solution: $P(W = rainy) = 3/10$ and $P(R = win \mid O = weak) = 7/11$
