COMP 6721 - Artificial Intelligence

Neural Networks

Solutions

Question 1 Assume that your local bank is trying to determine if a client should be considered to be a low-risk borrower or a high-risk borrower. Jim, the bank director, compiled a few data from previous experience based on the following criteria:

- The client has a mortgage or not.
- The client has an income inferior to 40,000$ or not.
- The client has a university degree or not.

Client	Has a mortgage?	Income ($< 40,000\$$?)	Has a University Degree?	Type of client?
A	No	Yes	Yes	low-risk
B	Yes	Yes	Yes	low-risk
C	No	Yes	Yes	low-risk
D	No	Yes	No	high-risk
E	Yes	No	No	high-risk

Assume that the weights of a perceptron are initialized this way:

- -0.6 for the Mortgage feature
- -0.2 for the Income feature
- +0.3 for the University degree feature

Show how the weights will be modified after each observation (each client) has been taken into account. Do only one iteration over the training data (1 epoch). Assume that a sign function is used and that all weights are always adjusted by a constant value of 0.1. Show all your work. *let's assume that if* $f(net) \geq 0$, *we classify as low-risk; and otherwise, we classify as high-risk. Note that the opposite could have been assumed but the results would be different.*

for A: $f(0 \times -.6 + 1 \times -.2 + 1 \times .3) = f(+.1) \geq 0$
So we conclude low-risk, which is correct. We don't change the weights at this point.

for B: $f(1 \times -.6 + 1 \times -.2 + 1 \times .3) = f(-.5) < 0$
So we conclude high-risk, which is incorrect. $f(net)$ is too low, we need to increase the weights of the active inputs (they are all active here).

$$w1 = -.6 + .1 = -.5$$
$$w2 = -.2 + .1 = -.1$$
$$w3 = +.3 + .1 = +.4$$

for C: $f(0 \times -.5 + 1 \times -.1 + 1 \times .4) = f(.5) \geq 0$
So we conclude low-risk, which is correct. We don't change the weights at this point.

for D: $f(0 \times -.5 + 1 \times -.1 + 0 \times .4) = f(-.1) < 0$
So we conclude high-risk, which is correct. We don't change the weights at this point.

for E: $f(1 \times -.5 + 0 \times -.1 + 0 \times .4) = f(-.5) < 0$
So we conclude high-risk, which is correct. We don't change the weights at this point.

The final weights are:

$$w1 = -.5$$
$$w2 = -.1$$
$$w3 = +.4$$

Question 2 Can a perceptron learn the SAME function of three binary inputs, defined to be 1 if all inputs are the same value and 0 otherwise? Either argue/show that this is impossible or construct a perceptron that correctly represents this function.

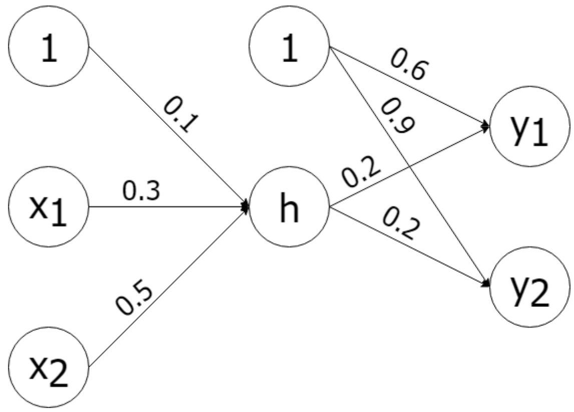
No. SAME is the complement of XOR. It is not linearly separable and therefore cannot be represented by a perceptron. Proof is by a figure showing the 3D space.

Question 3 Can a perceptron learn to correctly classify the following data, where each consists of three binary input values and a binary classification value: (111,1), (110,1), (011,1), (010,0), (000,0)? Either show that this is impossible or construct such a perceptron.

Yes. Output is 1 if at least 2 of the 3 inputs are 1. Therefore a perceptron with all three weights equal to 0.5 and a threshold value of 0.8 will work.

$$0.5i_1 + 0.5i_2 + 0.5i_3 \geq 0.8 \textit{ then output } = 1$$
$$0.5i_1 + 0.5i_2 + 0.5i_3 < 0.8 \textit{ then output } = 0$$

Question 4 Consider the neural network shown below. It consists of 2 input nodes, 1 hidden node, and 2 output nodes, with addition to a bias at the input layer and a bias at the hidden layer. All nodes in the hidden and output layers use the sigmoid activation function (σ).



(a) Calculate the output of y1 and y2 if the network is fed $x = (1, 0)$ as input.

$$h_{in} = b_h + w_{x_1\text{-}h}x_1 + w_{x_2\text{-}h}x_2 = (0.1) + (0.3 \times 1) + (0.5 \times 0) = 0.4$$

$$h = \sigma(h_{in}) = \sigma(0.4) = \frac{1}{1 + e^{-0.4}} = 0.599$$

$$y_{1,in} = b_{y_1} + w_{h\text{-}y_1}h = 0.6 + (0.2 \times 0.599) = 0.72$$

$$y_1 = \sigma(0.72) = \frac{1}{1 + e^{-0.72}} = 0.673$$

$$y_{2,in} = b_{y_2} + w_{h\text{-}y_2}h = 0.9 + (0.2 \times 0.599) = 1.02$$

$$y_2 = \sigma(1.22) = \frac{1}{1 + e^{-1.02}} = 0.735$$

As a result, the output is calculated as $y = (y1, y2) = (0.673, 0.735)$.

(b) Assume that the expected output for the input $x = (1, 0)$ is supposed to be $t = (0, 1)$, calculate the updated weights after the backpropagation of the error for this sample. Assume that the learning rate $\eta = 0.1$.

$$\delta_{y_1} = y_1(1 - y_1)(y_1 - t_1) = 0.673(1 - 0.673)(0.673 - 0) = 0.148$$
$$\delta_{y_2} = y_2(1 - y_2)(y_2 - t_2) = 0.735(1 - 0.735)(0.735 - 1) = -0.052$$
$$\delta_h = h(1-h) \sum_{i=1,2} w_{h\text{-}y_i} \delta_{y_i} = 0.599(1-0.599)[0.2\times 0.148+0.2\times(-0.052)] = 0.005$$

$$\Delta w_{x_1\text{-}h} = -\eta\delta_h x_1 = -0.1 \times 0.005 \times 1 = -0.0005$$
$$\Delta w_{x_2\text{-}h} = -\eta\delta_h x_2 = -0.1 \times 0.005 \times 0 = 0$$
$$\Delta b_h = -\eta\delta_h = -0.1 \times 0.005 = -0.0005$$

$$\Delta w_{h\text{-}y_1} = -\eta\delta_{y_1} h = -0.1 \times 0.148 \times 0.599 = -0.0088652$$
$$\Delta b_{y_1} = -\eta\delta_{y_1} = -0.1 \times 0.148 = -0.0148$$

$$\Delta w_{h\text{-}y_2} = -\eta\delta_{y_2} h = -0.1 \times (-0.052) \times 0.599 = 0.0031148$$
$$\Delta b_{y_2} = -\eta\delta_{y_2} = -0.1 \times (-0.052) = 0.0052$$

$$w_{x_1\text{-}h,new} = w_{x_1\text{-}h} + \Delta w_{x_1\text{-}h} = 0.3 + (-0.0005) = 0.2995$$
$$w_{x_2\text{-}h,new} = w_{x_2\text{-}h} + \Delta w_{x_2\text{-}h} = 0.5 + 0 = 0.5$$
$$b_{h,new} = b_h + \Delta b_h = 0.1 + (-0.0005) = 0.0995$$

$$w_{h\text{-}y_1,new} = w_{h\text{-}y_1} + \Delta w_{h\text{-}y_1} = 0.2 + (-0.0088652) = 0.1911348$$
$$b_{y_1,new} = b_{y_1} + \Delta b_{y_1} = 0.6 + (-0.0148) = 0.5852$$

$$w_{h\text{-}y_2,new} = w_{h\text{-}y_2} + \Delta w_{h\text{-}y_2} = 0.2 + 0.0031148 = 0.2031148$$
$$b_{y_2,new} = b_{y_2} + \Delta b_{y_2} = 0.9 + 0.0052 = 0.9052$$

Question 5 Recalculate your answers of the previous question, using matrix notation.

First, we need to visualize our parameters in matrix format.

$$x = \begin{bmatrix} 1 & 0 \end{bmatrix} \qquad t = \begin{bmatrix} 0 & 1 \end{bmatrix}$$

$$w_{x\text{-}h} = \begin{bmatrix} 0.3 \\ 0.5 \end{bmatrix} \qquad b_h = \begin{bmatrix} 0.1 \end{bmatrix} \qquad w_{h\text{-}y} = \begin{bmatrix} 0.2 & 0.2 \end{bmatrix} \qquad b_y = \begin{bmatrix} 0.6 & 0.9 \end{bmatrix}$$

(a) In the formulas below, $\sigma(X)$ stands for element-wise sigmoid of matrix X.

$$h_{in} = xw_{x\text{-}h} + b_h = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0.3 \\ 0.5 \end{bmatrix} + \begin{bmatrix} 0.1 \end{bmatrix} = \begin{bmatrix} 0.4 \end{bmatrix}$$

$$h = \sigma(h_{in}) = \sigma\left(\begin{bmatrix} 0.4 \end{bmatrix}\right) = \begin{bmatrix} 0.599 \end{bmatrix}$$

$$y_{in} = hw_{h\text{-}y} + b_y = \begin{bmatrix} 0.599 \end{bmatrix} \begin{bmatrix} 0.2 & 0.2 \end{bmatrix} + \begin{bmatrix} 0.6 & 0.9 \end{bmatrix} = \begin{bmatrix} 0.72 & 1.02 \end{bmatrix}$$

$$y = \sigma\left(\begin{bmatrix} 0.72 & 1.02 \end{bmatrix}\right) = \begin{bmatrix} 0.673 & 0.735 \end{bmatrix}$$

(b) In the formulas below, $X \circ Y$ stands for element-wise multiplication of matrices X and Y, and $1 - X$ is equivalent to subtracting the matrix X from a matrix with the same size of X, but with all elements equal to 1.

$$\delta_y = y \circ (1 - y) \circ (y - t) =$$

$$\begin{bmatrix} 0.673 & 0.735 \end{bmatrix} \circ \left(1 - \begin{bmatrix} 0.673 & 0.735 \end{bmatrix}\right) \circ \left(\begin{bmatrix} 0.673 & 0.735 \end{bmatrix} - \begin{bmatrix} 0 & 1 \end{bmatrix}\right) =$$

$$\begin{bmatrix} 0.148 & -0.052 \end{bmatrix}$$

$$\delta_h = h \circ (1 - h) \circ (\delta_y w_{h\text{-}y}^T) =$$

$$\begin{bmatrix} 0.599 \end{bmatrix} \circ \left(1 - \begin{bmatrix} 0.599 \end{bmatrix}\right) \circ \left(\begin{bmatrix} 0.148 & -0.052 \end{bmatrix} \begin{bmatrix} 0.2 \\ 0.2 \end{bmatrix}\right) = \begin{bmatrix} 0.005 \end{bmatrix}$$

$$\Delta w_{x\text{-}h} = -\eta(x^T \delta_h) = -0.1 \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0.005 \end{bmatrix}\right) = \begin{bmatrix} -0.0005 \\ 0 \end{bmatrix}$$

$$\Delta b_h = -\eta \delta_h = -0.1 \begin{bmatrix} 0.005 \end{bmatrix} = \begin{bmatrix} -0.0005 \end{bmatrix}$$

$$\Delta w_{h\text{-}y} = -\eta(h^T \delta_y) = -0.1 \left(\begin{bmatrix} 0.599 \end{bmatrix} \begin{bmatrix} 0.148 & -0.052 \end{bmatrix}\right) =$$

$$\begin{bmatrix} -0.0088652 & 0.0031148 \end{bmatrix}$$

$$\Delta b_y = -\eta \delta_y = -0.1 \begin{bmatrix} 0.148 & -0.052 \end{bmatrix} = \begin{bmatrix} -0.0148 & 0.0052 \end{bmatrix}$$

$$w_{x\text{-}h,new} = w_{x\text{-}h} + \Delta w_{x\text{-}h} = \begin{bmatrix} 0.3 \\ 0.5 \end{bmatrix} + \begin{bmatrix} -0.0005 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.2995 \\ 0.5 \end{bmatrix}$$

$$b_{h,new} = b_h + \Delta b_h = \begin{bmatrix} 0.1 \end{bmatrix} + \begin{bmatrix} -0.0005 \end{bmatrix} = \begin{bmatrix} 0.0995 \end{bmatrix}$$

$$w_{h\text{-}y,new} = w_{h\text{-}y} + \Delta w_{h\text{-}y} = \begin{bmatrix} 0.2 & 0.2 \end{bmatrix} + \begin{bmatrix} -0.0088652 & 0.0031148 \end{bmatrix} =$$

$$\begin{bmatrix} 0.1911348 & 0.2031148 \end{bmatrix}$$

$$b_{y,new} = b_y + \Delta b_y = \begin{bmatrix} 0.6 & 0.9 \end{bmatrix} + \begin{bmatrix} -0.0148 & 0.0052 \end{bmatrix} = \begin{bmatrix} 0.5852 & 0.9052 \end{bmatrix}$$