Homework 03 Neural Network

Let X = { $x^{(1)}$,..., $x^{(m)}$ } be a dataset of m samples with 2 features, i.e $x^{(i)} \in R^2$. The samples are classied into 2 categories with labels $y^{(i)} \in \{0,1\}$. A scatter plot of the dataset is shown in Fig.1:

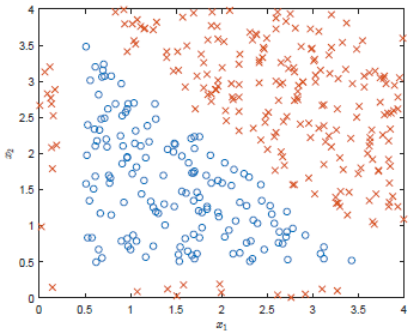

Figure 1: Plot of dataset X.

The examples in class 1 are marked as as "×" and examples in class 0 are marked as "∘". We want to perform binary classication using a simple neural network with the architecture shown in Figure 1:

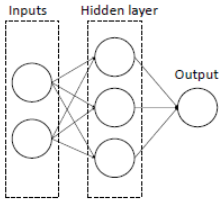

Figure 2: Architecture for our simple neural network.

Denote the two features x_1 and x_2 ,the three neurons in the hidden layer h_1 , h_2 , and h_3 ,and the output neuron as o. Let the weight from x_i to h_j be $w_{i,j}^{[1]}$ for $i \in \{1,2\}$, $j \in \{1,2,3\}$, and the weight from h_j to o be $w_j^{[2]}$. Finally, denote the intercept weight for h_j as $w_{0,j}^{[1]}$ and the intercept weight for o as $w_{0,j}^{[2]}$.

For the loss function, we'll use average squared loss instead of the usual negative log-likelihood:

$$l = \frac{1}{m}\sum_{i=1}^{m}(o^{(i)} - y^{(i)})^2$$

where $o^{(i)}$ is the result of the output neuron for example i.

(a) Suppose we use the sigmoid function as the activation function for h_1, h_2, h_3 and o. What is the gradient descent update to $w_{1,2}^{[1]}$, assuming we use a learning rate of α ? Your answer should be written in terms of $x^{(i)}, o^{(i)}, y^{(i)}$, and the weights.

(b) Now, suppose instead of using the sigmoid function for the activation function for h_1, h_2, h_3 and o, we instead used the step function $f(x)$, defined as

$$f(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases}$$

What is one set of weights that would allow the neural network to classify this dataset with 100% accuracy? Please specify a value for the weights in the order given below and explain your reasoning.

$$w_{0,1}^{[1]} = ?, w_{1,1}^{[1]} = ?, w_{2,1}^{[1]} = ?$$
$$w_{0,2}^{[1]} = ?, w_{1,2}^{[1]} = ?, w_{2,2}^{[1]} = ?$$
$$w_{0,3}^{[1]} = ?, w_{1,3}^{[1]} = ?, w_{2,3}^{[1]} = ?$$
$$w_0^{[2]} = ?, w_1^{[2]} = ?, w_2^{[2]} = ?, w_3^{[2]} = ?$$

Hint: There are three sides to a triangle, and there are three neurons in the hidden layer.

(c) Let the activation functions for h_1, h_2, h_3 be the linear function $f(x) = x$ and the activation function for o be the same step function as before. Is there a specic set of weights that will make the loss 0? If yes, please explicitly state a value for every weight. If not, please explain your reasoning.