Homework 4

Question 1: Linear Classifiers
Recall that a linear threshold function or a linear classifier is given by:

If $(w_0 + \sum_i w_i x_i) > 0$ then class is positive, otherwise it is negative.

Assume that 1 is true and 0 is false.

Q1.1 Consider a function over n Binary features, defined as follows. If at least k variables are false, then the class is positive, otherwise the class is negative. Can you represent this function using a linear threshold function.

If your answer is YES, then give a precise numerical setting of the weights. Otherwise, clearly explain, why this function cannot be represented using a linear threshold function.

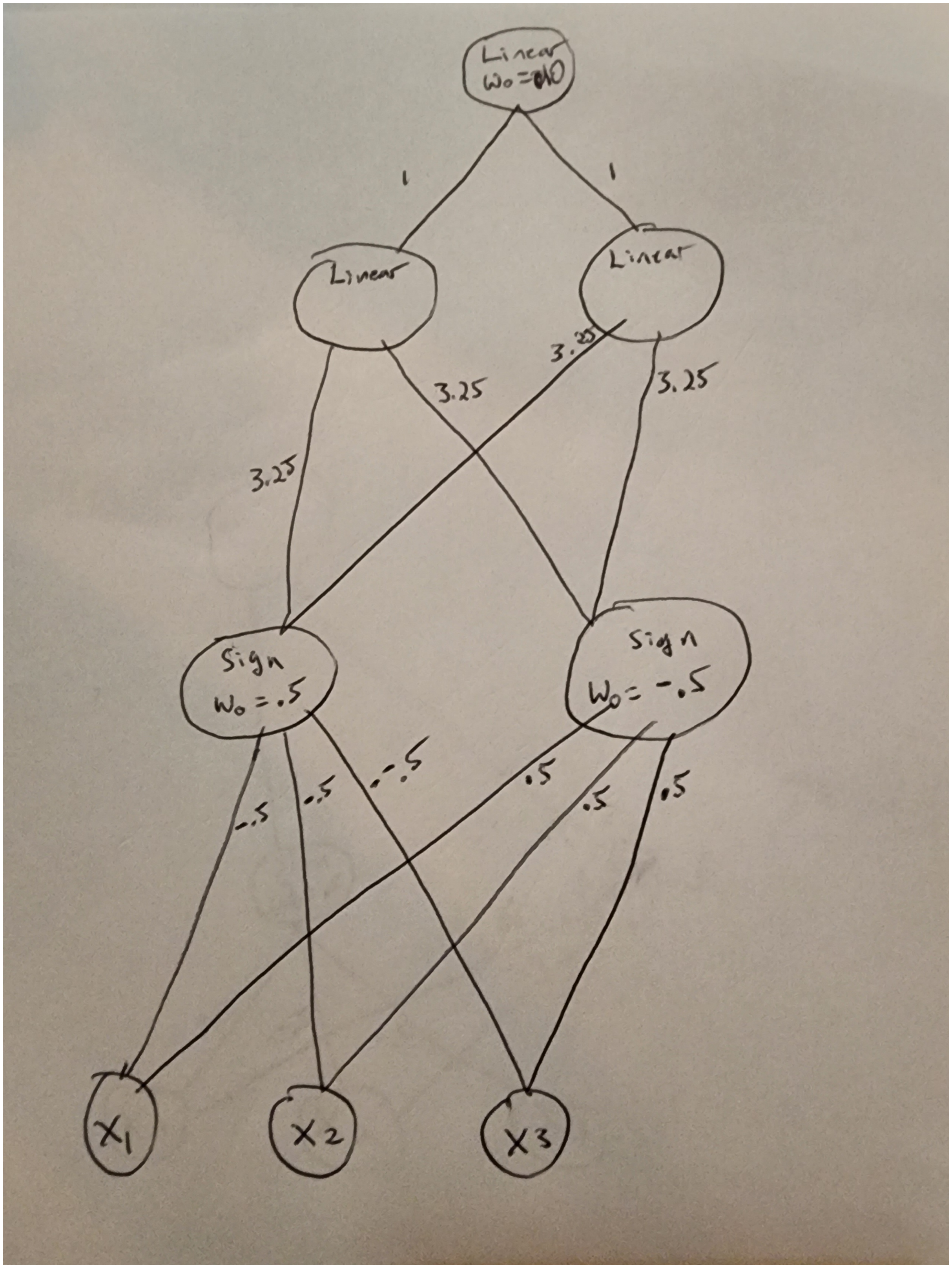
Answer: Yes, we can do this if we take advantage of the bias weight. Let $w_0 = n + 1 - k$, this represents the most amount of positive variables we can have for our example to be classified as positive. Let $w_i = -\frac{1}{n-k}$. If we have k or more false variables, our linear classifier will classify correctly.

Q1.2 Consider a linear threshold function over n real-valued features having $w_0 = 0$ (namely the bias term is zero). Can you always represent such a function using a linear threshold function having only $n - 1$ features? Answer YES or NO and briefly explain your answer. Note that no credit will be given if your explanation is incorrect.

Answer:
No. Say we have two features, or in other words, two dimensions and a line separating examples that goes through the origin and a slope of one. If we had a point above the line and to the left of a point below the line, and a point above the line and to the right (positive x) of the same point below the line and we collapsed these points to one dimension using only their x value, there would be no straight line that could separate the point that was below the line from the two points that were above the line.

Question 2: Neural Networks

Q2.1 Draw a neural network that represents the function $f(x1, x2, x3)$ defined below. You can only use two types of units: linear units and sign units. Recall that the linear unit takes as input weights and attribute values and outputs $w_0 + \sum_i w_i x_i$, while the sign unit outputs +1 if $(w_0 + \sum_i w_i x_i) > 0$ and -1 otherwise.



Q2.2 Derive the back-propagation algorithm for the neural network with three output nodes given on class slides.

Question 3: Linear Regression

Q3.1 Derive the expression for univariate linear regression (see class slides on linear regression).
$Loss(w_0, w_1) = \sum_{i=1}^{n}(y_i - (w_0 + w_1x_i))^2$
$\frac{\delta Loss}{\delta w_0} = 2\sum_{i=1}^{n}(y_i - w_0 - w_1x_i)$
$\frac{\delta Loss}{\delta w_1} = 2\sum_{i=1}^{n}(y_i - w_0 - w_1x_i)x_i$
From this we get that:
$w_0 = w_0 - 2\sum_{i=1}^{n}(y_i - w_0 - w_1x_i)$
$w_1 = w_1 - 2\sum_{i=1}^{n}(y_i - w_0 - w_1x_i)x_i$