Problem 1: Support Vector Machines [10 points]

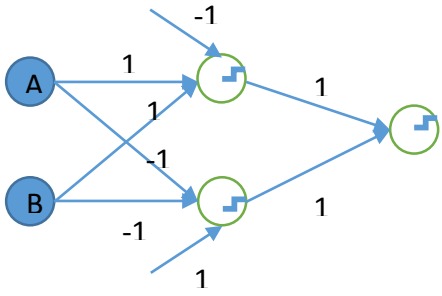
Do Exercise 18.17 in the textbook: Construct a support vector machine that computes the XOR function. Use values of +1 and -1 (instead of 1 and 0) for both inputs and outputs, so that an example looks like ([-1, 1], 1) or ([-1, -1], -1). Map the input [x_1, x_2] into a space consisting of x_1 and x_1x_2 . Draw the four input points in this space, and the maximal margin separator. What is the margin? Now draw the separating line back in the original Euclidean input space.

The examples map inputs from [x_1, x_2] to [x_1, x_1x_2] as follows:
[−1,−1] (negative) maps to [−1,+1]
[−1,+1] (positive) maps to [−1,−1]
[+1,−1] (positive) maps to [+1,−1]
[+1,+1] (negative) maps to [+1,+1]
Thus, the positive examples have $x_1x_2 = -1$ and the negative examples have $x_1x_2 = +1$. The maximum margin separator is the line $x_1x_2 = 0$, with a margin of 1. The separator corresponds to the $x_1 = 0$ *and* $x_2 = 0$ axes in the original space — this can be thought of as the limit of a hyperbolic separator with two branches.

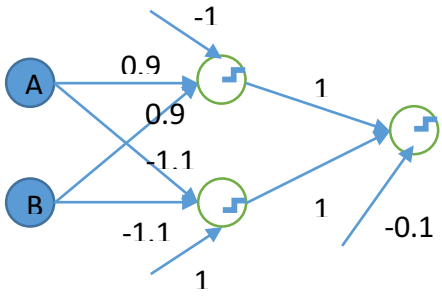
Problem 2: Neural Networks [20 points]

a).
Use a step function (LTU) as the activation function. Draw a figure that shows your network topology and the weights and bias values used. Use as less hidden layers as you can.

$$(A \vee \neg B) \wedge (\neg A \vee B) \equiv A \Leftrightarrow B \equiv \neg(A \; XOR \; B)$$



This one is given if the threshold point has value of 1. (threshold:0)



This one is given if the threshold point has value of 0. (threshold:0)

2b)

$D = g(w_{ad}A + w_{bd}B + w_{cd}C + w_{id}\cdot 1) = g(0.56) = 0.6365$

$E = g(w_{ae}A + w_{be}B + w_{cd}C + w_{id}\cdot 1) = g(0.08) = 0.52$

$F = g(w_{df}D + w_{ef}E + w_{if}\cdot 1) = g(0.3389) = 0.5839$

$\delta_F = \frac{-\partial Loss}{\partial net F} \cdot \frac{\partial net F}{\partial F} = (Y-F)F(1-F) = 0.1011$

$\delta_D = D(1-D)\,\delta_F\, w_{df} = 0.0070$

$\delta_E = E(1-E)\,\delta_F\, w_{ef} = -0.0025$

$\Delta w_{df} = \alpha \delta_F D = 0.0129$ $\quad w_{df} = 0.3129$

$\Delta w_{ef} = \alpha \delta_F E = 0.0105$ $\quad w_{ef} = -0.0895$

$\Delta w_{if} = \alpha \delta_F \cdot 1 = 0.0202$ $\quad w_{if} = 0.2202$

$\Delta w_{ad} = \alpha \delta_D A = 0.0004$ $\quad w_{ad} = 0.3004$

$\Delta w_{bd} = \alpha \delta_D B = 0.0011$ $\quad w_{bd} = 0.3011$

$\Delta w_{cd} = \alpha \delta_D C = 0.0001$ $\quad w_{cd} = 0.3001$

$\Delta w_{id} = \alpha \delta_D 1 = 0.0014$ $\quad w_{id} = 0.2014$

$\Delta w_{ae} = \alpha \delta_E A = -0.0002$ $\quad w_{ae} = -0.1002$

$\Delta w_{be} = \alpha \delta_E B = -0.0004$ $\quad w_{be} = -0.1004$

$\Delta w_{ce} = \alpha \delta_E C = -0.0001$ $\quad w_{ce} = -0.1001$

$\Delta w_{ie} = \alpha \delta_E 1 = -0.0005$ $\quad w_{ie} = 0.1995$