QUESTION 1

Consider the single layered network shown here. Suppose the values of the various quantities for a given input are the following:

\[ x_0 = 1 \quad v_0 = ?? \quad \hat{y} = ?? \]
\[ x_1 = 10 \quad v_1 = ?? \quad y = 0.25 \]
\[ x_2 = 4 \quad v_2 = ?? \quad t = 0 \]

where \( t \) denotes the true class for the data sample \( x \). Use the nonlinearity \( y = \tanh(\hat{y}) \) and error functional \( E = \frac{1}{2}(y - t)^2 \). What are the quantities

\[ \frac{\partial E}{\partial y}, \frac{\partial y}{\partial \hat{y}}, \frac{\partial \hat{y}}{\partial v_j}, j = 0, 1, 2, \]

and the resulting update to the coefficients \( v \)?

Repeat the above if we replace the items as follows: \( y = \text{sigmoid}(\hat{y}) = \frac{1}{1 + e^{-\hat{y}}}, \quad E = -t \log y - (1-t) \log(1-y) \).

QUESTION 2

Consider the nonlinearly-separable problem with points

<table>
<thead>
<tr>
<th>( k )</th>
<th>( x_1^{[k]} )</th>
<th>( x_2^{[k]} )</th>
<th>label = ( Y^{[k]} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>+1</td>
</tr>
<tr>
<td>2</td>
<td>-1</td>
<td>0</td>
<td>-1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>-1</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>-1</td>
<td>+1</td>
</tr>
</tbody>
</table>

and kernel \( K(u, v) = (1 + \langle u, v \rangle)^2 \) (quadratic kernel)

1. Find the explicit embedding \( \phi(x) \) of the training points in high dimensional space.
2. Find a weight vector \( w \) such that the inner product \( \langle w, \phi(x^{[k]}) \rangle + b = y^{[k]} \). Here \( b \) is a bias to be determined.
3. Write the weight vector as a linear combination of the training points \( w = y^{[1]} \alpha \phi(x^{[1]}) + y^{[2]} \alpha \phi(x^{[2]}) + y^{[3]} \alpha \phi(x^{[3]}) + y^{[4]} \alpha \phi(x^{[4]}) \). By symmetry, all the \( \alpha \)'s are the same. What is \( \alpha \)?