Question 1

$$h_1 = \sigma(2 + 4 + 2) = 8$$

$$h_2 = \sigma(-2 + 12 - 8) = 2$$

$$y_1 = \sigma(4 - 1) = 3$$

$$y_2 = \sigma(-4 + 1) = 0$$

Question 2-a

$$\hat{y^i} = \max(0, z^i) = \max(0, \sum_{j=0}^{P} w_j x_j^i)$$

Question 2-b

i)

$$\frac{\partial f(w)}{\partial w_j} = \sum_{i=1}^{n} \frac{\partial f}{\partial \hat{y}_i} \cdot \frac{\partial \hat{y}_i}{\partial z_i} \cdot \frac{\partial z_i}{\partial w_j}$$

$$\frac{\partial f}{\partial \hat{y}_i} = (\hat{y}_i - y_i)$$

$$\frac{\partial \hat{y}_i}{\partial z_i} = \begin{cases} 1 & \text{if } z_i > 0 \\ 0 & \text{if } z_i \leq 0 \end{cases}$$

$$\frac{\partial z_i}{\partial w_j} = x_j^i$$

$$\frac{\partial f(w)}{\partial w_j} = \sum_{i=1}^{n} (\hat{y}_i - y_i) \cdot 1_{z_i > 0} \cdot x_{ij}$$

ii)

```
for epoch in range(epochs):
    for i in range(n):
        z_i = sum(w[j] * X[i][j] for j in range(p+1))
        y_hat = max(0, z_i)
        error = y_hat - y[i]

        for j in range(p+1):
            grad = error * X[i][j] if z_i > 0 else 0
```

```
w[j] = w[j] - alpha * grad
```

Question 2-c

$$\frac{\partial f^i(w)}{\partial w^i} = \frac{\partial f^i(w)}{\partial y^i} \cdot \frac{\partial y^i}{\partial z^i} \cdot \frac{\partial z^i}{\partial w^i}$$

$$\frac{\partial f^i(w)}{\partial y^i} = (\hat{y}_i - y_i) + 2\lambda w_j$$

$$\frac{\partial y^i}{\partial z^i} = \hat{y}_i(1 - \hat{y}_i)$$

$$\frac{\partial z^i}{\partial w^i} = x^i_j$$

$$\frac{\partial f_i(w)}{\partial w^i} = (\hat{y}_i - y_i) \cdot \hat{y}_i(1 - \hat{y}_i) \cdot x_{ij} + 2\lambda w_j$$