DATA.ML.300 Computer Vision Exercise 2

Väinö Pollari, H274459

January 2023

1

a) Output of the hidden unit and the output unit, x = input unit and b = bias

$$x1 = x * w + b = 1 * (-2) + 2 = 0 \tag{1}$$

$$x2 = \frac{1}{1 + exp(-x1)} = \frac{1}{1 + exp(0)} = \frac{1}{2} \tag{2}$$

The network output unit is

$$y = x2 * w2 + b2 = \frac{1}{2} * 4 + 0 = 2 \tag{3}$$

b) the loss for this training case, t = 1

$$E = \frac{1}{2} * (t - y)^2 = \frac{1}{2} * (1 - 2)^2 = \frac{1}{2} \tag{4}$$

c) derivative of the loss with respect to w2 using the chain rule

$$E = \frac{dE}{dw2} = \frac{dE}{dy}\frac{dy}{dw2} => (\frac{d}{dy}\frac{1}{2} * (t-y)^2) * \frac{x2}{dw2}(y * w2 + b2) => (-t+y) * (x1) = \frac{1}{2} \tag{5}$$

d) The derivative of the loss with respect to w1

$$\frac{dE}{dw1} = \frac{dE}{dy}\frac{dy}{dx2}\frac{dx2}{dx1}\frac{dx1}{dw1} => (-t+y)(w2)(mu(x1) * (1-mu(x1))(x) => 1 * 4 * \frac{1}{2} * (1-\frac{1}{2}) * 1 = 1 \tag{6}$$

2

a) Euclidean distances and cosine similarities

$$d(Q, A) = \sqrt{(2-1)^2 + (1-2)^2 + (6-3)^2 + (4-4)^2 + (2-1)^2} = 3,464 \tag{7}$$

$$d(Q,B) = \sqrt{(2-3)^2 + (1-1)^2 + (6-4)^2 + (4-1)^2 + (2-4)^2} = 4,795 \quad (8)$$

$$similarity(Q,A) = \frac{QxA}{||Q||x||A||} = \frac{40}{\sqrt{61} * \sqrt{31}} = 0.919 \quad (9)$$

$$similarity(Q,B) = \frac{QxB}{||Q||x||B||} = \frac{45}{\sqrt{61} * \sqrt{52}} = 0.799 \quad (10)$$

b) A is better since its euclidean distance is smaller than B and its cosine similarity is bigger than B