60006 - Tutorial 3

Interest Point Detection

Xin Wang

March 15, 2022

Question 1

Question 1.1: Calculate the derivative of $f(x)$ with respect to x.

$$\frac{dy}{dx} = \frac{e^x}{(e^x + 1)^2}$$

Question 1.2: Describe a problem with the sigmoid function when we train a neural network using the gradient descent algorithm.

The problem is the Vanishing Gradient Problem which occurs during backpropagation when value of the weights are changed. When $f(x)$ saturates at either 0 or 1, its derivative $f'(x)$ becomes nearly 0. This causes the learning process to be very slow and cause it to converge to their optimum.

Question 2

Question 2.1:

$$\max(\sigma(\max(x, x)), \sigma(x))$$
$$\max(\sigma(x), \sigma(x))$$
$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

Question 2.2:

$$M.M.(M^2.N).N = O(M^4.N^2)$$

Question 2.3:

$$\beta = 0 : \frac{x}{2}$$
$$\beta = \infty : \begin{cases} 0, x < 0 \\ x, x \geq 0 \end{cases}$$

Question 3

Question 3.1: Check whether p fulfils the properties of a probability vector, i.e. it is non-negative and its elements sum to 1.

Given that $p_i = \frac{e^{c_i}}{\sum_k e^{c_k}}$:

- The exponentials ensure the numerator and denominator are always positive
- The sum in the denominators ensure the average always added up to 1

Question 3.2: Derive the derivative.

Question 4

Question 4.1:

Data shape and size:

- Layer 1: $1 \times 24 \times 24 \times 20$ - $46.08KB$
- Layer 2: $1 \times 12 \times 12 \times 20$ - $11.52KB$
- Layer 3: $1 \times 8 \times 8 \times 50$ - $12.80KB$
- Layer 4: $1 \times 4 \times 4 \times 50$ - $3.20KB$
- Layer 5: $1 \times 1 \times 1 \times 500$ - $2.00KB$
- Layer 6: $1 \times 1 \times 1 \times 10$ - $0.04KB$
- Layer 7: 1 - $0.004KB$

Question 4.2: Calculate the receptive fields for neurons in the following layers and fill in the table.