Question 1 ..... **(7 Marks)**

Check the correct box (or boxes) for MCQs. Overwriting will result in a zero. Use the space provided for answering short questions.

(a) **(1 Mark)** Using gradient descent always guarantee the same solution.

- ◯ True.
- ◯ False.

(b) **(1 Mark)** A traditional neural network is nothing more than stacks of linear classifier with sigmoid non-linearity.

- ◯ True.
- ◯ False.

(c) **(2 Marks)** What is so different in Artificial Neural Networks than machine learning, be brief ?

(d) **(1 Mark)** Let's suppose that $z = x^T\theta$ (θ being a unit vector) what does the value of z signifies ?

(e) For $\sigma(x) = \frac{1}{1+\exp(-x)}$, as

i. **(1 Mark)** $x \to \infty$, $\sigma(x) \to$ ____

ii. **(1 Mark)** $x \to -\infty$, $\sigma(x) \to$ ____

Continue...

Question II ..... **(10 Marks)**

Given following cost function of two variables,

$$f(x, y) = (1 - x)^2 + 100(y - x^2)^2$$

(a) **(2 Marks)** What will be its gradient vector ?

(b) **(4 Marks)** Write Python code (or pseudocode) for finding the minimum of this function using brute-force method.

(c) **(4 Marks)** Write Python code (or pseudocode) for finding the minimum via gradient-descent.

0061

Question III ..... **(25 Marks)**

Given the two equations, exponential is represented by exp and all other symbols(except x) are weights, x is input data

$$f1 = \frac{c}{1 + exp^{-ax-k}} + \frac{d}{1 + exp^{-ax-k}} + z \quad (1)$$

$$f2 = \frac{e}{1 + exp^{-ax-k}} + \frac{f}{1 + exp^{-ax-k}} + y \quad (2)$$

f1 and f2 are outputs of neurons.

(a) **(5 Marks)** Draw the network represented by the f1 and f2.

(b) i. **(10 Marks)**

$$f(f1, f2) = \frac{exp(f1)}{exp(f1) + exp(f2)} \quad (3)$$

where f1 and f2 are given in above equations

Find the partial derivative of f with respect to all weights, at x=2, a=5, b=4, c=9 , d=12 , e = -3, f= 6 , g= -2, h= -1

Continue...

ii. **(10 Marks)** Write the code for finding the partial derivative of f given values of x and all weights. The function takes 11 numbers as arguments and returns tuple of 10 values(derivatives w.r.t all weights)