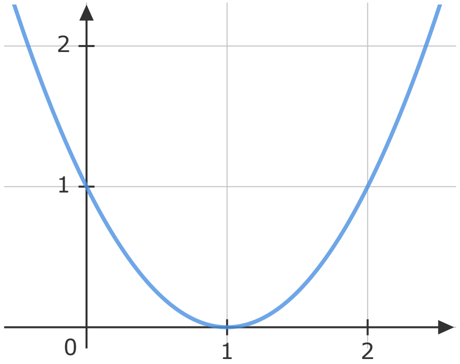
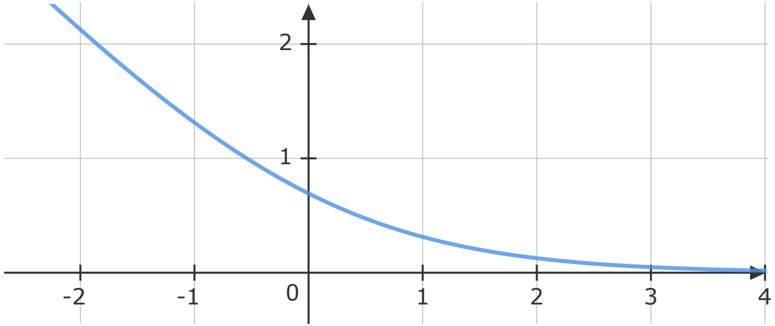
Week 12 | Solve With Instructors

Q1) Match the following loss functions graph with their correct mathematical representation

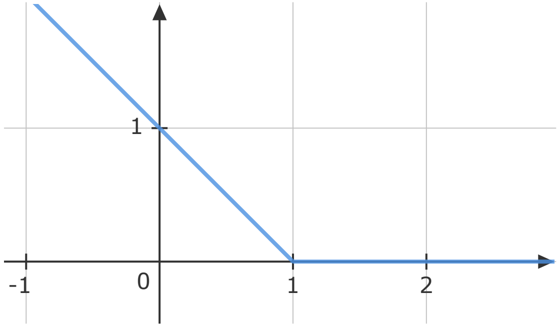
A)



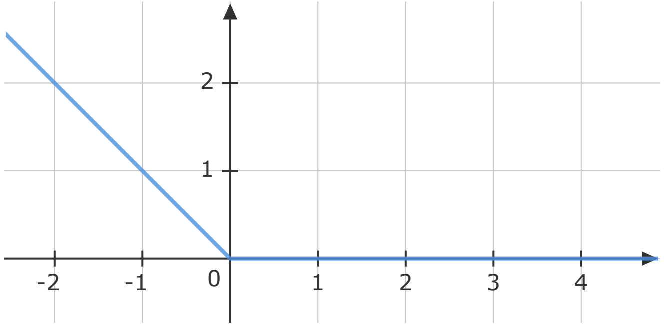
B)



C)



D)



a) A $\rightarrow$ Perceptron Loss, B $\rightarrow$ Logistic Loss

b) A $\rightarrow$ Logistic Loss, D $\rightarrow$ Perceptron Loss

c) C $\rightarrow$ Hinge Loss, B $\rightarrow$ Perceptron Loss

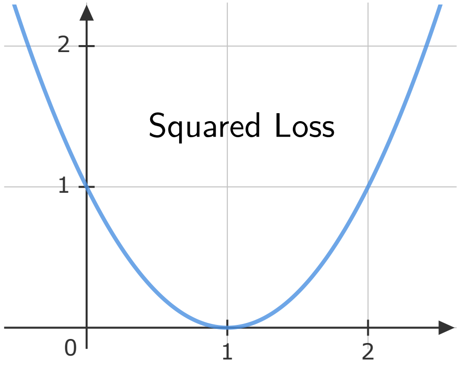
d) D $\rightarrow$ Perceptron Loss, C $\rightarrow$ Hinge Loss

e) B $\rightarrow$ Logistic Loss, A $\rightarrow$ Hinge Loss

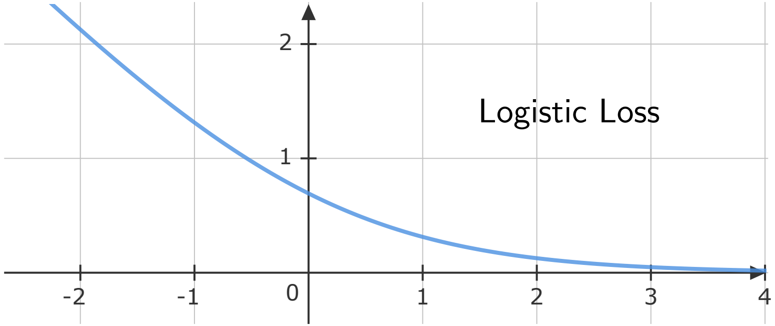
f) C $\rightarrow$ Squared Loss, D $\rightarrow$ Perceptron Loss

Solution: Below is the correct matching for each loss function with its graphical representation

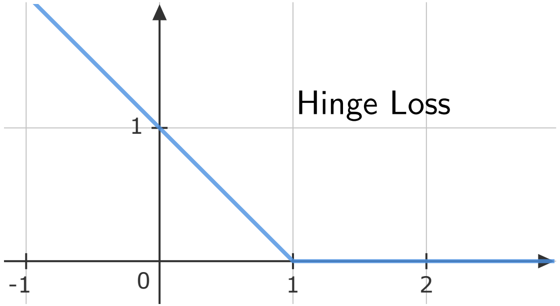
A)



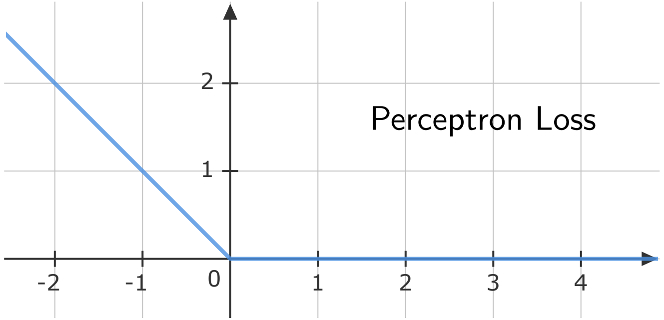
B)



C)



D)



Q2) Which of the following statements about the sigmoid function

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

is (are) true?

a) The function is continuously differentiable
b) The function is monotonic
c) The function attains its maximum when $z \to -\infty$
d) The function is bounded between 0 and 1

Solution:

1. The function is continuously differentiable
- True. The sigmoid function is smooth and continuous for all real values of z. Its derivative also exists everywhere, so it is continuously differentiable.

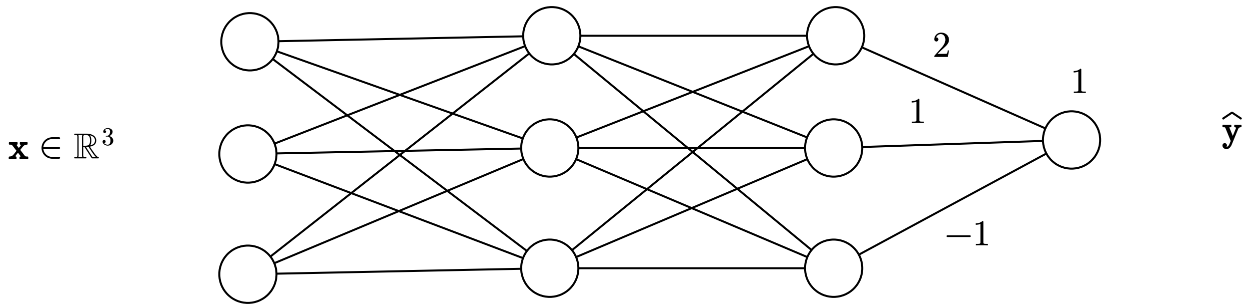
2. The function is monotonic
- True* The sigmoid function is monotonic. Specifically, it is increasing as z increases. The derivative of the sigmoid function is always positive, indicating that the function is monotonically increasing over the entire domain of z.

3. The function attains its maximum when $z \to -\infty$
- False. As $z \to -\infty$, the sigmoid function approaches 0, not its maximum. The function's maximum value is attained as $z \to \infty$, where it approaches 1. Thus, the maximum value is 1, and this is approached when z becomes very large (positive infinity).

4. The function is bounded between 0 and 1
- True. The sigmoid function is bounded between 0 and 1 for all real values of z. As $z \to -\infty$, $z \to 0$ and as $z \to -\infty$, $z \to 1$. Thus, $0 < \sigma(z) < 1$ for all values of z.

Q3) The diagram below shows a neural network used for a classification problem. The network contains two hidden layers and one output layer. The input to the network is a column vector $\mathbf{x} \in \mathbb{R}^3$. The first hidden layer contains 3 neurons, the second hidden layer contains 3 neurons and the output layer contains 1 neurons. Each neuron in the l^{th} layer is connected to all the neurons in the $(l+1)^{\text{th}}$. Each neuron has a bias connected to it except the input layer. The true output is 1.

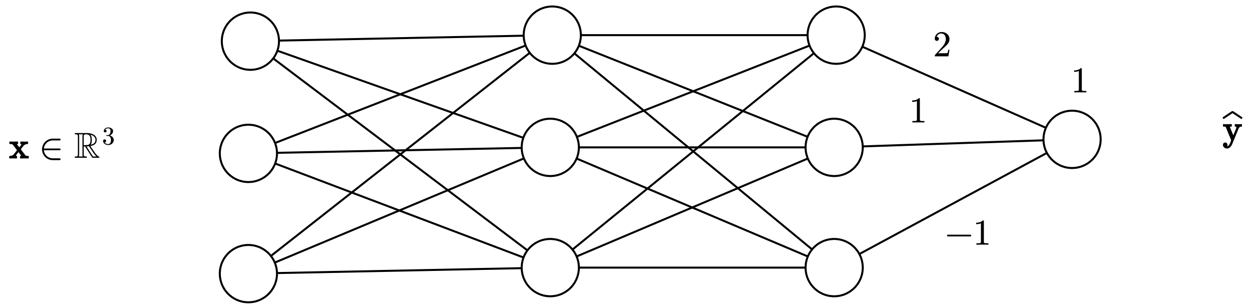
All the neurons in the hidden and output layers use Sigmoid activation function. Assume that the network uses the cross entropy loss (use natural log).



i) How many parameters are there in the network?

ii) Let the output of 2^{nd} hidden layer (after applying activation function) is $\begin{bmatrix} 0.2 & 0.5 & 0.8 \end{bmatrix}^T$. Find the cross entropy loss if $\begin{bmatrix} 2 & 1 & -1 \end{bmatrix}^T$ is the weight of the i^{th} element and 1 is the bias. Enter your answer up to two decimal points accuracy.

Solution:



Total number of weights $(3 \times 3) + (3 \times 3) + (3 \times 1) = 21$

Total number of bias $3 + 3 + 1 = 7$

Hence total number of parameters are 28.

ii) The output of 2^{nd} hidden layer (after applying activation function) is $\begin{bmatrix} 0.2 & 0.5 & 0.8 \end{bmatrix}^T$. After applying weights and bias

$$\begin{bmatrix} 0.2 \\ 0.5 \\ 0.8 \end{bmatrix} \begin{bmatrix} 2 & 1 & -1 \end{bmatrix} = 0.1$$

$$0.1 + 1 = 1.1$$

To get the final output, we need to apply activation function which is Sigmoid

$$\frac{1}{1 + e^{-1.1}} \approx 0.75$$

The cross entropy loss is given by

$$-y\ln(\widehat{y}) - (1 - y)\ln(1 - \widehat{y})$$

$$= -1\ln(0.75) - (1 - 1)\ln(1 - 0.75) \approx 0.29$$

Q4) What will be the result of applying ReLU to the following values?:

$$-2.7,\ 3.9,\ -1.0,\ 4.2,\ 6.4,\ -7.3$$

a) $1,\ 3.9,\ 1,\ 4.2,\ 6.4,\ 1$
b) $0,\ 1,\ 0,\ 1,\ 1,\ 0$
c) $-2.7,\ 0,\ -1.0,\ 0,\ 0,\ -7.3$
d) $-1,\ +1,\ -1,\ +1,\ +1,\ -1$
e) $0,\ 3.9,\ 0,\ 4.2,\ 6.4,\ 0$

Solution: The ReLU function is

$$g(x) = \max(0, x) = \begin{cases} x & x \geqslant 0 \\ 0 & x < 0 \end{cases}$$

So the values $-2.7,\ -1.0,\ -7.3$ will be mapped to 0 while rest will remain as it as.

Hence the final answer is $0,\ 3.9,\ 0,\ 4.2,\ 6.4,\ 0$.