**Module 5**

1.What activation function should I use in the case of ?

1. Classifiers

Ans: For Binary classification: Sigmoid or Tanh Activation functions

For More than 2classification: Softmax activation function

1. Vanishing gradient problem

Ans: Relu Activation function

1. Generalised

Ans: Y = Continuous use Linear or identity activation function

Y = 2 Categories use Sigmoid or Tanh activation function

Y = >2 Categories use Softmax activation function

1. Only in hidden layers

Ans: The modern default activation function for hidden layer is the Relu function

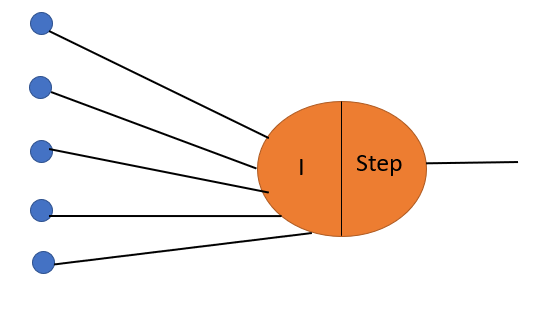
1. Only in Output layer

Ans: Generally using Relu for most of the cases

2. Calculate the error of a following perceptron

Whose values are given as below

|  |  |  |  |
| --- | --- | --- | --- |
| Sno | Inputs(x) | Weights | Output |
| 1 | 0.5 | 0.6 | 1.62 |
| 2 | 0.006 | 0.001 |  |
| 3 | 0.1 | 0.005 |  |
| 4 | 0.4 | 0.67 |  |
| 5 | 0.8 | 0.45 |  |



Ans: Calculate predicted value of Y= W1\*X1 + W2\*X2 + W3\*X3 + W4\*X4 + W5\*X5 + Bias\*1

= 0.5\*0.6 + 0.006\*0.001 + 0.1\*0.005 + 0.4\*0.67 +0.8\* 0.45+1

= 1.481

Error or Loss function = Actual value – Predicted value

= 1.62 – 1.481

= 0.139

3. Write down the steps involved in Back Propagation and the disadvantages of using Back Propagation.

Ans: Steps involved in Back Propagation:

1. The weights are randomly initialised for each and every neuron

2. Calculate the summation or integration function for each neuron

3. Take the function of integration that is nothing but with the help of required activation function

4. Finally calculate the loss function

5. Calculate derivative of old weights with the help of chain rule

6. Now update old weight by subtracting old weight with the derivatives

7.Finally will get new updated weight

Disadvantages of using Back Propagation:

1.The actual performance of backpropagation on a specific problem is dependent on the input data.

2.Back propagation algorithm in data mining can be quite sensitive to noisy data.

3. You need to use the matrix-based approach for backpropagation instead of mini-batch.