## **Loss function**

### **Loss function for Binary Classification**

What is the loss function that you can use for a binary classification problem

1. In normal cases, the number of neurons in the output layer would be equal to the number of classes
2. However a shortcut in the case of binary classification would be to use only one output neuron that uses a sigmoid function. Here is a diagrammatic representation of that configuration
3. Here, ŷ = P(y = 1)
4. Therefore, we can obtain P(y = 0) = 1 - P(y = 1)
5. Consider the following values for the variables
   1. y = 1

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | 0.9 | 0.2 | 0.4 | 0.3 |  |
| W1= | -0.5 | 0.4 | 0.3 | 0.3 |
|  | 0.1 | 0.1 | -0.1 | 0.2 |
|  | -0.2 | 0.5 | 0.5 | / |

* 1. W2 = [0.5 0.8 -0.6 0.3]
  2. x = [0.3 0.5 -0.4 0.3]

1. The output values are as follows
   1. In this case
   2. Cross Entropy Loss:
      1. In this case, since y = 1
2. Consider another case where x = [-0.6 -0.6 0.2 0.3] and true class y = 1
3. The output values are as follows
   1. In this case
   2. Cross Entropy Loss:
      1. In this case, since y = 0
      2. Here, even though the true value was 0, our neuron was outputting a very large value(0.7152) which was already indicative of a large loss value.