**Day 25**

**What to do?**

Learn about softmax and cross entropy for multi-class classification.

**Softmax:**

While sigmoid activation function is used for binary classification, softmax activation is used for multi-class classification. Let us say you want to classify between cats and dogs. That is binary classification, hence, sigmoid activation. What if you want to classify digits from 0 – 9? This is where softmax function comes, as the model must classify 10 different items. Softmax ensures that the values of the classes at the output layer add up to 1. The function of the softmax function is,

**Cross entropy:**

Now that the activation function is ready, it is time to choose a loss function that is appropriate for the multi – class classification problem, which is cross – entropy loss. The goal of the model is to minimize the cross – entropy value. The formula for cross – entropy is,

In the formula above, -p(X) is the probability of class X in target (actual value) and q(X) is the probability of class X in prediction (predicted value). However, the formula above is just for one class. To calculate the loss of multiple classes, the losses are added to equal cross – entropy.