Non-programming Assignment

# 1. What is the reason for softmax?

The reason for using softmax is to convert raw model outputs (logits) into probabilities. This is particularly useful for multi-class classification problems, where we want to determine the probabilities of each class and then choose the class with the highest probability as the model’s prediction. Softmax ensures that the sum of the output probabilities is equal to 1, which makes it easier to interpret the results as probabilities for each class.

# 2. What is softmax and how does it work?

Softmax is an activation function that transforms the raw outputs (logits) of a neural network into a probability distribution. It takes the exponentials of each raw score and divides by the sum of these exponentials across all classes. The resulting values represent the probabilities of each class.

Mathematically, for a set of logits z\_i:

softmax(z\_i) = exp(z\_i) / sum(exp(z\_j)) for all j

This ensures that each z\_i is mapped to a value between 0 and 1, and the sum of all the probabilities is 1.