

The softmax function is defined as follows:

$$\text{softmax}(x_i) = \frac{e^{x_i}}{\sum_{j=1}^n e^{x_j}}$$

This equation represents the softmax function applied to the element x_i of a vector x of length n . The numerator exponentiates x_i , while the denominator sums the exponential values of all elements in the vector. The result is the probability distribution over the elements, with each element being between 0 and 1, and the sum of all elements equal to 1.