1. **Implementation of Rectified Linear Unit (RELU)**

   Defintion of RELU layer

$$y = \max(0, x) \tag{1}$$

   Forward Propagation of RELU layer

$$RELU(x) = \begin{cases} x, & \text{if } x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

   Backward Propagation of RELU layer

$$\frac{d}{dx}(RELU(x)) = \begin{cases} 1, & \text{if } x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

2. **Implementation of Softmax Layer**

   In a multi-class classification problem, softmax layer is used to calculate the probabilities in each class.

   Forward Propagation of softmax layer
   The softmax function for any value a_i is expressed as:

$$p_i = \frac{exp(a_i)}{\sum_i \exp(a_k)} \tag{2}$$

   The computation of softmax value can be huge, because exponential of even a moderately large value is very large number. In order to avoid this, the largest value of x is subtracted from each x value.

   Back Propagation of softmax layer

   We need to calculate the derivative and pass it through previous layers during back propagation.

$$\frac{\partial p_i}{\partial a_j} = \begin{cases} p_i(1 - p_j), & \text{if } i = j \\ -p_j p_i, & \text{if } i \neq j \end{cases}$$