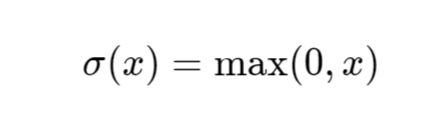
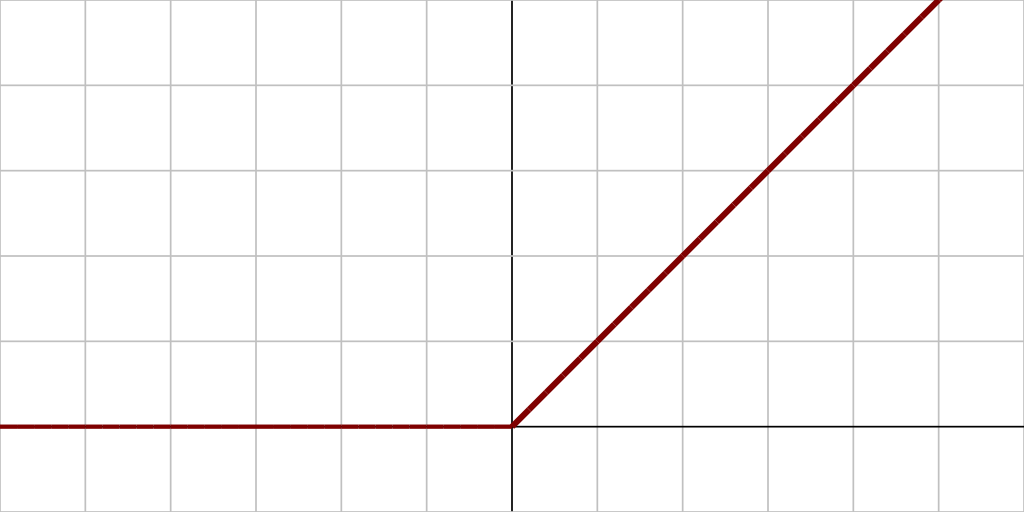
1. **Number of conv layers in model is** 2
2. **Number of dense layers in model is** 1
3. **Which activation function is better and why:**

The better activation function is relu because It returns 0 when the input is smaller than 0, or the value if it’s greater than or equal to 0. In a formula:

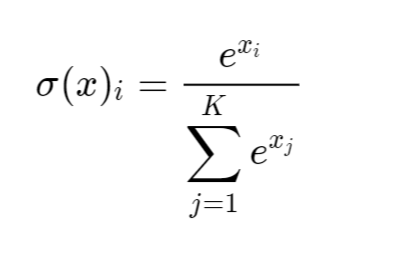


The formula for ReLUs



Almost every neural networks outputs values between 0 and 1 — the probability of an input belonging to some class. Since ReLUs allow values far greater than 0, the values need to be scaled.

The way this is usually done is by applying the softmax activation function. Mathematically, it’s defined as



The formula for softmax

With k as the number of inputs.

This function might seem complex, but the idea is quite simple. Basically, softmax scales the output to range 0 to 1, with all the numbers adding up to 1. This is reasonable because it means the neural network is 100% certain that the input image one of the output categories (note that if there are images in a dataset without a label, they should be labeled as ‘unclassified’ or ‘other’).

Another thing softmax does is scale the numbers based on the input value. The scaled value of the output node with the highest value will be much higher than the node with the second highest output.

ReLU activation functions are a very popular choice among deep learning practitioners because they are very cheap to compute.

1. **The pooling layer(s), where to add them:**

We add pooling layer after every Conv2D Layer like in the following image:



**The effect on the number of parameters and model size:**

The Pooling layer is responsible for reducing the spatial size of the Convolved Feature. This is to decrease the computational power required to process the data through dimensionality reduction. Furthermore, it is useful for extracting dominant features which are rotational and positional invariant, thus maintaining the process of effectively training of the model.