

Chapter 1

Getting Started with TensorFlow 2.x

1.1 Implementing activation functions

The activation functions live in the **neural network(nn)** library in TensorFlow.

The rectified linear unit, known as ReLU, is the most common and basic way to introduce non-linearity into neural networks. This function is just called $\max(0, x)$. It is continuous, but not smooth.

There are times where we'll want to cap the linearly increasing part of the preceding ReLU activation function. We can do this by nesting the $\max(0, x)$ function in a $\min()$ function. The implementation that TensorFlow has is called the ReLU6 function. This is defined as

$$\min(\max(0, x), 6)$$

This is a version of the hard-sigmoid function, is computationally faster, and does not suffer from vanishing (infinitesimally near zero) or exploding values.

The sigmoid function is the most common continuous and smooth activation function. It is also called a logistic function and has the form

$$\frac{1}{(1 + \exp(-x))}$$

The sigmoid function is not used very often because of its tendency to zero-out the backpropagation terms during training.

We should be aware that some activation functions, such as the sigmoid, are not zero-centered. This will require us to zero-mean data prior to using it in most computational graph algorithms

Another smooth activation function is the hyper tangent. The hyper tangent function is very similar to the sigmoid except that instead of having a range between 0 and 1, it has a range between -1 and 1. This function has the form of the ratio of the hyperbolic sine over the hyperbolic cosine. Another way to write this is as follows:

$$\frac{\exp(x) - \exp(-x)}{\exp(x) + \exp(-x)} \tag{1.1}$$

The softsign function is also used as an activation function. The form of this function is

$$\frac{x}{|x| + 1} \quad (1.2)$$

The softsign function is supposed to be a continuous (but not smooth) approximation to the sign function.

Another function, the softplus function, is a smooth version of the ReLU function. The form of this function is

$$\log(\exp(x) + 1)$$

The **Exponential Linear Unit (ELU)** is very similar to the softplus function except that the bottom asymptote is -1 instead of 0. The form is

$$ELU = \begin{cases} \exp(x) + 1, & x < 0 \\ x, & \text{other} \end{cases} \quad (1.3)$$