Activation function - sigmoid
108-W3-P4-171RMC216

Audars Kārlis Klints

February 2018

0.1 The sigmoid function

The sigmoid function is defined as follows:

$$\sigma(x) = \frac{1}{1+e^{-x}}$$

The sigmoid function derivative $\sigma'(x)$ is defined as follows:

$$\sigma'(x) = \frac{d}{dx}dx = \sigma(x)(1-\sigma(x))$$

Here's a detailed derivation:

$$\frac{d}{dx}dx = \frac{d}{dx}\left[\frac{1}{1+e^{-x}}\right] =$$

$$= \frac{d}{dx}(1+e^{-x})^{-1} =$$

$$= -(1+e^{-x})^{-2}(-e^{-x}) =$$

$$= \frac{e^{-x}}{(1+e^{-x})^2} =$$

$$= \frac{1}{1+e^{-x}} \times \frac{e^{-x}}{1+e^{-x}} =$$

$$= \frac{1}{1+e^{-x}} \times \frac{(1+e^{-x})-1}{1+e^{-x}} =$$

$$= \frac{1}{1+e^{-x}} \times \left(\frac{1+e^{-x}}{1+e^{-x}} - \frac{1}{1+e^{-x}}\right) =$$

$$= \frac{1}{1+e^{-x}} \times \left(1 - \frac{1}{1+e^{-x}}\right) =$$

$$= \sigma(x)(1-\sigma(x))$$

0.2 Function graph

As represented in following graph, it has small output changes in (0 and 1) when the varies in (-∞ and +∞). Mathematically, the function is continuous. A typical sigmoid function is represented in the following graph:

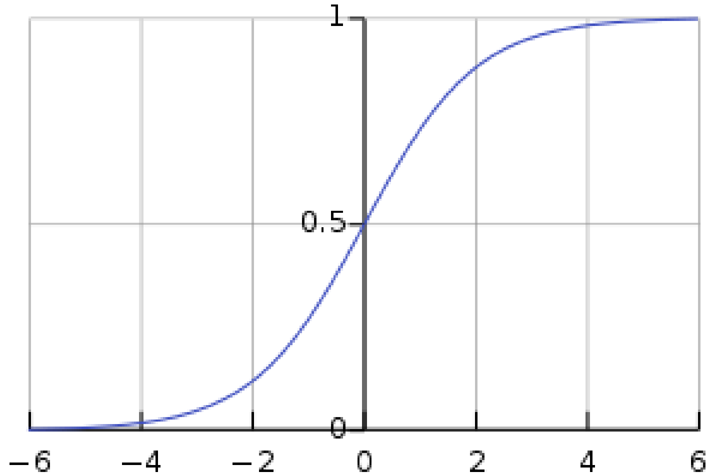

Figure 1: Sigmoid graph

0.3 Sigmoid graph usage

A neuron can use the sigmoid for computing the nonlinear function $\sigma(z = wx + b)$. Note that, if $z = wx + b$ is very large and positive, then $e^{-z} \rightarrow 0$, so $\sigma(z) \rightarrow 1$, while $z = wx + b$ if is very large and negative $e^{-z} \rightarrow \infty$ so $\sigma(z) \rightarrow 0$. In other words, a neuron with sigmoid activation has a behavior similar to the perceptron, but the changes are gradual and output values, such as 0.5539 or 0.123191, are perfectly legitimate. In this sense, a sigmoid neuron can answer *maybe*. [1]

Bibliography

[1] A. Gulli and S. Pal. *Deep Learning with Keras.* Packt Publishing, 2017.