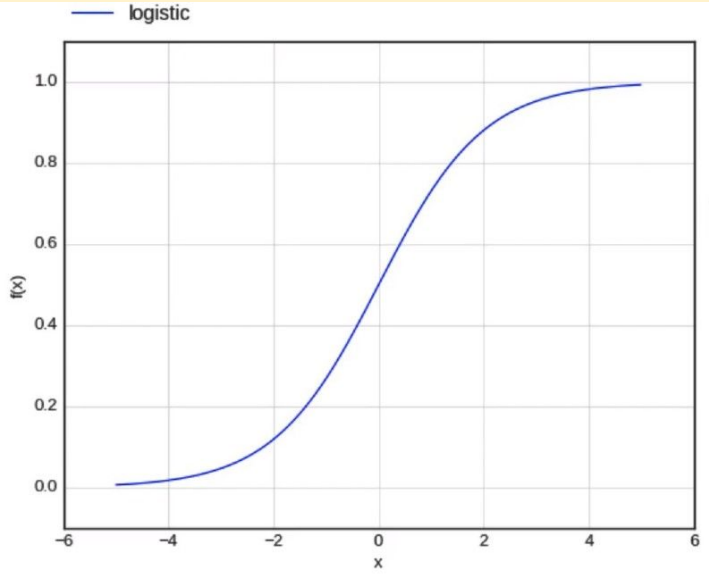
PadhAI: Activation Functions & Initialization Methods

One Fourth Labs

Saturation in logistic neuron

Let's look at the logistic function

1. The following figure illustrates the logistic function



   a. $f(x) = \frac{1}{1+e^{-x}}$
   b. $f'(x) = \frac{\partial f(x)}{\partial x} = f(x) * (1 - f(x))$
2. A logistic neuron is said to be saturated when it reaches its peak values when it is given high extremes of positive or negative values as inputs.
   a. When $f(x) = 0$
   b. And hence $f'(x) = 0$
   c. In the case where we are calculating the gradient w.r.t a weight associated with a saturated neuron, the saturated neuron's derivative is 0, thus resulting in the entire gradient becoming 0
   d. This is because the term associated with the saturated neuron in the chain rule for gradient calculation becomes 0, thus making the entire gradient 0
   e. Due to this, the weights are not updated.
   f. This is called the **Vanishing Gradient Problem**, because the gradient vanishes or becomes 0 due to the presence of a saturated neuron.