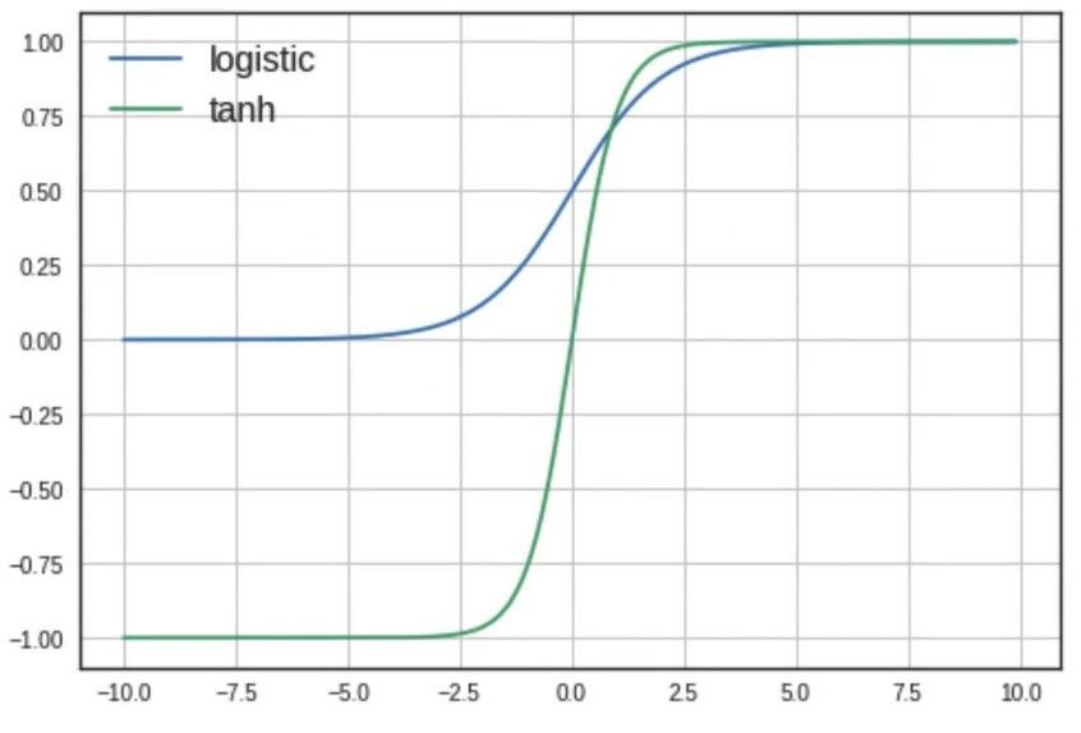
PadhAI: Activation Functions & Initialization Methods

One Fourth Labs

Introducing Tanh and ReLU activation functions

What are the other alternatives to the Logistic function?

1. **tanh**
   a. The following figure illustrates the tanh function



   b. $f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$
   c. $f'(x) = \frac{\partial f(x)}{\partial x} = (1 - (f(x))^2)$
   d. The tanh function ranges from -1 to +1, whereas the logistic function ranges from 0 to 1
   e. It is a zere centered function.
   f. The function saturates at f(x) = -1 or 1, thus causing the gradients to vanish.
   g. tanh is computationally expensive because of e^x
   h. However, it is still preferred over the logistic function

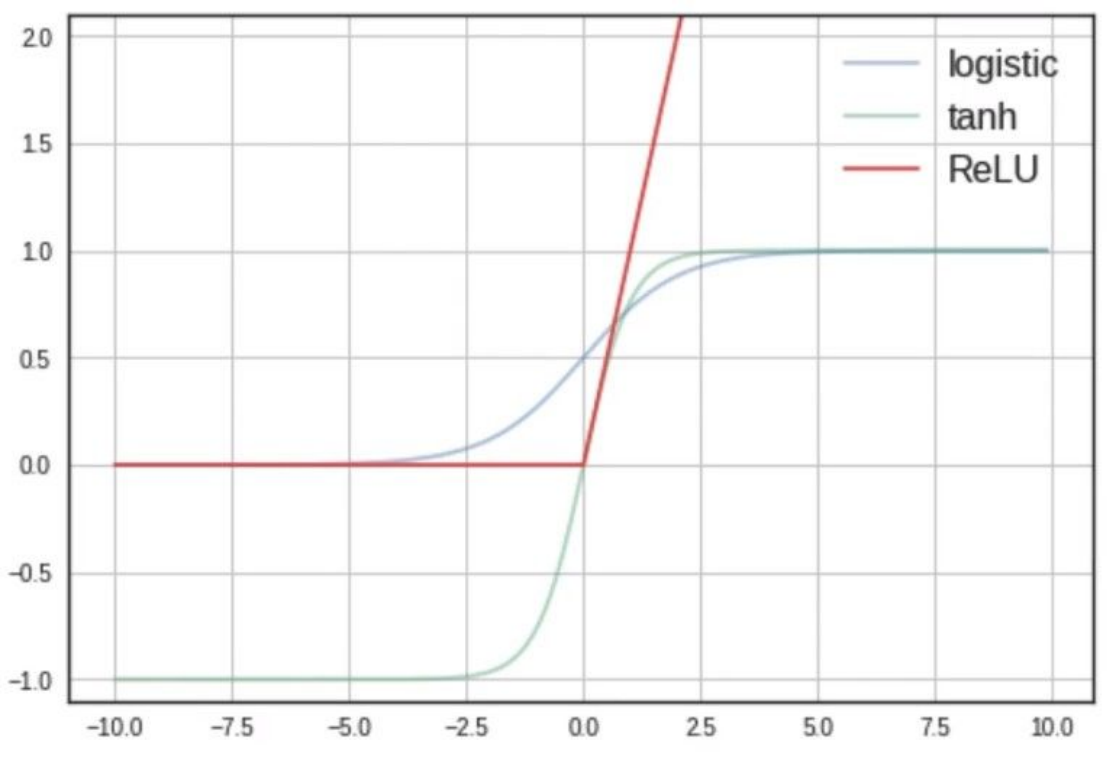
PadhAI: Activation Functions & Initialization Methods

One Fourth Labs

---

2. ReLU

a. The following figure illustrates the ReLU function



b. $f(x) = max(0, x)$

c. $f'(x) = \frac{\partial f(x)}{\partial x} = 0 \text{ if } x < 0 \mid 1 \text{ if } x > 0$

d. ReLU outputs the input value itself if it is positive, else it outputs zero, i.e. f(1) = 1, f(-1) = 0

e. It does not saturate in the positive region

f. It is not zero centered

g. Easy to compute (no expensive e^x)