Activation Functions

 Romeo Kienzler

$$f(x) = sigmoid(x)$$

$$z_2 = XW_1$$

$$a_2 = f(z_2)$$

$$z_3 = a_2 W_2$$

$$\hat{y} = a_3 = f(z_3)$$

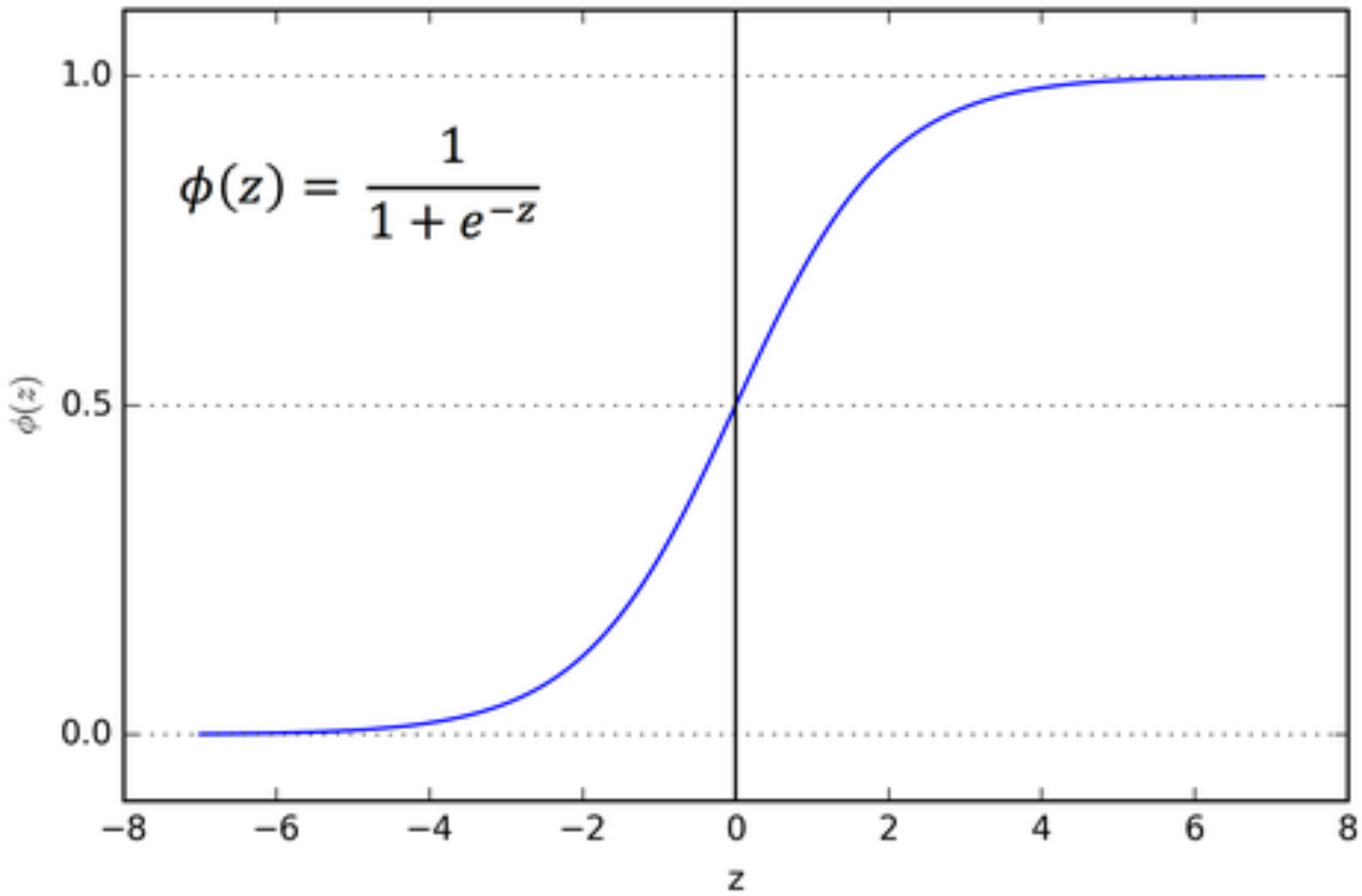


Image Credit:
https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6

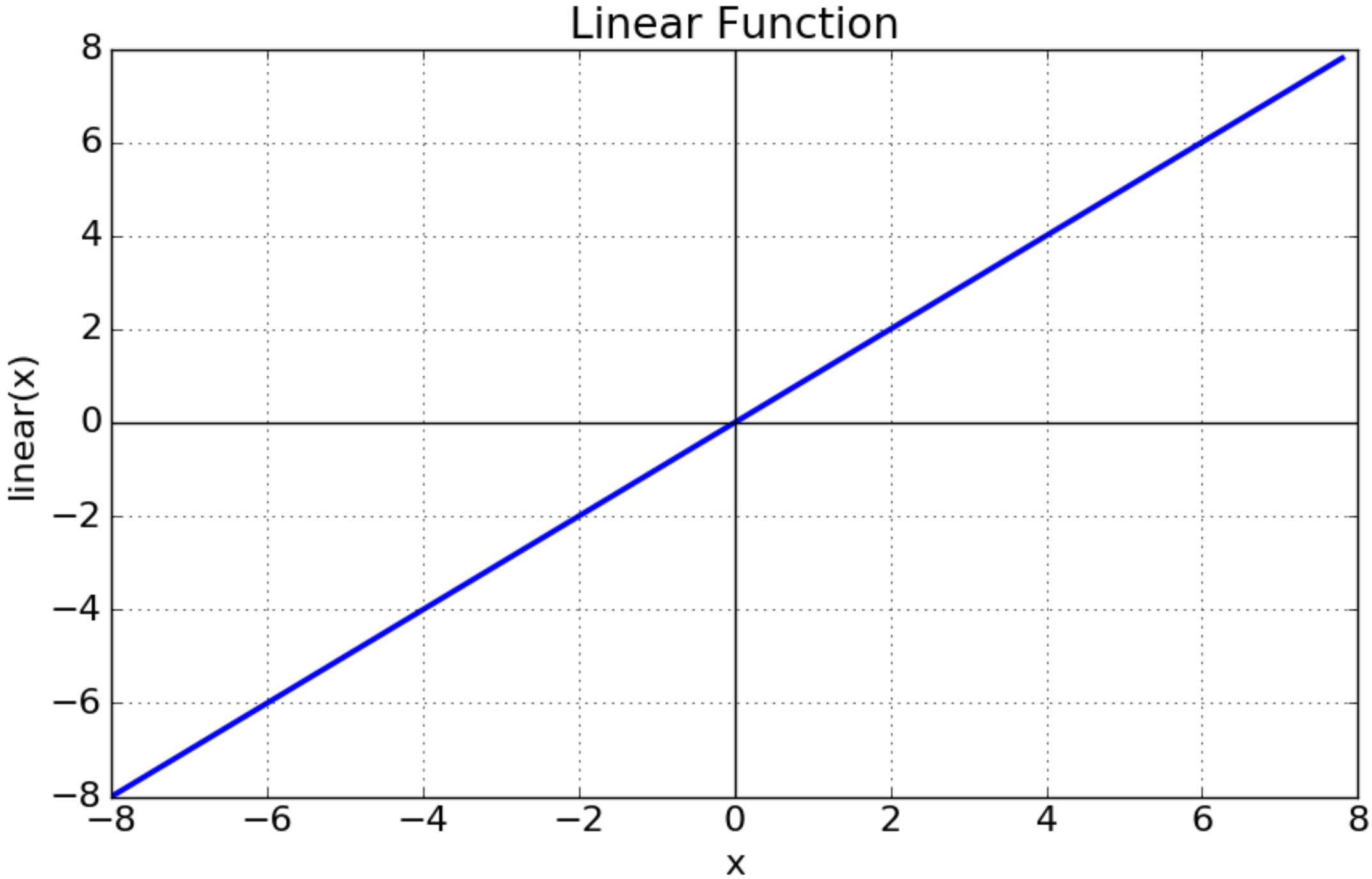


Image Credit:
https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6

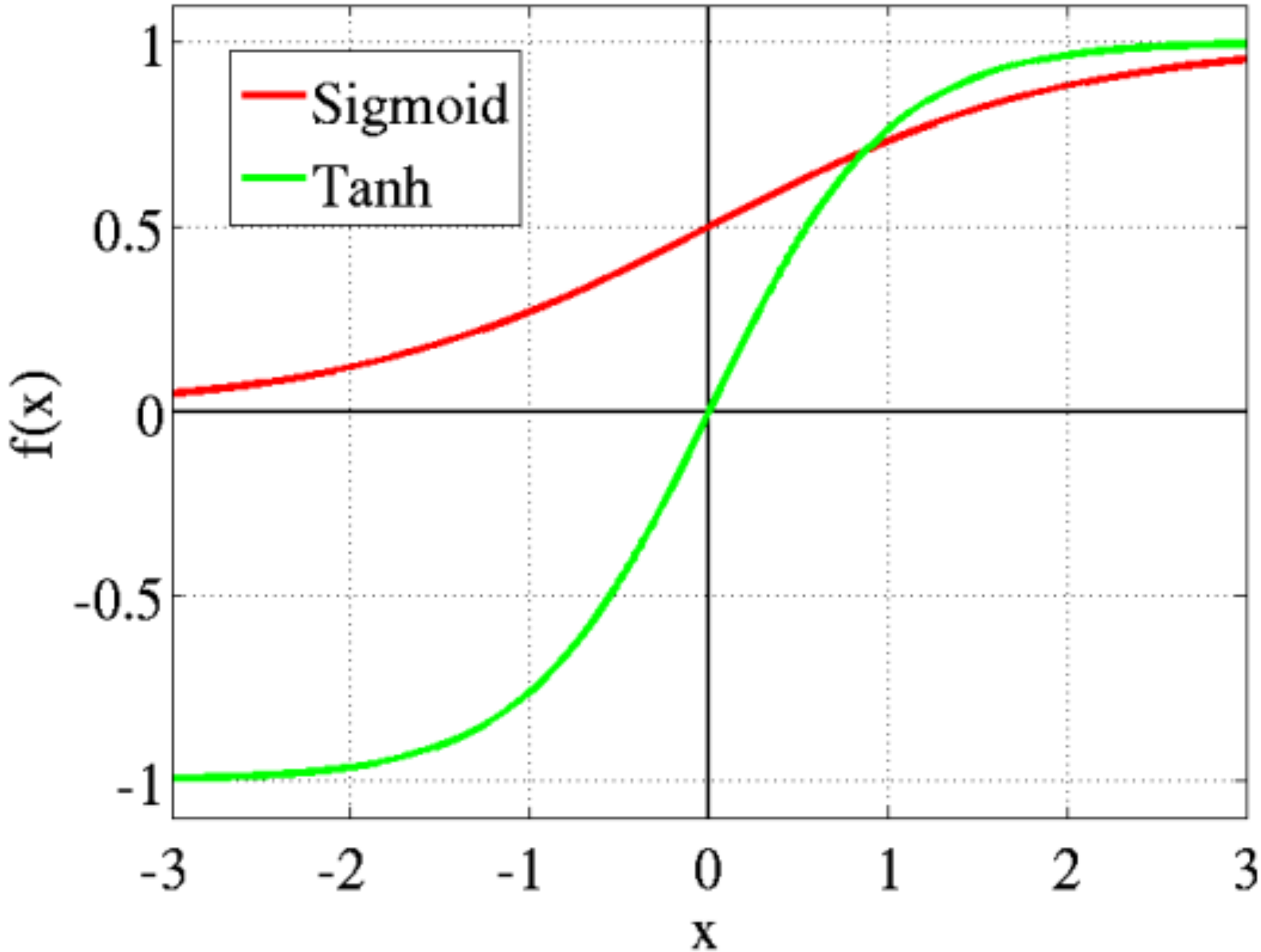


Image Credit:
https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6

$$\sigma : \mathbb{R}^K \to [0,1]^K$$

$$\sigma(\mathbf{z})_j = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}} \quad \text{for } j = 1, \ldots, K.$$

$$\sigma : \mathbb{R}^K \to [0,1]^K$$

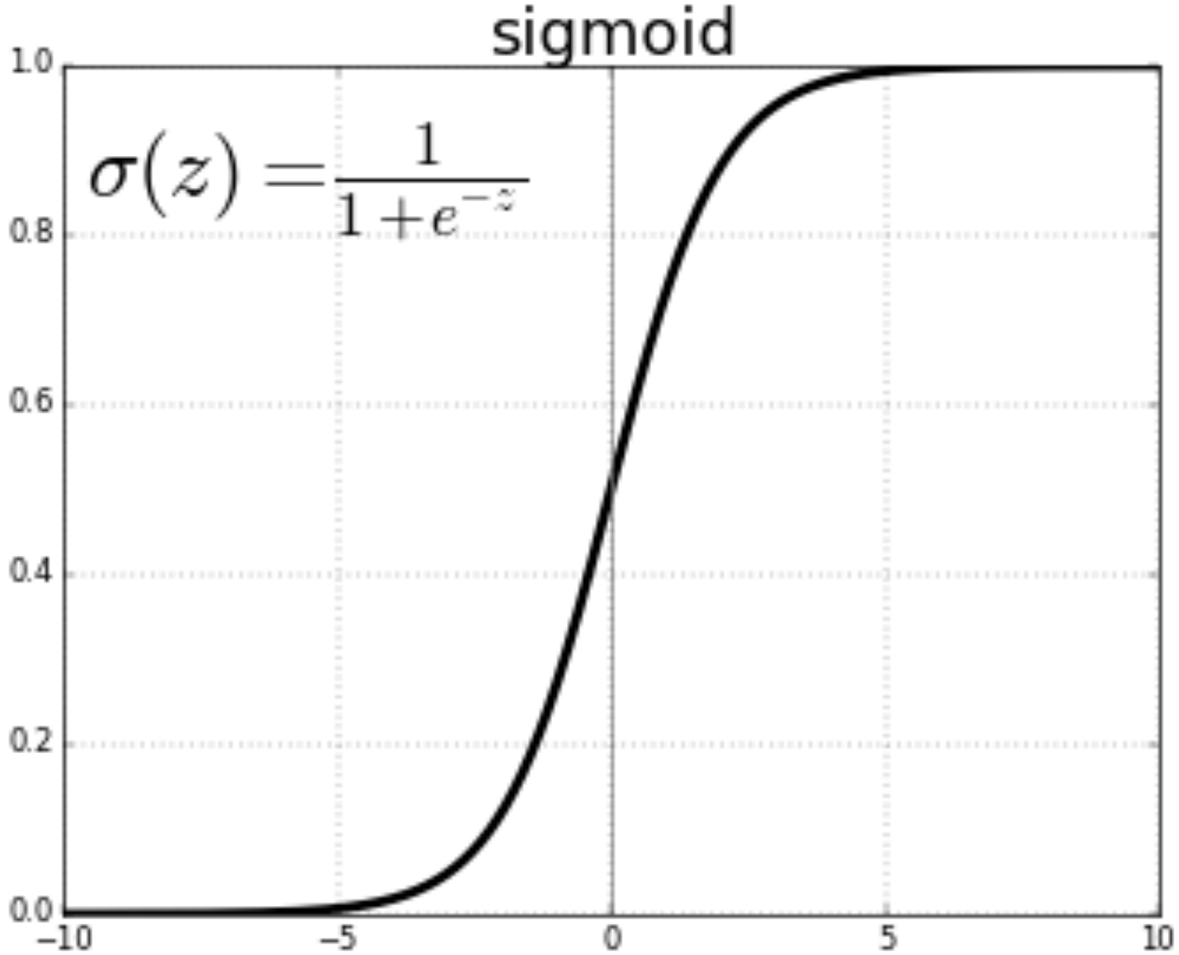
$$\sigma(\mathbf{z})_j = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}} \quad \text{for } j = 1, \ldots, K.$$

$$\sigma : \mathbb{R}^K \to [0,1]^K$$

$$\sigma(\mathbf{z})_j = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}} \quad \text{for } j = 1, \ldots, K.$$

$$\sigma : \mathbb{R}^K \to [0,1]^K$$

$$\sigma(\mathbf{z})_j = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}} \quad \text{for } j = 1, \ldots, K.$$



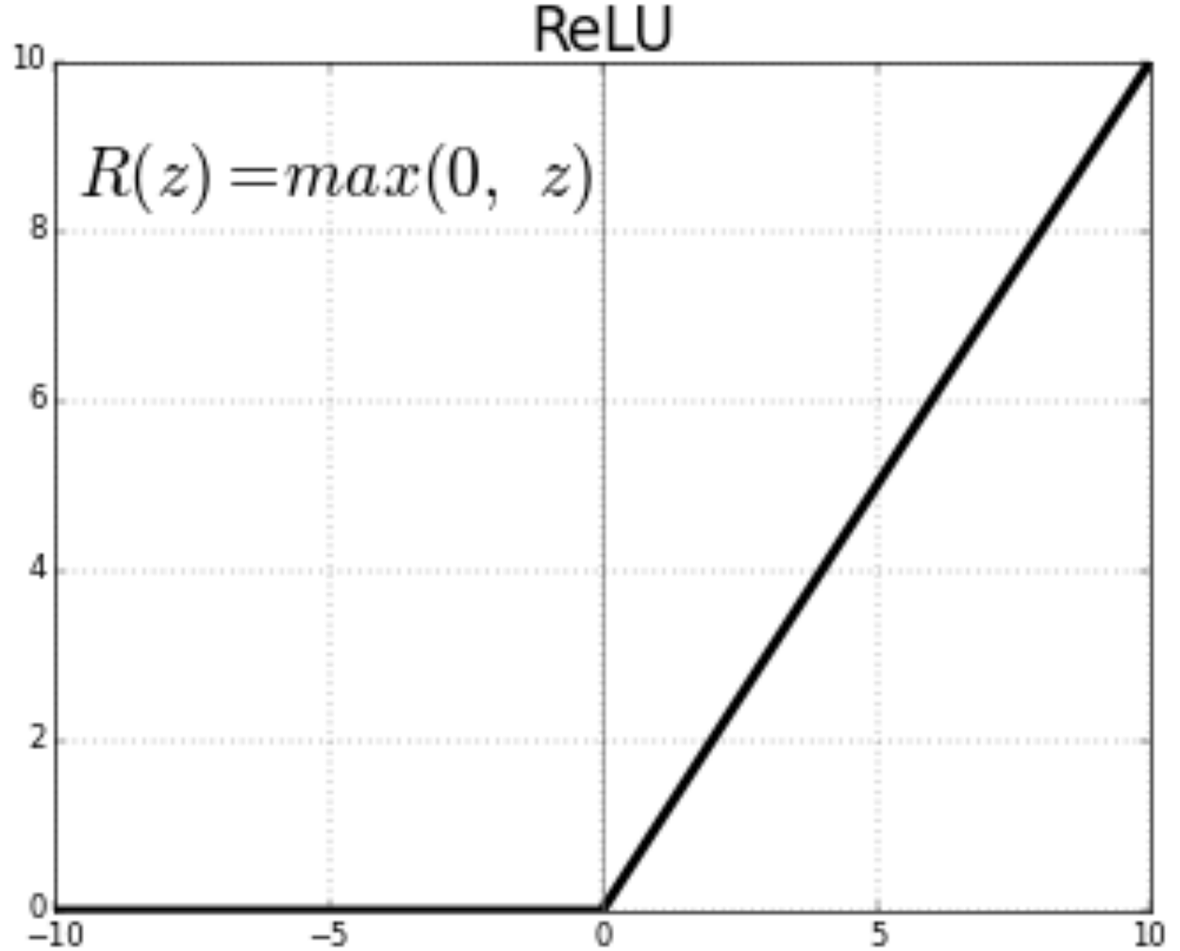
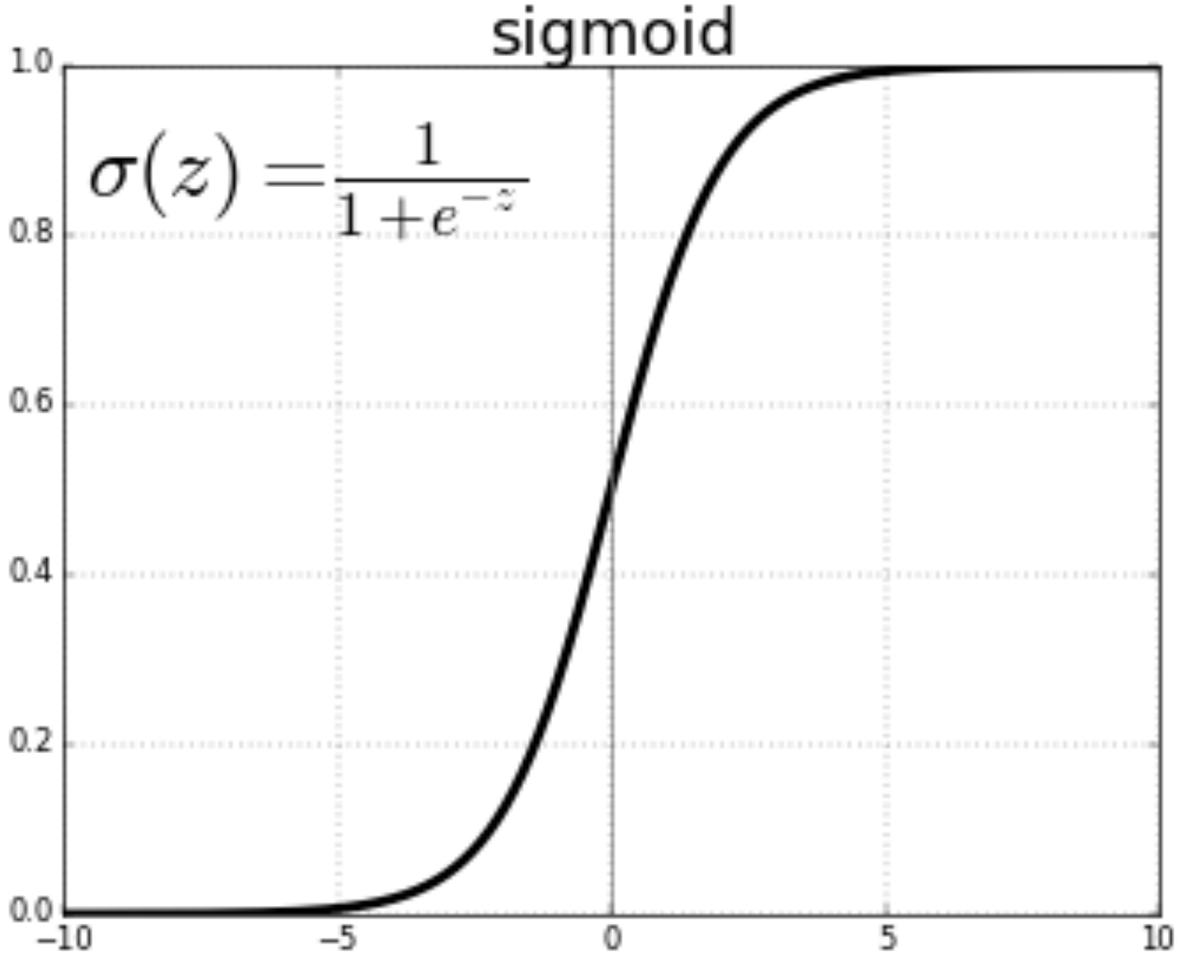


Image Credit:
https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6



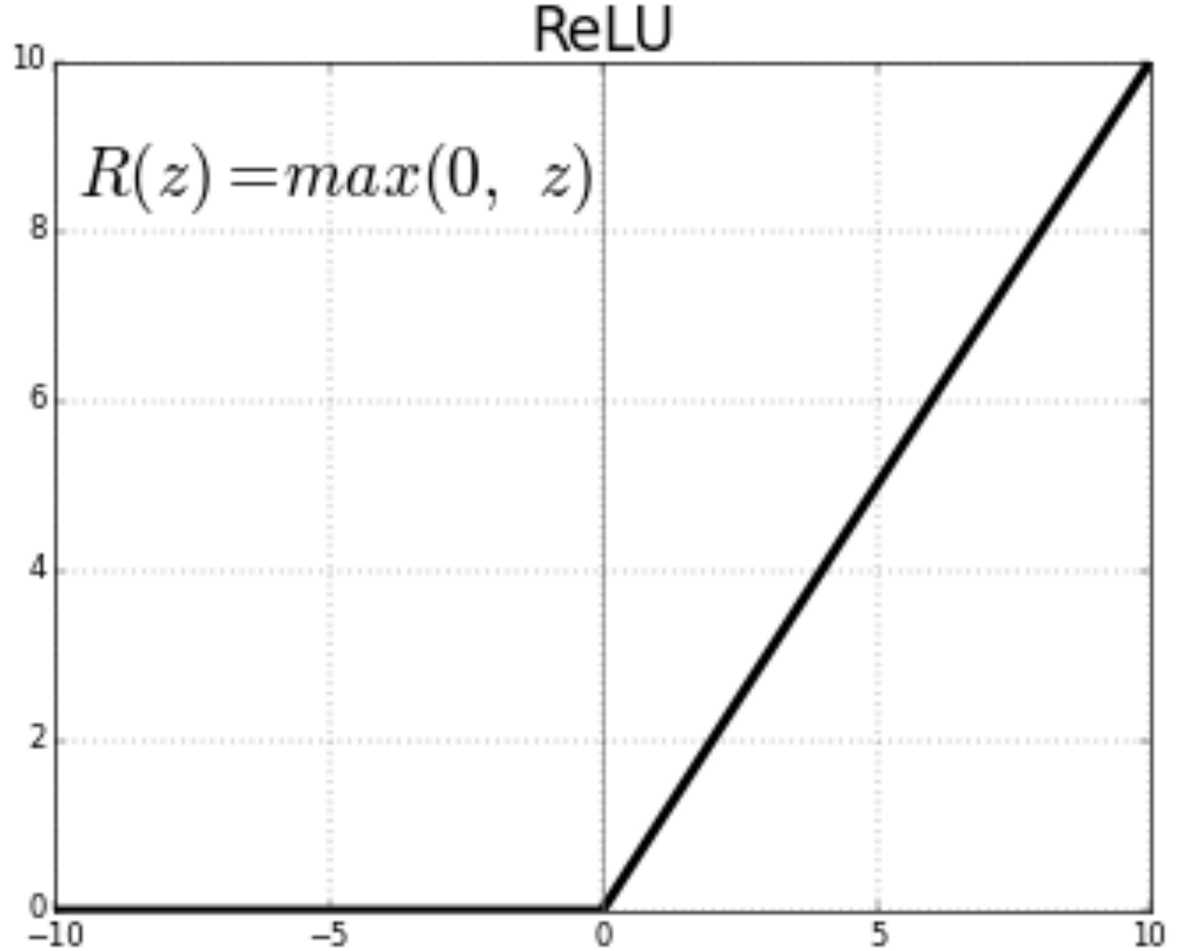
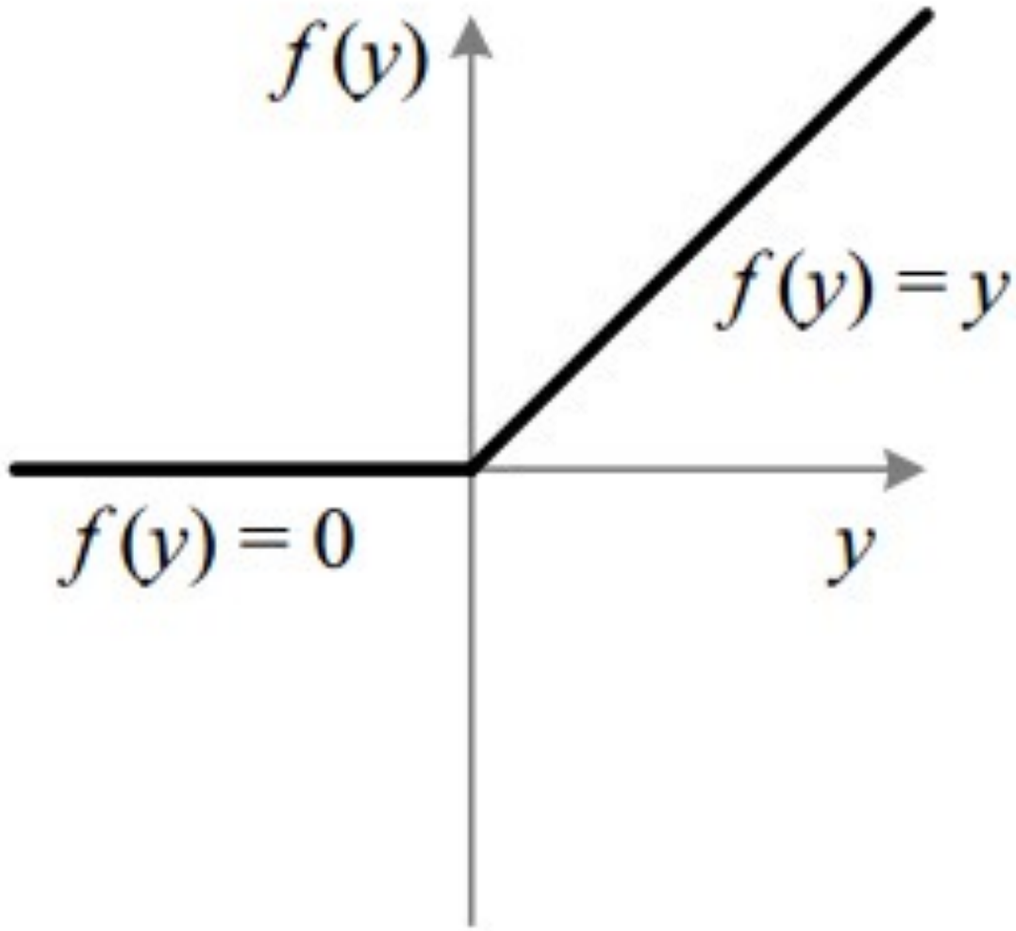


Image Credit:
https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6



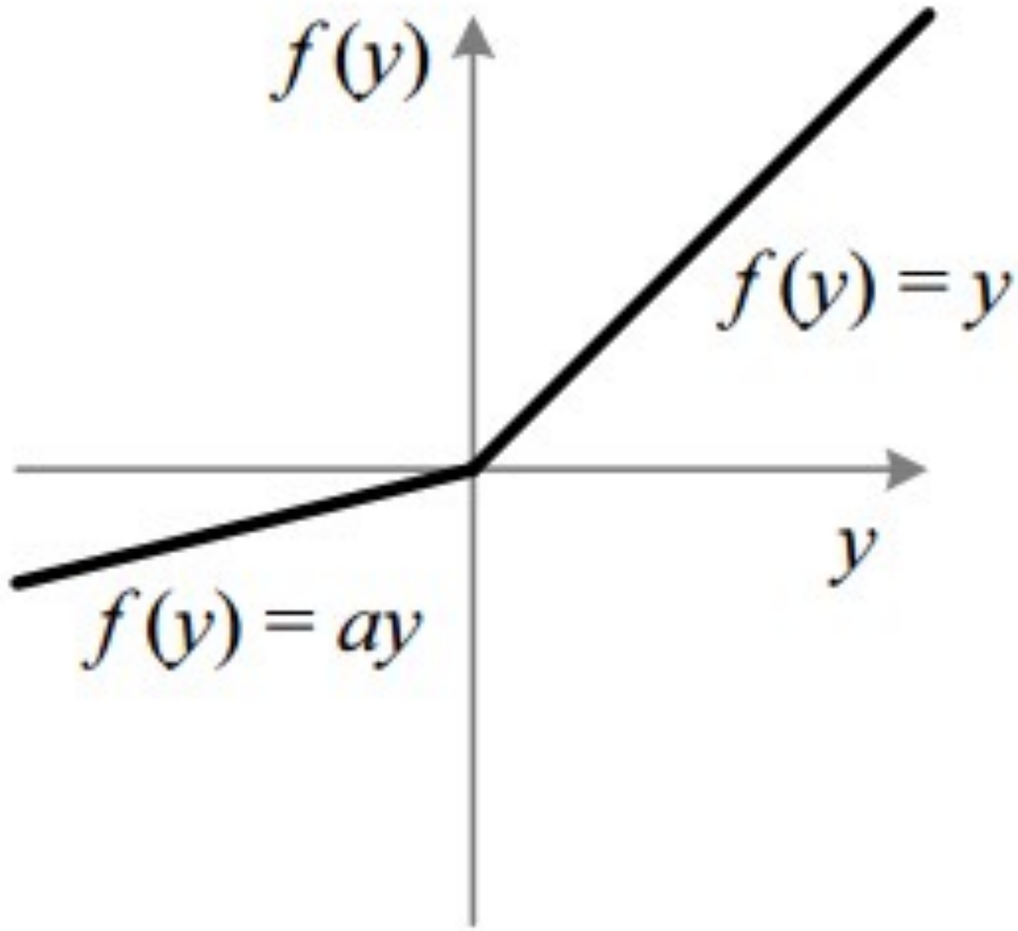


Image Credit:
https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6

Summary

Bias-Variance Tradeoff

 Romeo Kienzler