# Logistic Regression

Logistic regression is a learning algorithm used in a supervised learning problem when the output $y$ are all either zero or one. The goal of logistic regression is to minimize the error between its predictions and training data.

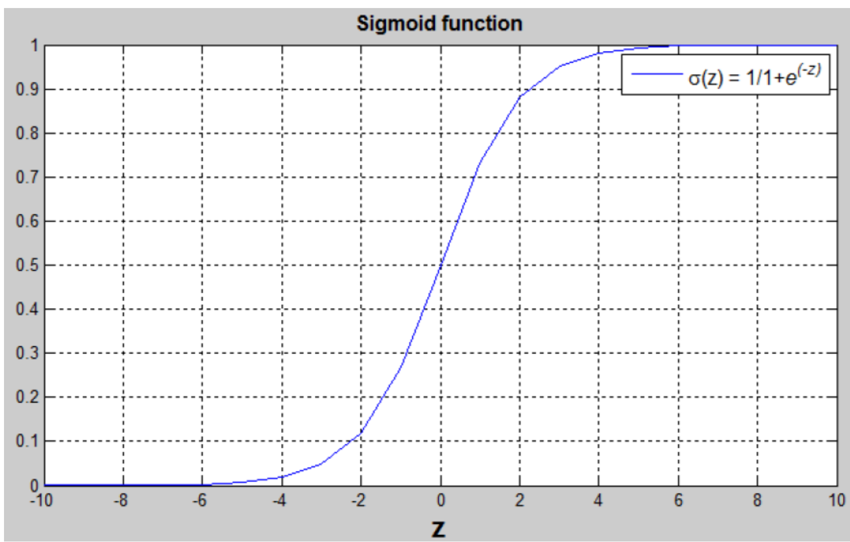
## Example: Cat vs No - cat

Given an image represented by a feature vector $x$, the algorithm will evaluate the probability of a cat being in that image.

$$Given\ x\ ,\ \hat{y} = P(y = 1|x), \text{where } 0 \ \leq \hat{y} \leq 1$$

The parameters used in Logistic regression are:

- The input features vector: $x \ \in \mathbb{R}^{n_x}$, where $n_x$ is the number of features
- The training label: $y \in 0,1$
- The weights: $w \ \in \mathbb{R}^{n_x}$, where $n_x$ is the number of features
- The threshold: $b \ \in \mathbb{R}$
- The output: $\hat{y} = \sigma(w^T x + b)$
- Sigmoid function: s = $\sigma(w^T x + b)$ = $\sigma(z)$= $\frac{1}{1+\ e^{-z}}$




$(w^T x + b)$ is a linear function ($ax + b$), but since we are looking for a probability constraint between [0,1], the sigmoid function is used. The function is bounded between [0,1] as shown in the graph above.

Some observations from the graph:

- If $z$ is a large positive number, then $\sigma(z)$ = 1
- If $z$ is small or large negative number, then $\sigma(z)$ = 0
- If $z = 0$, then $\sigma(z)$ = 0.5