Binary Classification:

In this our dependent variable (y) is a discrete variable where it takes two values i.e. $y \in \{0,1\}$ where 0 represents negative class and 1 represents a positive class.

Examples:

Email : Spam or not ?

Online Transactions : Fraud or not ?

Tumor : malignant or Benign ?

Threshold classifier output : We use linear regression and map all predictions greater than 0.5 as a 1 and all less than 0.5 as a 0. However, this method doesn't work well because classification is not actually a linear function.

What if we use linear reg for classification problem ?

If we use linear regression model then out predicted output might be <0 or sometimes it might be >1, both the cases are not suitable for classification problem.

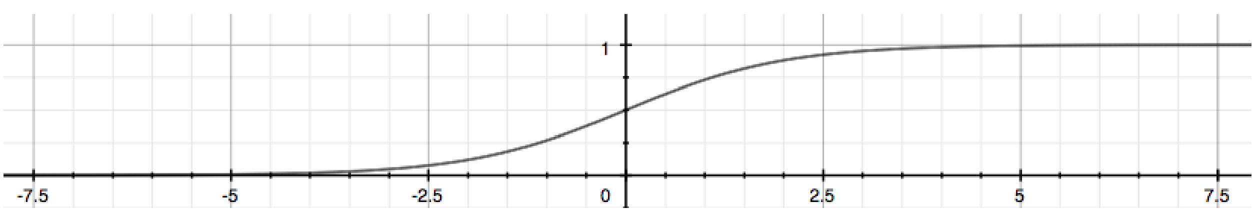
We can use classification algorithms such as **logistic Regression** for the binary classification.

Logistic Regression :

Our goal is to have our hypothesis function between 0 to 1, i.e. $0 <= h_\theta(x) <=1$.

θ

For Logistic regression we modify our hypothesis function as $h_\theta(x) = g(h_\theta(x))$, where g(z) is a logistic/sigmoid function, where $g(z) = \frac{1}{1+e^{-z}}$, the graph looks like below.



Interpretation of Hypothesis Output:

$h_\theta(x)$ = estimated probability that y=1 on input x.

Example of cancer prediction : $h_\theta(x) = 0.7$ → it tells that 70% chance of tumor being malignant tumor.

Another way of notation is $h_\theta(x) = P(y=1/x; \theta)$ → " *probability that y=1, given x, parameterized by* θ"

Also, $P(y=1/x; \theta) + P(y=0/x; \theta) = 1$.