# Linear Classification & Logistic Regression

Linear classifiers and the importance of class probabilities

For simplicity, we define a linear classifier as a binary classifier that separates two classes (positive and negative class) using a linear separator by computing a linear combination of the features and comparing against a set threshold.

Logistic Regression: Sigmoid, logit and the log-likelihood

Logistic regression is a linear algorithm that can be used for binary or multiclass classification. It is a discriminative classifier that estimates the probability that an instance belongs to a class using an s-shape function curve called the sigmoid function. The predicted values obtained after using a linear equation on the predictors by applying logistic regression can fall in the range of negative infinity to positive infinity. The sigmoid maps these results by shrinking the value to fall between 0 and 1. We can say that we use the sigmoid function to transform linear regression into logistic regression.

$$sigmoid\ \sigma(x)\ =\ \frac{1}{1\ +\ e^{-x}}$$
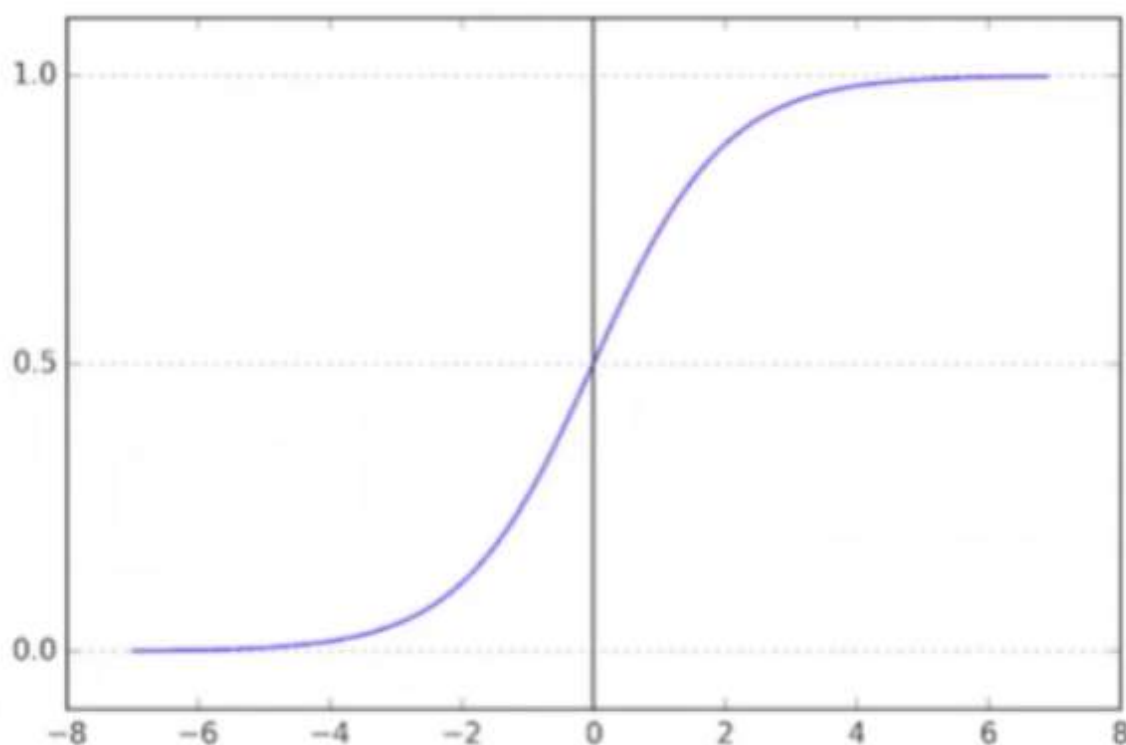
The sigmoid function can be applied to a linear equation

$$z = \beta_0 + \beta_1 x$$

to obtain values h between 0 and 1 such that:

$$h = \sigma(z) = \frac{1}{1+e^{-z}} = \frac{1}{1+e^{-(\beta_0+\beta_1 x)}}$$

For a binary classification task with classes A and B, if a threshold is set for 0.5 and the probability of an instance belonging to a class is p, we can say that if p < 0.5 the instance if of class A while it is of class B is p > 0.5.

Also known as the log of odds, logit is the logarithm of odds ratio where the odds ratio is the probability that an event occurs divided by the probability that the event does not occur. Logit is the inverse of the sigmoid such that it maps values from negative infinity to positive infinity.