Logistic Regression

- Supervised Learning
- Classification
- Binary
- Multi-nominal

An activation function is a function that is added into an artificial neural network in order to help the network learn complex patterns in the data.

Sigmoid: The sigmoid activation function is also called the logistic function. It is the same function used in the logistic regression classification algorithm. The function takes any real value as input and outputs values in range 0 to 1.

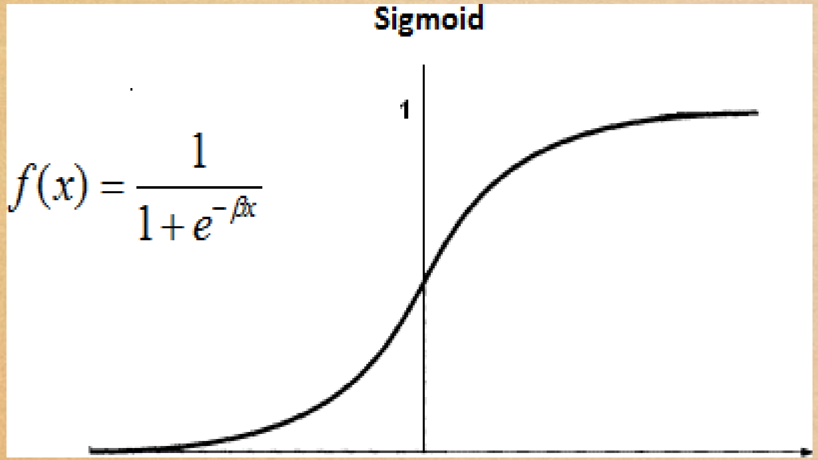
Activation Functions

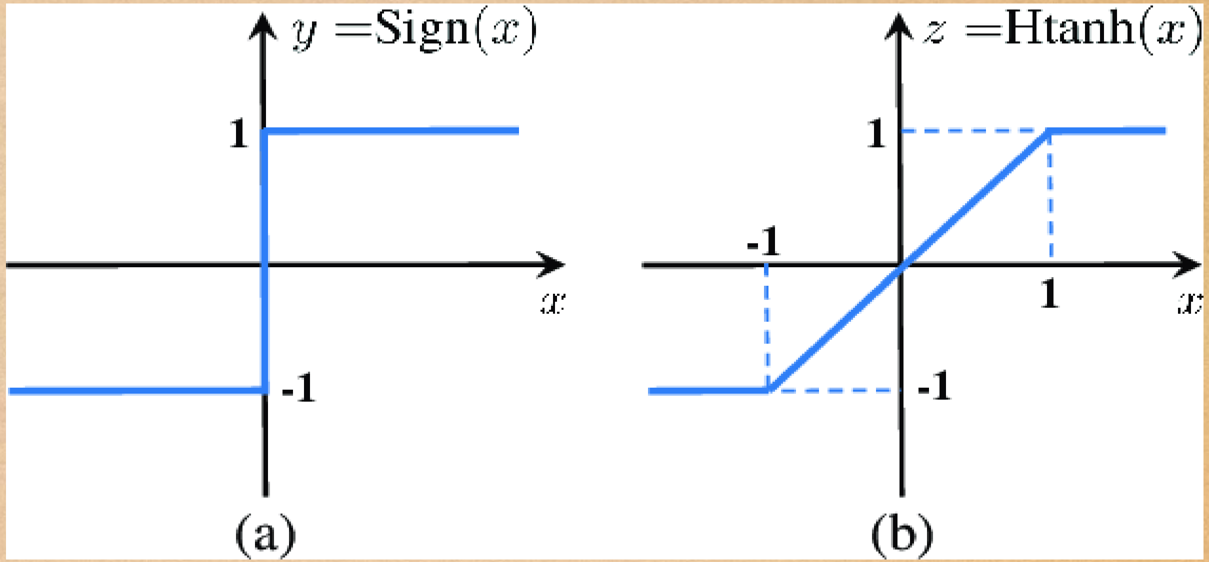
Sigmoid vs Softmax

- The sigmoid function is used for the two-class logistic regression, whereas the softmax function is used for the multiclass logistic regression.

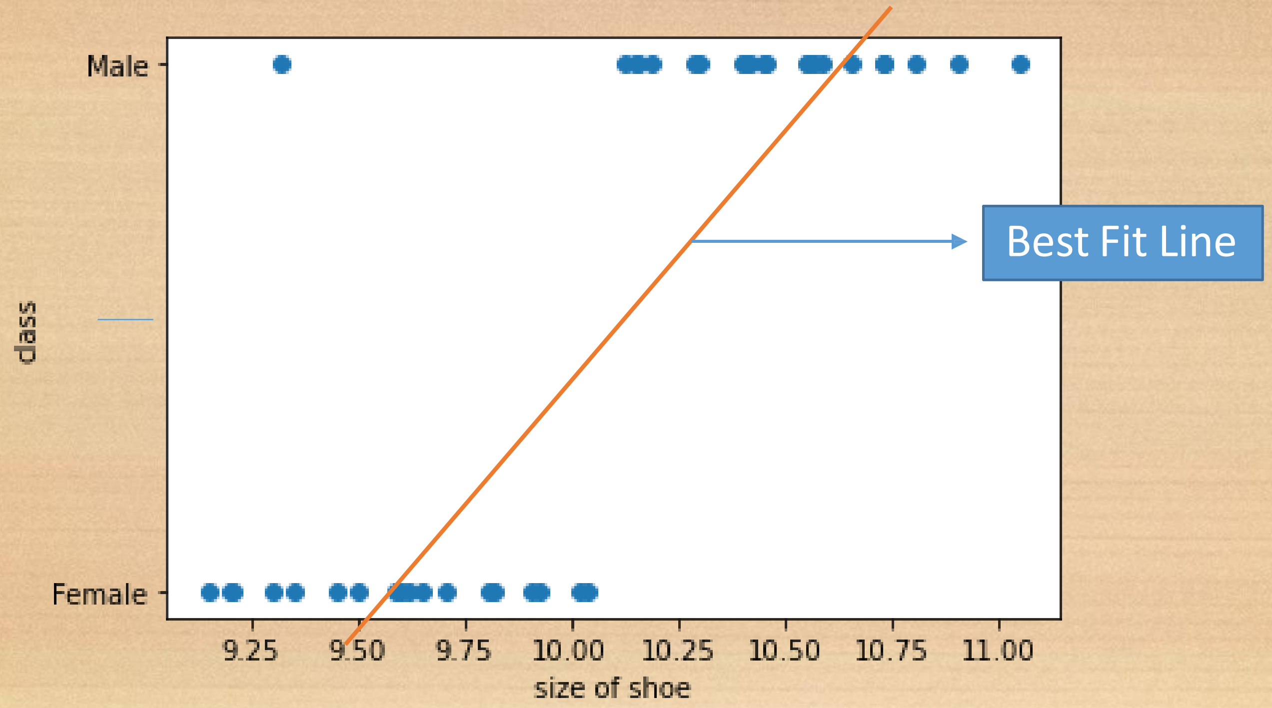
- The main advantage of using Softmax is the output probabilities range. The range will 0 to 1, and the sum of all the probabilities will be equal to one. If the softmax function used for multi-classification model it returns the probabilities of each class and the target class will have the high probability.

Activation Functions





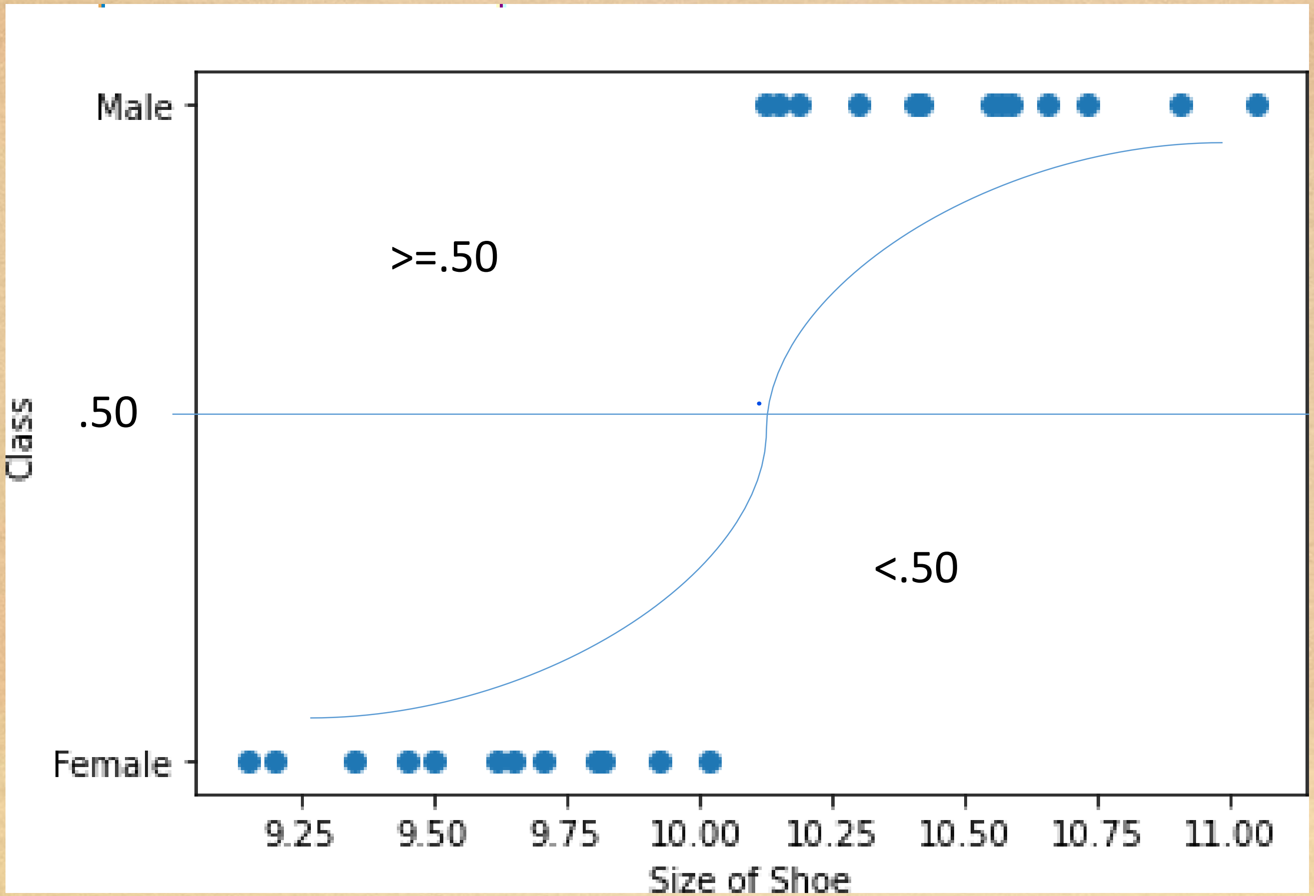
Linear Regression



$Y = MX + C$

M = Slope
C = Intercept
X = Data Point

Logistic Regression

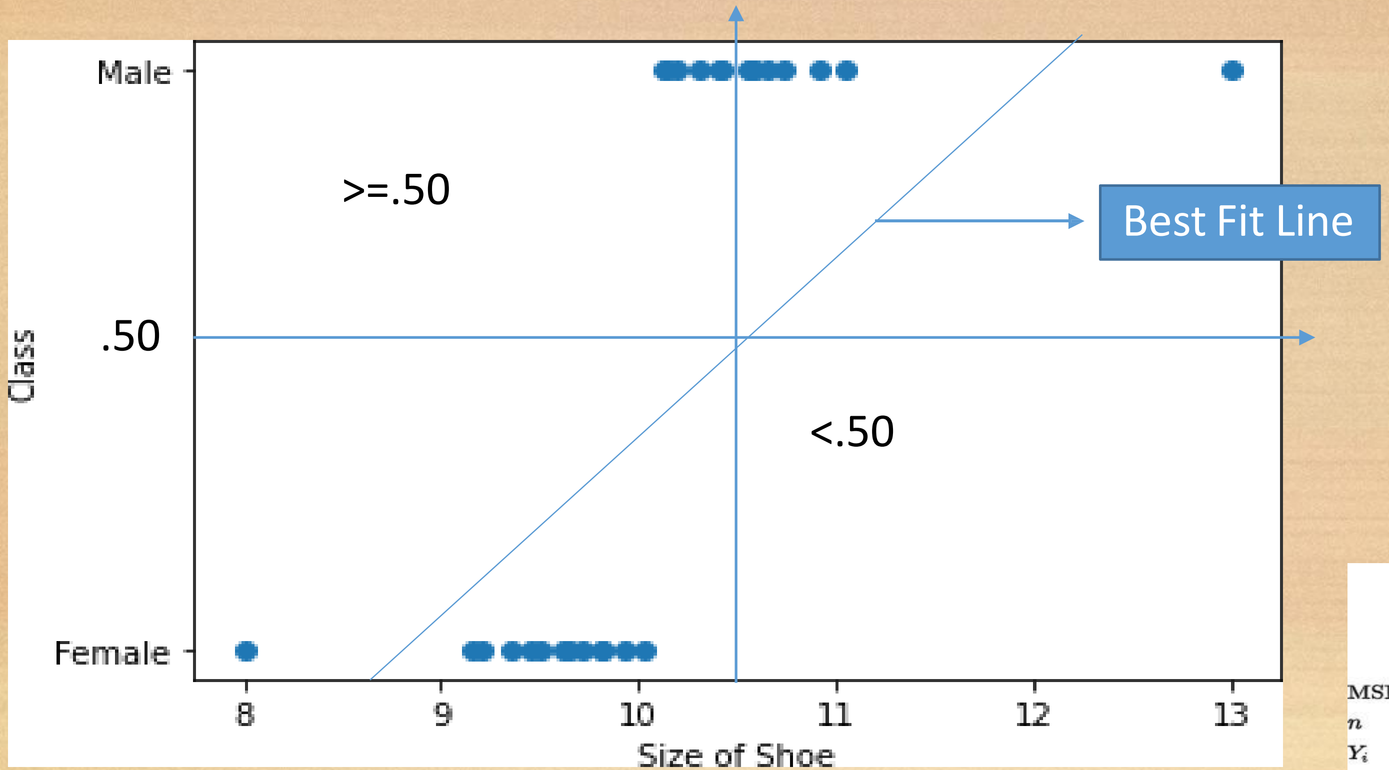


Logit(x) = MX + C

M = Slope
C = Intercept
X = Data Point

$$sigmoid_x = \frac{1}{1+e^{-x}}$$

Linear Regression



Y = MX + C

M = Slope
C = Intercept
X = Data Point

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$$

MSE = mean squared error
n = number of data points
Y_i = observed values
$\hat{Y}_i$ = predicted values

Logistic Regression

$$\boxed{\log\left(\frac{y}{1-y}\right) = mx + c}$$

1. Raising e to the power on both sides of the equation

$$\left(\frac{y}{1-y}\right) = e^{mx+c}$$

2. One divided by both sides of the equation

$$\left(\frac{1-y}{y}\right) = e^{-mx+c}$$

3. $\left(\frac{1}{y} - 1\right) = e^{-(mx+c)}$

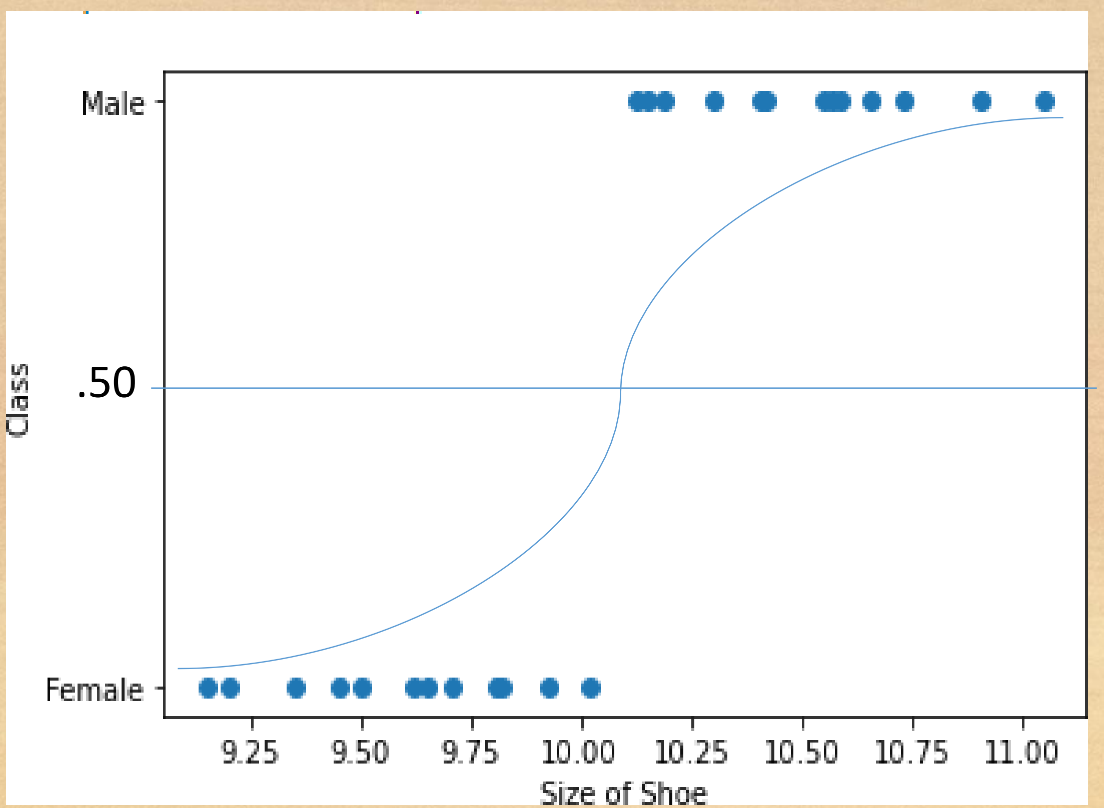
4. $\frac{1}{y} = 1 + e^{-(mx+c)}$

5. $1 = y(1 + e^{-(mx+c)})$

6. $y = \frac{1}{1+e^{-(mx+c)}}$

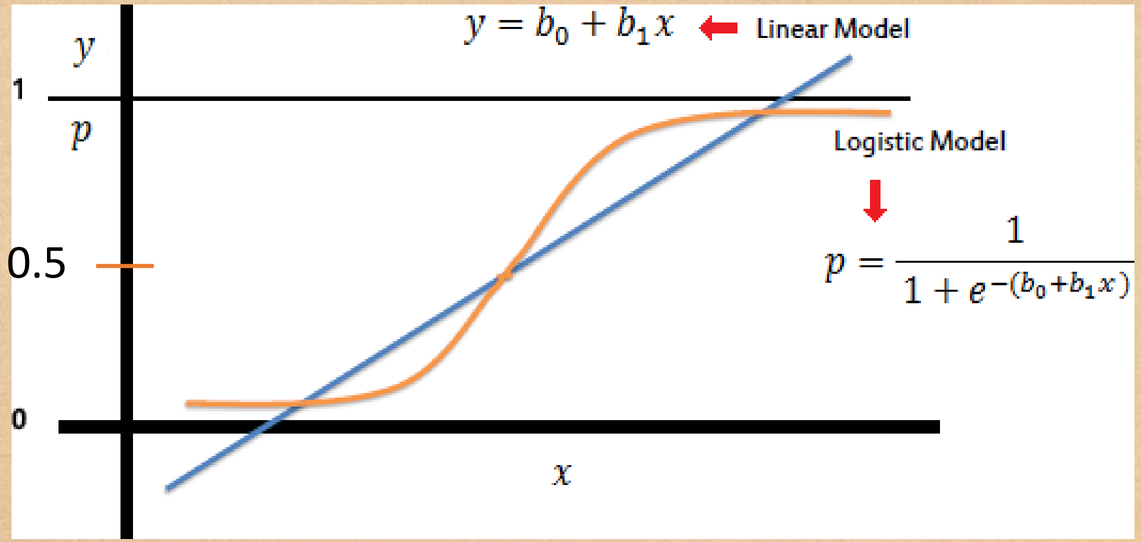
Logistic Regression

Logistic regression is a linear classifier, so you'll use a linear function $f(\mathbf{x}) = b_0 + bx$ or $f(\mathbf{x}) = b_0 + b_1x_1 + \cdots + b_rx_r$, also called the Logit. The variables b_0, b_1, …, b_r are the estimators of the regression coefficients, which are also called the predicted weights or just coefficients.



Linear vs Logistic Regression

$$y = mx + c$$



Abu Bakar Siddique Mahi