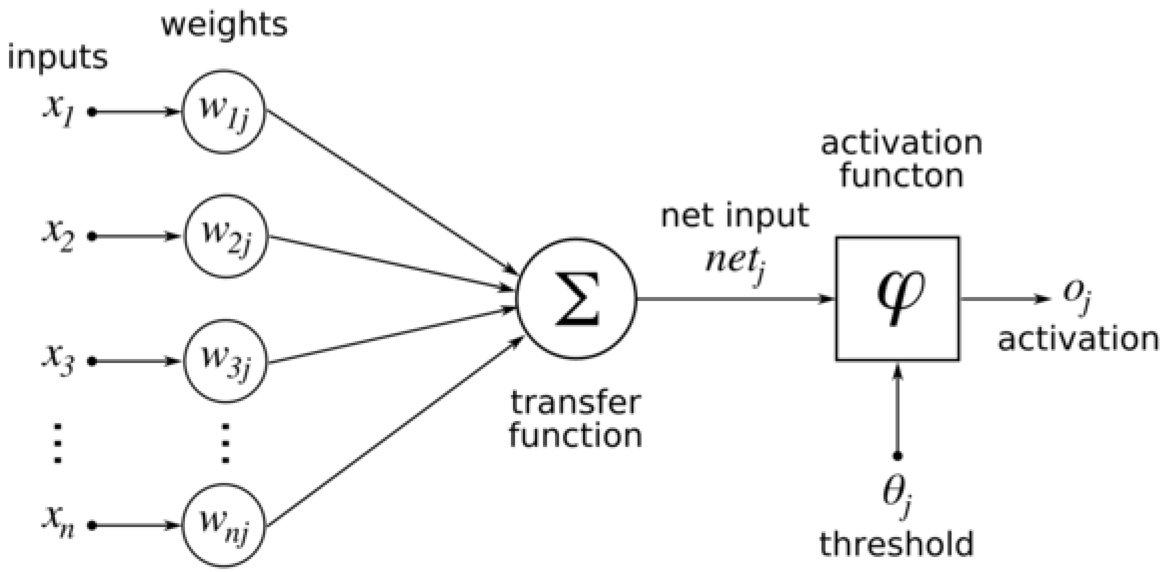
Logistic Regression

- Supervised Learning
- Classification
- Binary
- Multi-nominal

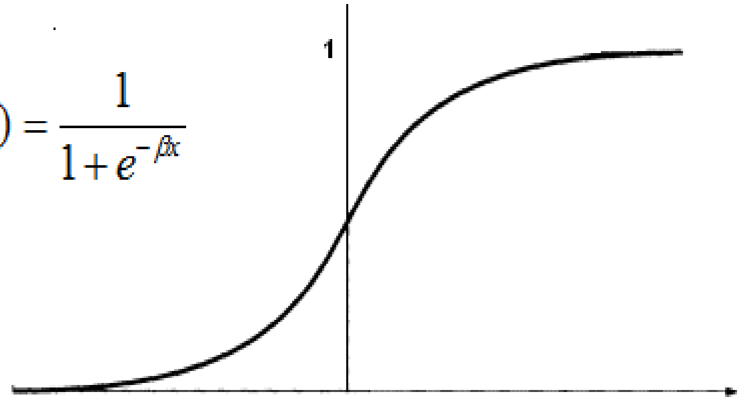
Neural Networks

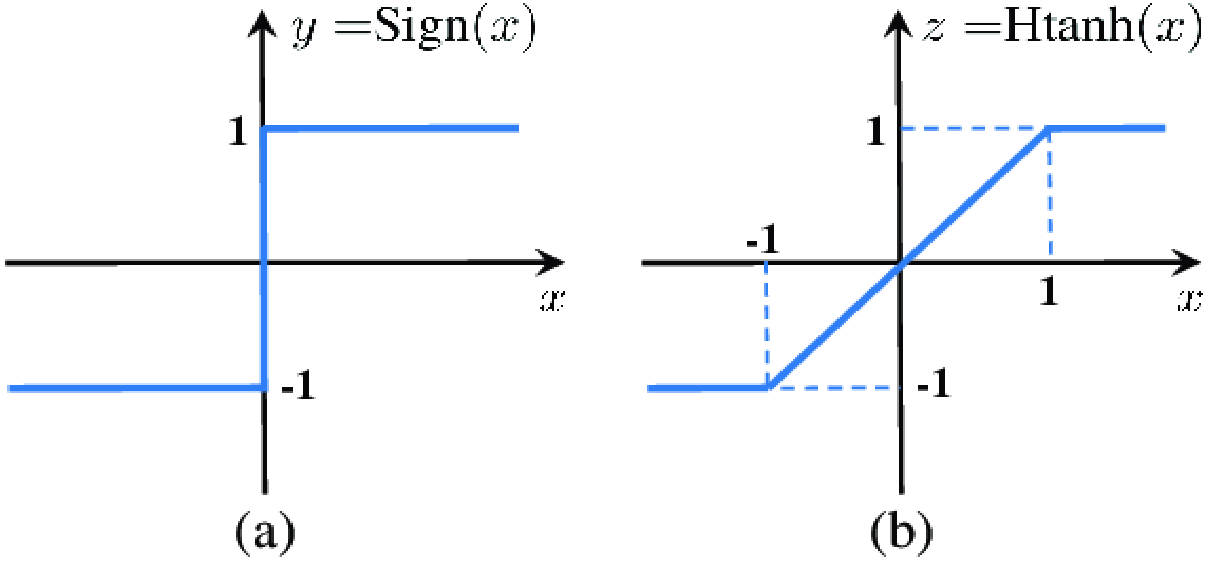


Activation Functions

Sigmoid

$$f(x) = \frac{1}{1+e^{-\beta x}}$$





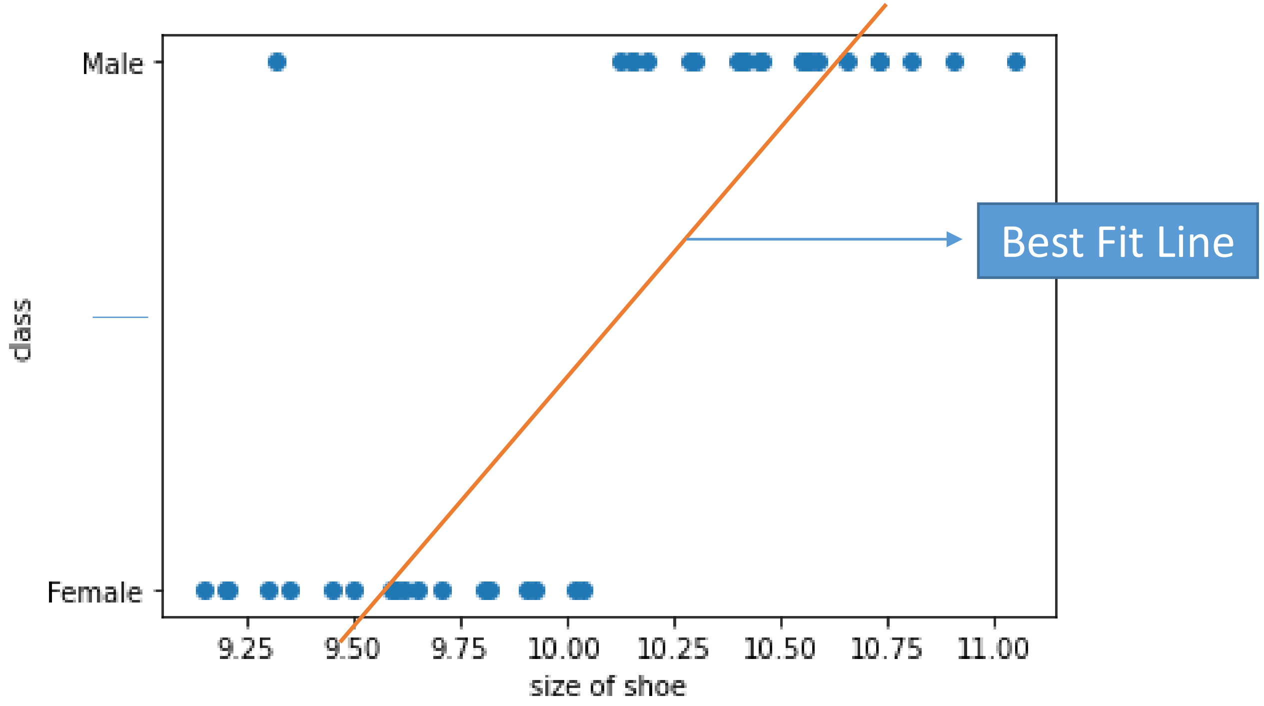
Activation Functions

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.

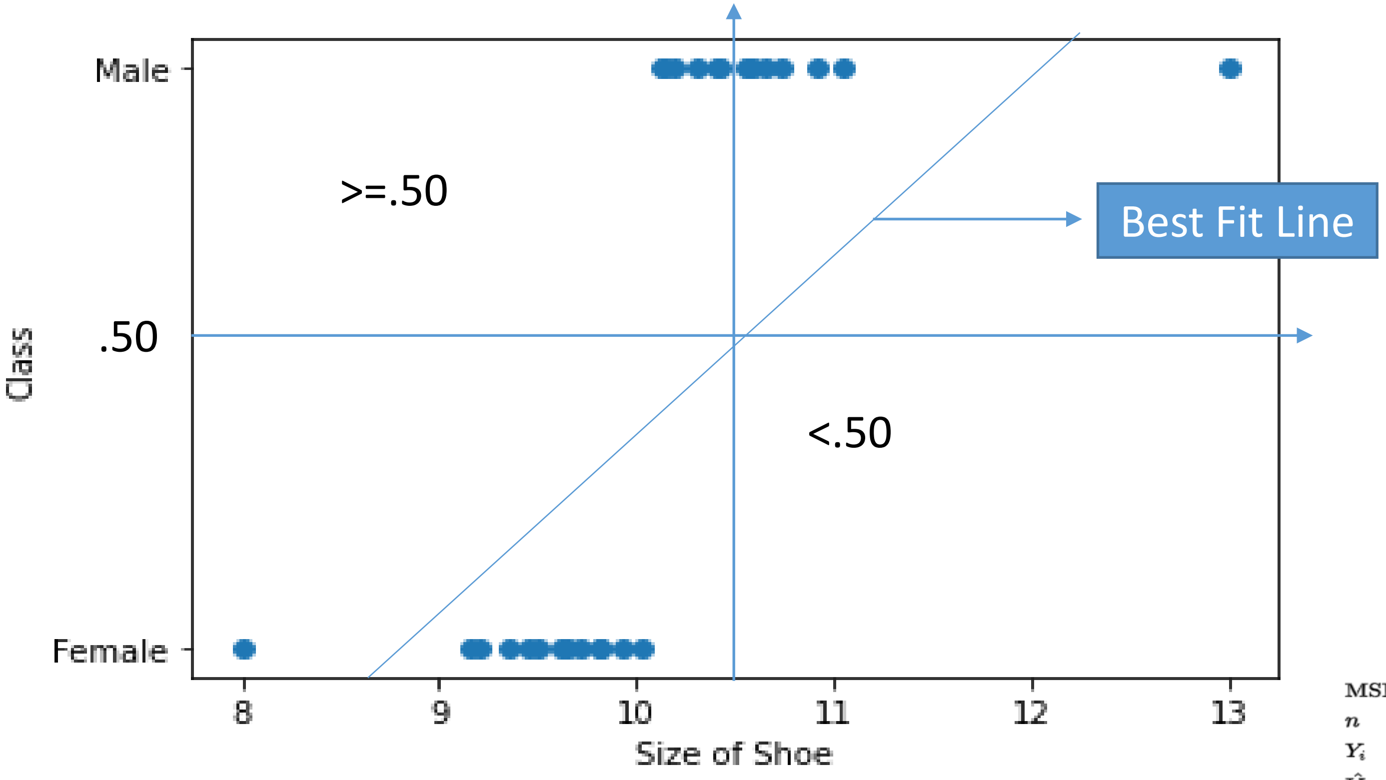
Linear Regression



Y = MX + C

M = Slope
C = Intercept
X = Data Point

Linear Regression



Y = MX + C

M = Slope
C = Intercept
X = Data Point

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

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

Logistic Regression

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

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

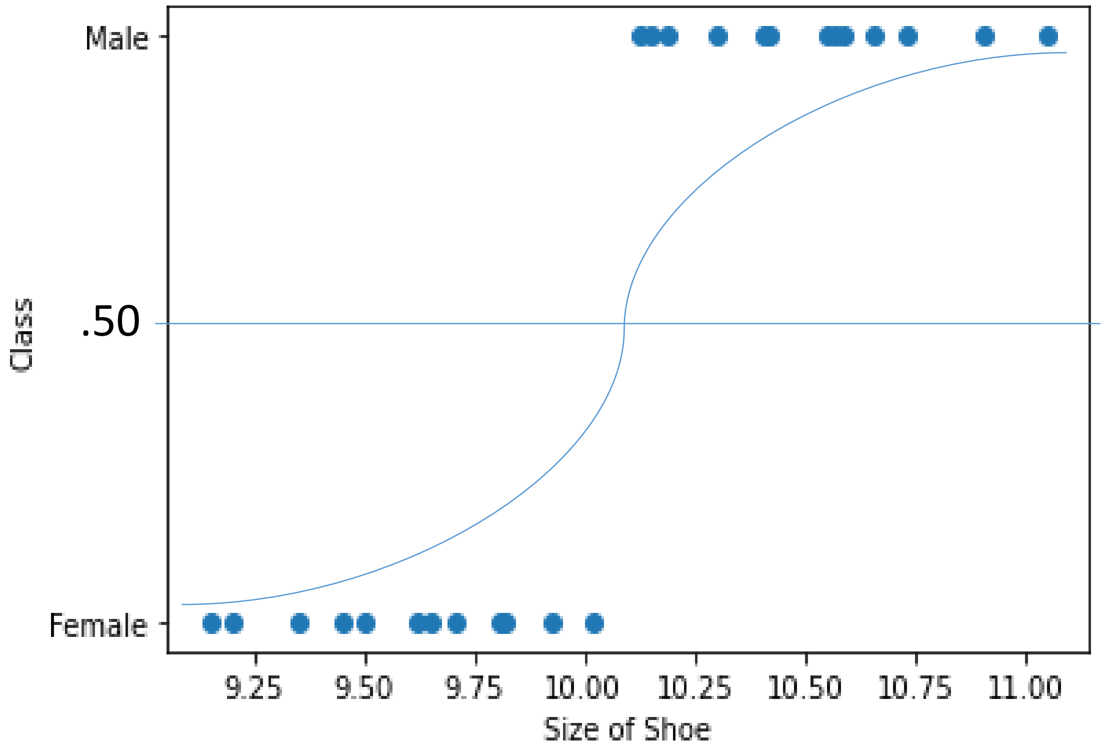
4. $\frac{1}{y} = 1 + e^{-(mx+c)}$

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

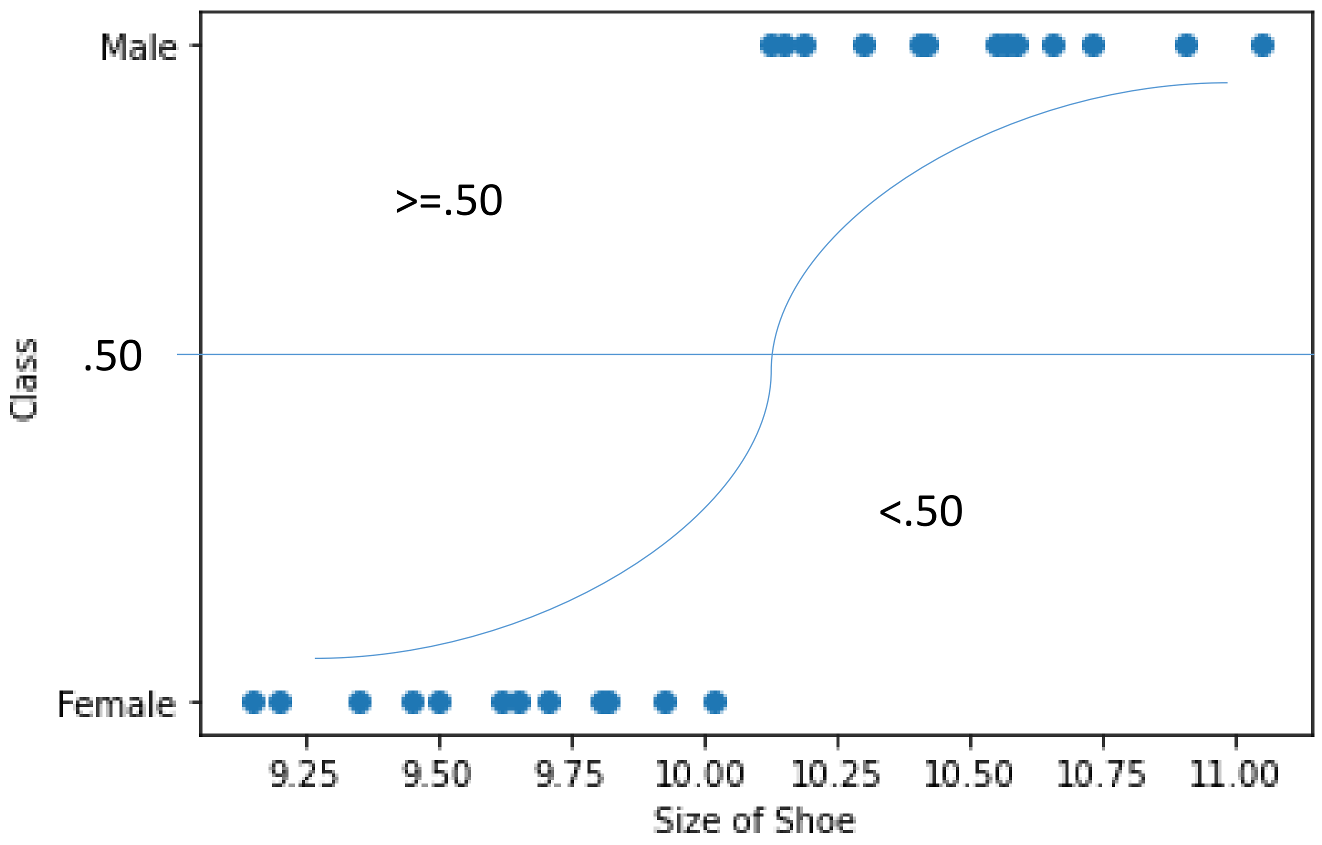
6. $y = \frac{1}{1+e^{-(mx+c)}}$

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 $b_0, b_1, \ldots, b_r$ are the estimators of the regression coefficients, which are also called the predicted weights or just coefficients.



Logistic Regression



Logit(x) = MX + C

M = Slope
C = Intercept
X = Data Point

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

Linear vs Logistic Regression