Multiple Linear Regression: -

$$y = B_0 + B_1 x_1 + B_2 x_2 + B_3 x_3$$

Y = dependent variable

B_0 = Constant

$B_{1,}, B_2, B_3$ = coefficients or weights

$X_{1,}, X_2, X_3$ = Independent features

$$y = B_0 + B_1 x_1 + B_2 x_2 + B_3 x_3$$

Salary = $B_0 + B_1 * \text{years of experience} + B_2 * \text{position} + B_3$ * area of expertise

But choosing the required features is one of the key feature that needs to be followed while building a model

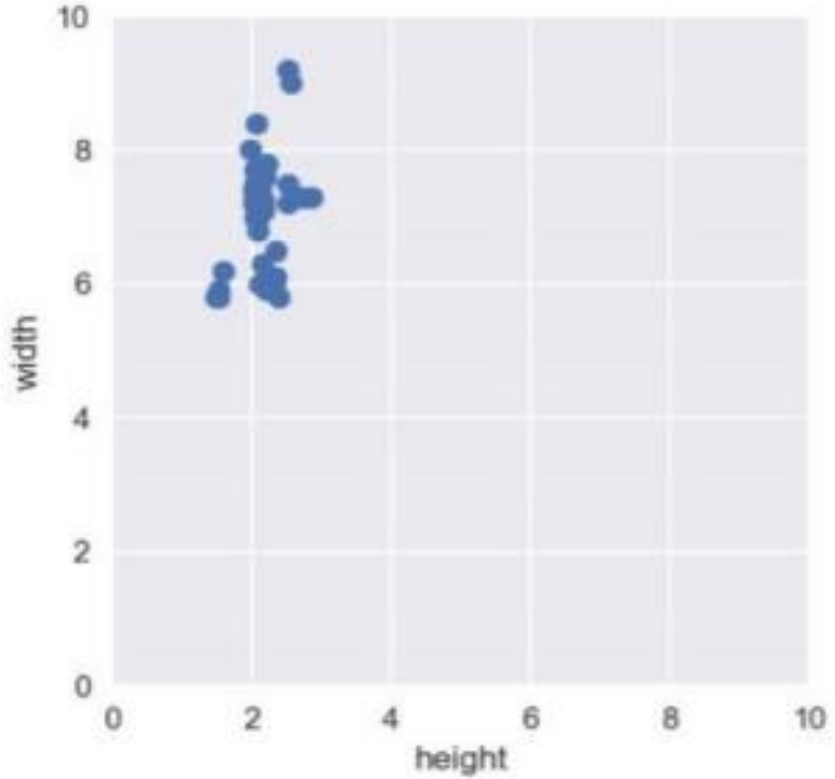
Ridge Regression: -

- Ridge regression adds a regularisation parameter in order to reduce the penalty while having large variables

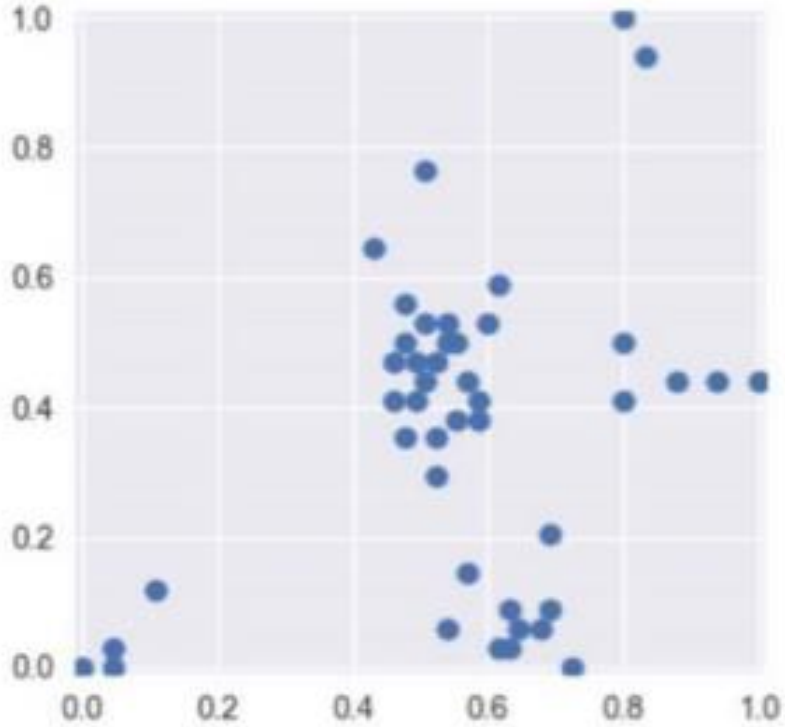
$$RIDGE = \xi(y - y_i)^2 + \alpha\xi B_i^2$$

- Ridge regression uses a L2 regularization which minimizes the sum of squares of "***B***" entities.
- The regularization is controlled by using a alpha α term
- Higher alpha means more regularization and simpler models

Feature Normalization with MinMaxScaler



Unnormalized data points



Normalized with MinMaxScaler

Lasso Regression: -

• Ridge regression adds a regularisation parameter in order to reduce the penalty while having large variables

$$LASSO = \xi(y - y_i)^2 + \alpha\xi \, |B_i|$$

• Lasso regression uses a L1 regularization which minimizes the sum of absolute values of the coefficients.

• Lasso Regression has the effect of setting parameter coefficients/ weights to zero for least influenced variables. This is called sparse solution: a kind of feature selection

• The regularization is controlled by using a alpha α term

• Higher alpha means more regularization and simpler models