**Assignment 7**

1. How sci-kit learn module performs feature scaling?

Ans:

Sci-Kit learn perfoms feature scaling using two ways, by Normalization using MinMaxScaler. And by Standardization using StandardScaler.

MinMaxScaler formula,

Xscaled = Xi – Xmax/Xmax - Xmin

StandardScaler formula,

Xscaled = Xi – mean /Std deviation

1. If you are performing feature scaling, what you will prefer Standardisation or Normalization?

Ans:

Standardization, as it is more robust to outliers. Also in Standardization there is no bounding range unlike Normalization range is between 0 and 1. This means normalization has some effect with outliers.

1. What is the difference between the residual sum of squares and regularization?

Ans:

Residual sum of squares = summation (Ypred –Yactuals)^2

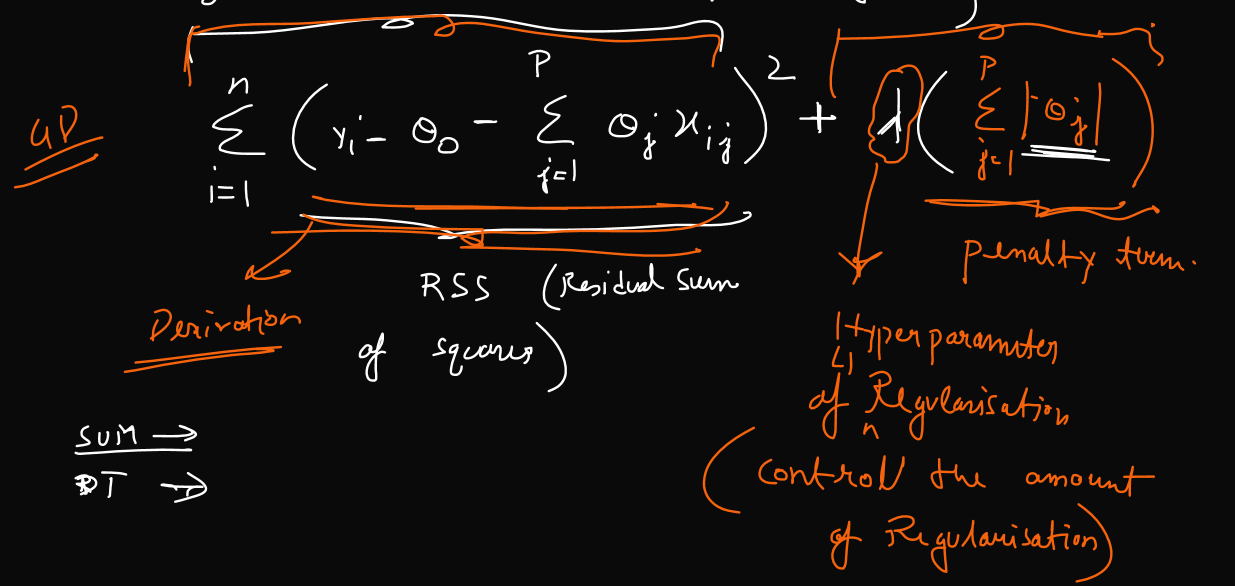
RSS is the squared sum of difference between predictions and actuals. Used for a Model evaluation metric for Regression problems.

Regularization is method of controlling/constraining the model with additional parameter. This is used to avoid overfitting of the model. Regularization also used for reducing model complexity, penalizing the loss function.

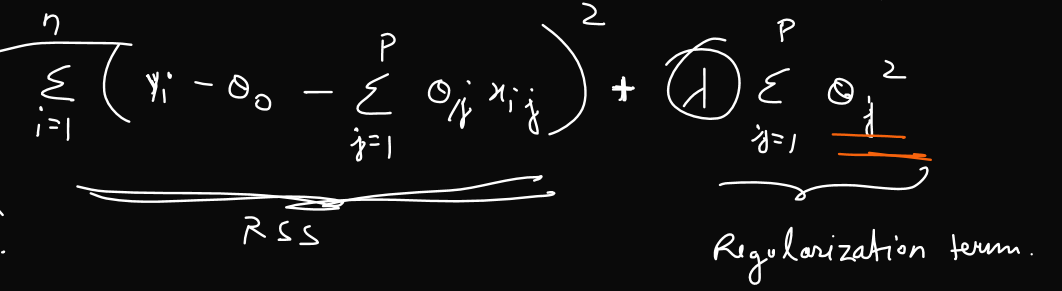
1. Differentiate between L1, L2, and Elastic Net regularization on the basis of the alpha parameter in scikit learn?

Ans:

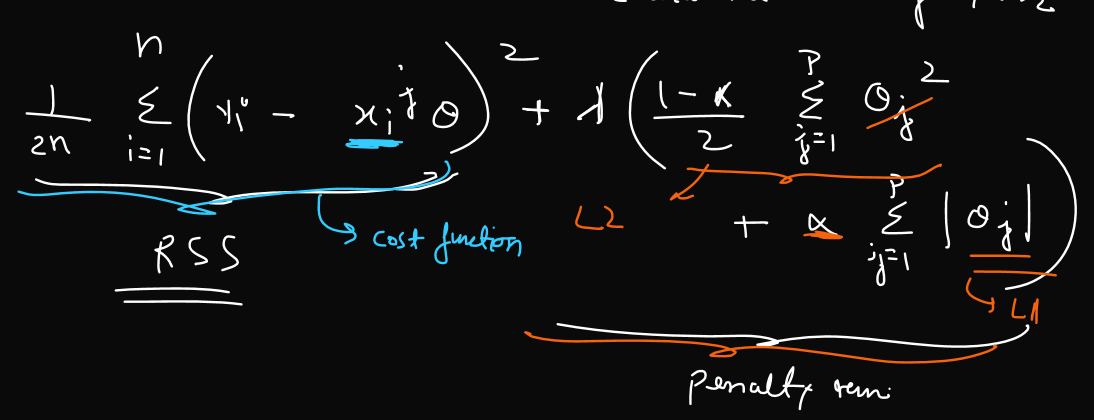
L1 Regularization also known as Lasso, adds penalty equal to absolute value of the magnitude of coefficients. It can yield in sparse models by setting some features coefficients to zero. Also used for Feature Selection technique.



L2 Regularization also known as Ridge, adds penalty equal to square of magnitude of coefficients. This will result in dense model, meaning features will not be set to zero.



ElasticNet is the combination of L1 and L2 with the addition of new parameter alpha. Where Alpth is the ratio of L1 and L2.



From above it can be seen that

if **alpha = 0, it will give a Ridge or L2 Regularization.**

if **Alpha = 1, it will give Lasso or L1 regularization.**

1. Suppose overfitting is happening, what kind of regularization you will prefer to perform?

Ans:

L1 regularization is preferable as it sets least important features to zero. Whereas L2 will still give a very small non-zero coefficients thereby making it more complez.