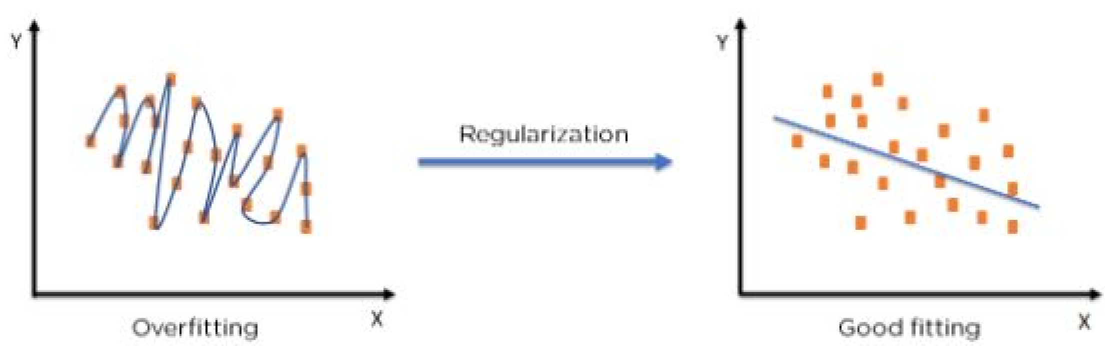
MACHINE LEARNING

1) A
2) A
3) B
4) B
5) A
6) B
7) D
8) A
9) A
10) B
11) B
12) A,B,C

Subjective Questions

13) Explain the term regularization?

Regularization refers to techniques that are used to calibrate machine learning models in order to minimize the adjusted loss function and prevent overfitting or underfitting.



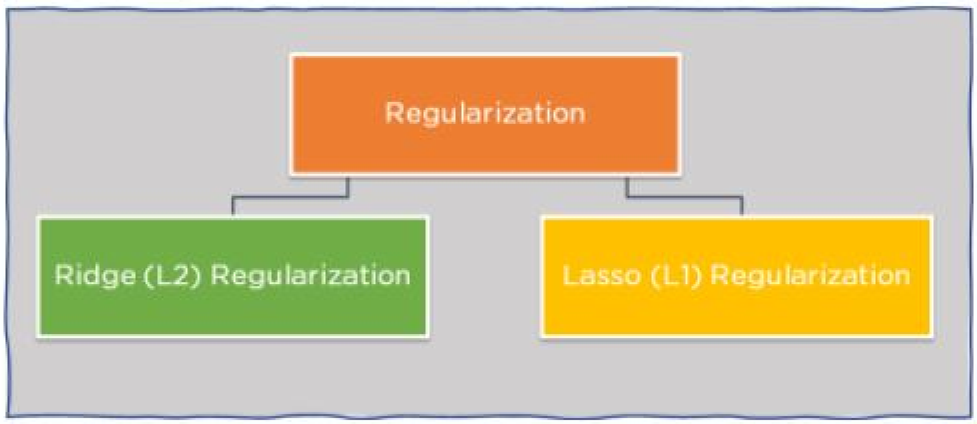
Regularization on an over-fitted model

Using Regularization, we can fit our machine learning model appropriately on a given test set and hence reduce the errors in it. Regularization is "any modification we make to a learning algorithm that is intended to reduce its generalization error but not its training error."

If a notwork performs perfectly in the training dataset it cannot be guranteed to perform as well in new datasets. But normally speaking what the network encounters is unseen data when we use it. The most staightforward strategy is just increase the dataset but it is not realistic most of the time. The modification is mostly some mathematical improvements, for instance a more proper loss function, early stopping, drop out and etc.

14) Which particular algorithms are used for regularization?

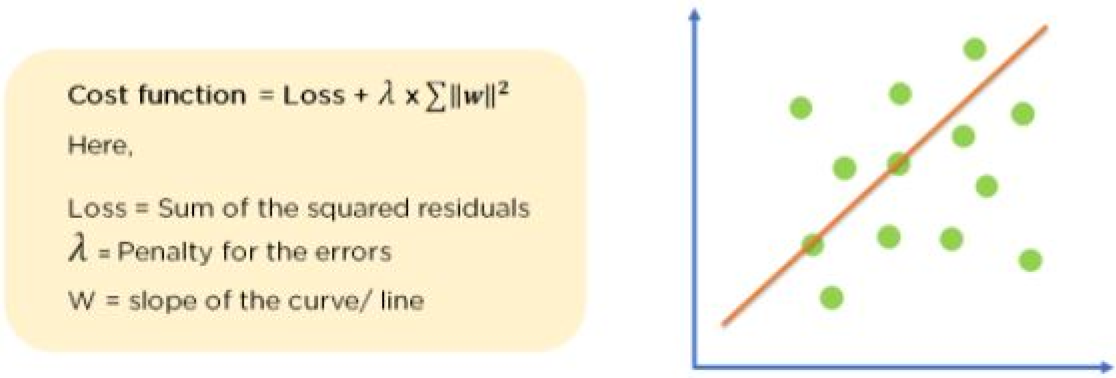
There are two main types of regularization techniques: Ridge Regularization and Lasso Regularization.



Ridge Regularization :

Also known as Ridge Regression, it modifies the over-fitted or under fitted models by adding the penalty equivalent to the sum of the squares of the magnitude of coefficients.

This means that the mathematical function representing our machine learning model is minimized and coefficients are calculated. The magnitude of coefficients is squared and added. Ridge Regression performs regularization by shrinking the coefficients present. The function depicted below shows the cost function of ridge regression :



Cost Function of Ridge Regression

In the cost function, the penalty term is represented by Lambda λ. By changing the values of the penalty function, we are controlling the penalty term. The higher the penalty, it reduces the magnitude of coefficients. It shrinks the parameters. Therefore, it is used to prevent multicollinearity, and it reduces the model complexity by coefficient shrinkage.

Consider the graph illustrated below which represents Linear regression :

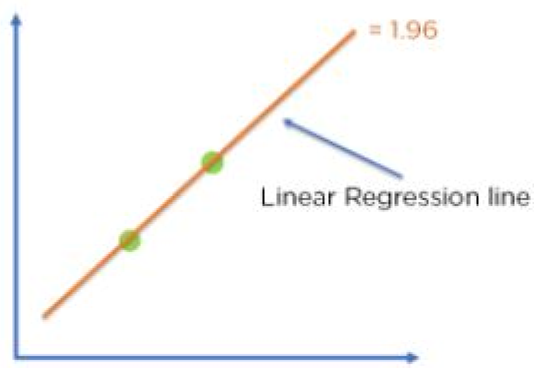


Figure 8: Linear regression model

Cost function = Loss + λ x$\sum\|w\|^2$

For Linear Regression line, let's consider two points that are on the line,

Loss = 0 (considering the two points on the line)

λ= 1

w = 1.4

Then, Cost function = 0 + 1 x 1.42

= 1.96

For Ridge Regression, let’s assume,

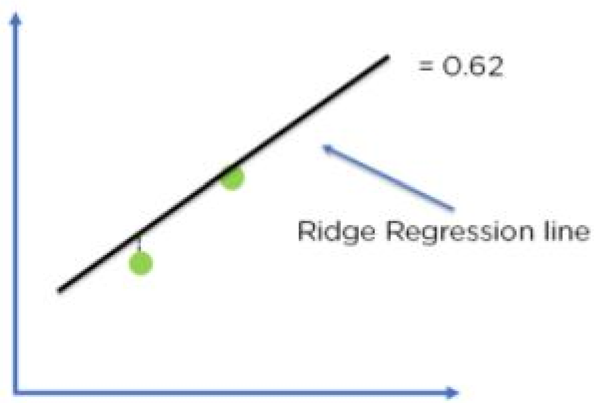
Loss = 0.32 + 0.22 = 0.13

λ = 1

w = 0.7

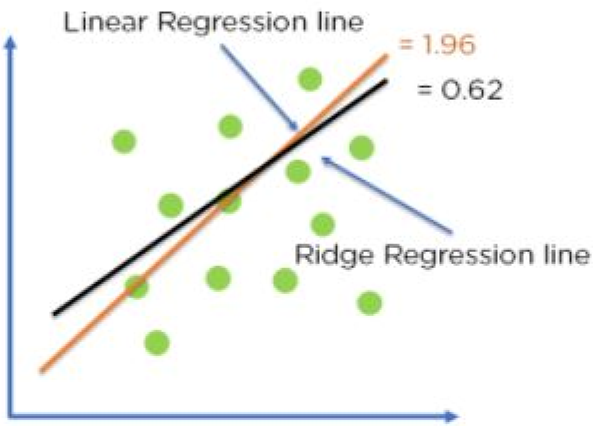
Then, Cost function = 0.13 + 1 x 0.72

= 0.62



Ridge regression model

Comparing the two models, with all data points, we can see that the Ridge regression line fits the model more accurately than the linear regression line.



Optimization of model fit using Ridge Regression

Lasso Regression

It modifies the over-fitted or under-fitted models by adding the penalty equivalent to the sum of the absolute values of coefficients.

Lasso regression also performs coefficient minimization, but instead of squaring the magnitudes of the coefficients, it takes the true values of coefficients. This means that the coefficient sum can also be 0, because of the presence of negative coefficients. Consider the cost function for Lasso regression :

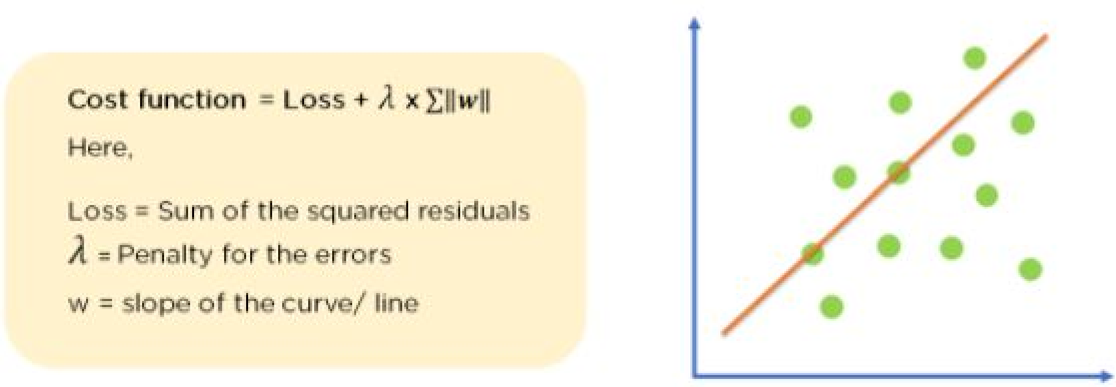


Figure 11: Cost function for Lasso Regression

We can control the coefficient values by controlling the penalty terms, just like we did in Ridge Regression. Again consider a Linear Regression model :

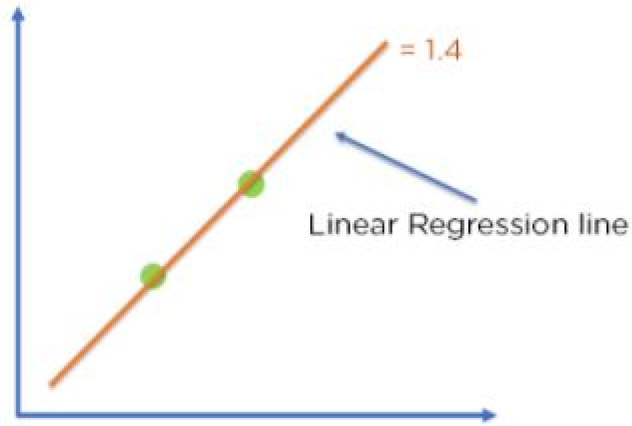


Figure 12: Linear Regression Model

Linear Regression Model

Cost function = Loss + λ x ∑‖w‖

For Linear Regression line, let's assume,

Loss = 0 (considering the two points on the line)

λ = 1

w = 1.4

Then, Cost function = 0 + 1 x 1.4

= 1.4

For Ridge Regression, let's assume,

Loss = 0.32 + 0.12 = 0.1

λ = 1

w = 0.7

Then, Cost function = 0.1 + 1 x 0.7

= 0.8

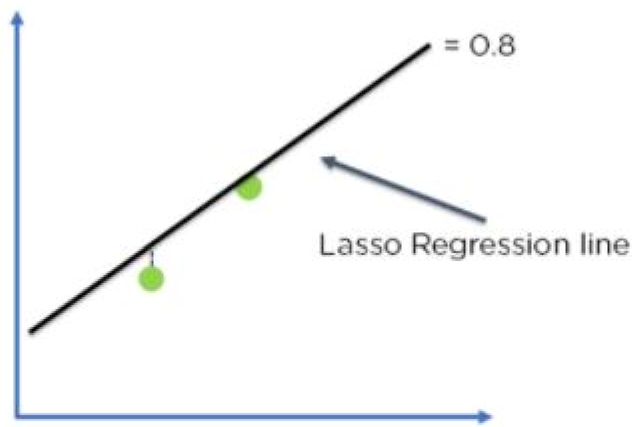


Figure 13: Lasso Regression

Comparing the two models, with all data points, we can see that he Lasso regression line fits the model more accurately than the linear regression line.

15) Explain the term error present in linear regression equation?

An error term represents the margin of error within a statistical model; it refers to the sum of the deviations within the regression line, which provides an explanation for the difference between the theoretical value of the model and the actual observed results.
A regression line always has an error term because, in real life, independent variables are never perfect predictors of the dependent variables. Rather the line is an estimate based on the available data. So the error term tells you how certain you can be about the formula.