Worksheet-1(Assignment-39)

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

13. **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.

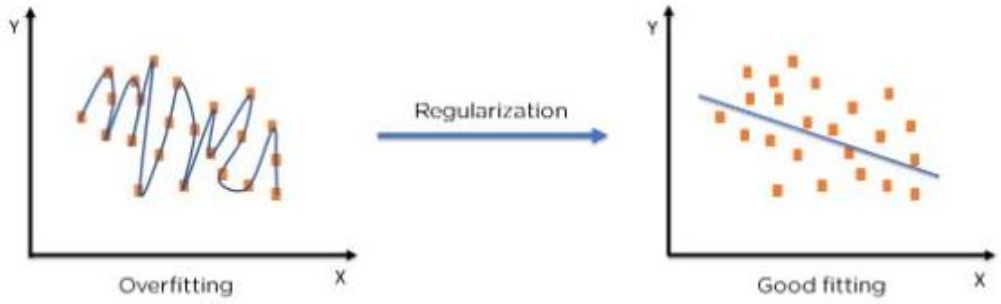


Figure 5: Regularization on an over-fitted model

14. Different techniques of regularizations are: - 1) Ridge regularization

2) 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.

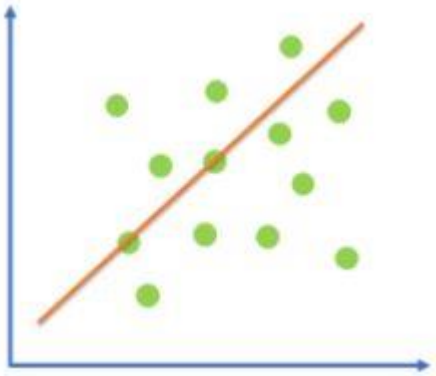
Cost function = Loss + λ x $\sum \|w\|^2$

Here,

Loss = Sum of the squared residuals

λ = Penalty for the errors

W = slope of the curve/ line



Lasso Regularization: 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

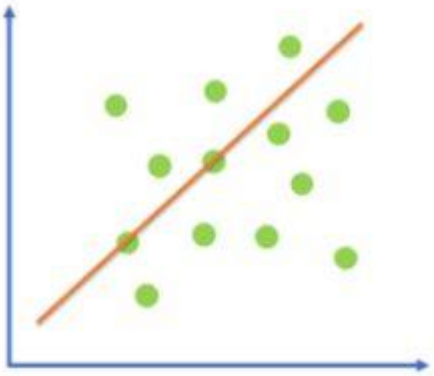
Cost function = Loss + λ x $\sum \|w\|$

Here,

Loss = Sum of the squared residuals

λ = Penalty for the errors

w = slope of the curve/ line



15. The error term is the difference between the data predicted by the model and actual value.

There are 3 types of metrics used in Linear Regression used for predicting error in model:

a) **Mean Absolute Error**: The Mean absolute error represents the average of the absolute difference between the actual and predicted values in the dataset. It measures the average of the residuals in the dataset.

b) **Mean Squared Error**: Mean Squared Error represents the average of the squared difference between the original and predicted values in the data set. It measures the variance of the residuals.

c) **Root Mean Squared Error**: Root Mean Squared Error is the square root of Mean Squared error. It measures the standard deviation of residuals.