MACHINE LEARNING ASSIGNMENT

Q1-Q11:

1. A

2. A

3. B

4. B

5. C

6. B

7. D

8. D

9. A

10. B

11. A, B, C

Q12:

12. A, B, C

Q13-Q15:

13. Regularization is one of the most important concepts of machine learning. It is a technique to prevent the model from overfitting by adding extra information to it. Sometimes the machine learning model performs well with the training data but does not perform well with the test data. It means the model is not able to predict the output when deals with unseen data by introducing noise in the output, and hence

the model is called overfitted. This problem can be deal with the help of a regularization technique.

This technique can be used in such a way that it will allow to maintain all variables or features in the model by reducing the magnitude of the variables. Hence, it maintains accuracy as well as a generalization of the model.

14. Ridge Regression:

- Ridge regression is one of the types of linear regression in which a small amount of bias is introduced so that we can get better long-term predictions.

- Ridge regression is a regularization technique, which is used to reduce the complexity of the model. It is also called as L2 regularization.

- In this technique, the cost function is altered by adding the penalty term to it. The amount of bias added to the model is called Ridge Regression penalty. We can calculate it by multiplying with the lambda to the squared weight of each individual feature.

Lasso Regression:

- Lasso regression is another regularization technique to reduce the complexity of the model. It stands for Least Absolute and Selection Operator.

- It is similar to the Ridge Regression except that the penalty term contains only the absolute weights instead of a square of weights.

- Since it takes absolute values, hence, it can shrink the slope to 0, whereas Ridge Regression can only shrink it near to 0.

- It is also called as L1 regularization.

15. An error term in statistics is a value which represents how observed data differs from actual population data. It can also be a variable which represents how a given statistical model differs from reality.

The error term includes everything that separates your model from actual reality. This means that it will reflect non-linearities, unpredictable effects, measurement errors and omitted variables.