ASSIGNMENT 6
MACHINE LEARNING

1.C

2.B

3.C

4.A

5.B

6.A & D

7.B & C

8.A & C

9.A & C

10. Explain how does the adjusted R-squared penalize the presence of unnecessary predictors in the model?

The adjusted R-squared compensates for the addition of variables and only increases if the new predictor enhances the model above what would be obtained by probability. Conversely, it will decrease when a predictor improves the model less than what is predicted by chance.

11. Differentiate between Ridge and Lasso Regression.

Similar to the lasso regression, ridge regression puts a similar constraint on the coefficients by introducing a penalty factor. However, while lasso regression takes the magnitude of the coefficients, ridge regression takes the square. Ridge regression is also referred to as L2 Regularization.

12.What is VIF? What is the suitable value of a VIF for a feature to be included in a regression modelling?

A variance inflation factor is a measure of the amount of multicollinearity in regression analysis. Multicollinearity exists when there is a correlation between multiple independent variables in a multiple regression model.

The suitable value of a VIF for a feature to be included in a regression modelling is less than 10

13. Why do we need to scale the data before feeding it to the train the model?

To ensure that the gradient descent moves smoothly towards the minima and that the steps for gradient descent are updated at the same rate for all the features, we scale the data before feeding it to the model.

14. What are the different metrics which are used to check the goodness of fit in linear regression?

- MEAN ABSOLUTE ERROR
- MEAN SQUARED ERROR
- ROOT MEAN SQUARED ERROR

15. From the following confusion matrix calculate sensitivity, specificity, precision, recall and accuracy.

ACTUAL/PREDICTED	TRUE	FALSE
TRUE	1000 TP	50 FP
FALSE	250 FN	1200 TN

- PRECISION = TP/ TP + FP

=1000/1000+ 50

=1000/1050

=0.9528

- RECALL = TP/TP+FN

=1000/1000+250

=1000/1250

=0.80

- ACCURACY=(TP+TN)/(TP+TN+FP+FN)

=1000+1200/1000+1200+250+50

0.88

- SPECIFICITY=TN / TN + FP

=1200/1200+50

=0.96