Machine Learning Assignment Answers

1. A
2. B
3. D
4. B
5. B
6. A,d
7. B,c
8. A,b
9. A,c
10. 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. 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.
12. A VIF(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.
13. Scaling the target value is a goo idea in regression modelling, scaling the data makes it easy for a model to learn and understand the problem.
14. 3 statistics are used in Ordinary Least Squares (OLS) regression to evaluate model fit: R-squared, the overall F-test, and the Root Mean Square Error(RMSE).
15. Sensitivity: 0.8000
    Specificity: 0.9600
    Precision: 0.9524
    Negative Predictive value: 0.8276
    False positive rate: 0.0400
    False Discovery rate: 0.0476
    False Negative rate: 0.2000
    Accuracy: 0.8800
    F1 score: 0.8696
    Matthews correlation coefficient: 0.7699