WORKSHEET 6

MACHINE LEARNING

Q1.B)

Q2.B)

Q3.C)

Q4.A)

Q5.B)

Q6.A) AND D)

Q7.B) AND C)

Q8.D)

Q9.B) AND C)

Q10**.** 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. It will decrease when a predictor improves the model less than what is predicted by chance.

Q11. a) Ridge regression takes the square of the coefficients and lasso takes the magnitude.

**b)** Lasso regression can be used for automatic feature selection, as the geometry of its constrained region allows coefficient values to inert to zero.

**c)** An alpha value of zero in either ridge or lasso model will have results similar to the regression model.

Q12. Variance inflation factor (VIF) is a measure of the amount of multicollinearity in a set of multiple regression variables. A VIF of three or below is not a cause for concern. As VIF increases, the less reliable our regression results will be.

Q13. 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.

Q14. There are three error metrics that are commonly used for evaluating and reporting the performance of a regression model; they are:

* Mean Squared Error (MSE).
* Root Mean Squared Error (RMSE).
* Mean Absolute Error (MAE)

**Q15. Sentivity : TP/TP+FN**

**1000/(1000+250) = 1000/1250 = 0.8**

**Specificity: TN/TN+FP**

**1200/1250 = 0.96**

**Accuracy = TP + TN / TP + TN + FP + FN**

**2200/2500 = 0.88**

**Precision = TP / TP + FP = 1000/1050 = 0.95**