MACHINE LEARNING

Assignment 6

Q1 C

Q2 B

Q3 C

Q4 B

Q5 B

Q6 A,D

Q7 B,C

Q8 A,C

Q9 A,B

Q10 adjusted r square: when we add new variable to our regression model but there is no shifting in r square value or marginal shifting in r square value, it means adjusted value is decreased. So adjusted r square value indicates there is no use of such variable or it is not good for our regression model. This is sole purpose of adjusted r square.

Q11

In ridge penalty scale down regression coefficient or almost zero

Where as in lasso regression coefficient becomes zero for some variables thus this is used for features/variable selection.(since regression coefficient is zero so independent variable also zero). So removed unnecessary independent variable to make model best fit.

Q12

VIF or variance of inflation factor is metric of correlation/ of one independent variable to the rest of independent variables.

If VIP is high it means the independent variable is well explained by other independent variable so there is no use of it so remove it to make model best fit.

The threshold value for VIP is 5

Q13

Machine learning algorithm works better when the variables on the same scale and closes to standard mean.

Q14

Mean absolute error (MAE)

Mean squared error (MSE)

Root mean squatted error (RMSE)

Root mean squared log error(RMSLE)

Q15

Sensitivity/recall=TP/(TP+FN)=1000/1050=.95 (BASE TRUTH)

Specificity=TN/(TN+FP)=1200/1450=.82

Precision= TP/(TP+FP)=1000/1250=.80(BASE PREDICTION)

Accuracy=TP/(TP+FN+FP+TN)=2200/2500=.88