WORKSHEET SET 4

MACHINE LEARNING

1. A)
2. B)
3. C)
4. A)
5. B)
6. A), D)
7. B), c)
8. A)
9. D)
10.
11.   Lasso- Due to the fact that coefficients will be shrunk towards a mean of zero, less important features in a dataset are eliminated when penalized. The shrinkage of these coefficients based on the alpha value provided leads to some form of automatic feature selection, as input variables are removed in an effective approach.
Ridge- 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.
Lasso can be Used for Model Selection, but not Ridge Regression

12.   A variance inflation factor (VIF) 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. This can adversely affect the regression results. Thus, the variance inflation factor can estimate how much the variance of a regression coefficient is inflated due to multicollinearity. As a rule of thumb, a VIF

of three or below is not a cause for concern. As VIF increases, the less reliable your regression results are going to be.

13. 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.