Machine learning Worksheet 4

1. C
2. B
3. C
4. B
5. B
6. A, D
7. B, C
8. A
9. A, B
10. Adjusted R-squared is a modified version of R-squared that has been adjusted for the number of predictors in the model. The adjusted R-squared increases when the new term improves the model more than would be expected by chance. It decreases when a predictor improves the model by less than expected. Typically, the adjusted R-squared is positive, not negative. It is always lower than the R-squared.

    Adding more independent variables or predictors to a regression model tends to increase the R-squared value, which tempts makers of the model to add even more variables. This is called overfitting and can return an unwarranted high R-squared value. Adjusted R-squared is used to determine how reliable the correlation is and how much it is determined by the addition of independent variables.

    In a portfolio model that has more independent variables, adjusted R-squared will help determine how much of the correlation with the index is due to the addition of those variables. 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.

Machine learning Worksheet 4

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.
    Ridge regression is also referred to as L2 Regularization.

$$\boxed{L_{hridge}(\hat{\beta}) = \sum_{i=1}^{n}(y_i - x_i'\hat{\beta})^2 + \lambda \sum_{j=1}^{m} w_j \hat{\beta}_j^2.}$$

One obvious advantage of lasso regression over ridge regression, is that it produces simpler and more interpretable models that incorporate only a reduced set of the predictors. However, neither ridge regression nor the lasso will universally dominate the other.

Lasso can be Used for Model Selection, but not Ridge Regression

ridge regression can be used for the analysis of prostate-specific antigen and clinical measures among people who were about to have their prostates removed. The performance of ridge regression is good when there is a subset of true coefficients which are small or even zero.
Lasso regression is a regularisation technique. It is used over regression methods for a more accurate prediction. This model uses shrinkage. Shrinkage is where data values are shrunk towards a central point as the mean.

Machine learning Worksheet 4

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.

    4 , 5 or above will be a suitable value of a VIF for a feature to be included in a regression modelling.

13. 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.
    Scaling the target value is a good idea in regression modelling; scaling of the data makes it easy for a model to learn and understand the problem. Scaling of the data comes under the set of steps of data pre-processing when we are performing machine learning algorithms in the data set.