MACHINE LEARNING(worksheet-11)

1.(B)

2.(B)

3.(A)

4.(C)

5.(B)

6.(A)

7.(B)

8.(C)

9.(A, B, D)

10.(B)

11. (A,B)

12. R-squared (R2) is a statistical measure that represents the proportion of the variance for a dependent variable that's explained by an independent variable or variables in a [regression](https://www.investopedia.com/terms/r/regression.asp)model. Whereas correlation explains the strength of the relationship between an independent and dependent variable, R-squared explains to what extent the variance of one variable explains the variance of the second variable. Adjusted R-squared compares the descriptive power of regression models that include diverse numbers of predictors. Every predictor added to a model increases R-squared and never decreases it. , a model with more terms may seem to have a better fit just for the fact that it has more terms, while the adjusted R-squared compensates for the addition of variables and only increases if the new term enhances the model above what would be obtained by probability and decreases when a predictor enhances the model less than what is predicted by chance

13.Regression metric which measures the average magnitude of errors in a group of predictions, without considering their directions. In other words, it’s a mean of absolute differences among predictions and expected results where all individual deviations have even importance. The function takes as an input two arrays of the same size: predictions and targets. The parameter m of the formula, which is the number of samples, equals to the length of sent arrays. Thanks to the fact that arrays have the same length it is possible to iterate over both of them at the same time. The absolute value of the difference between each prediction and target is calculated and added to accumulated\_error variable. After gathering errors from all pairs, the accumulated result is averaged by the parameter m which returns MAE error for given data

14. SST is the maximum sum of squares of errors for the data because the minimum information of Y itself was only used for the baseline model.

SSR is the additional amount of explained variability in Y due to the regression model compared to the baseline model.

SSE is the difference between SST and SSR is remaining unexplained variability of Y after adopting the regression model.

15. The various metrics used to evaluate the results of the prediction are :

Mean Squared Error(MSE)

Root-Mean-Squared-Error(RMSE).

Mean-Absolute-Error(MAE).

R² or Coefficient of Determination.

Adjusted R²