MACHINE LEARNING – WORKSHEET 11 (LINEAR REGRESSION)

1]B

2]D

3]A

4]C

5]A

6]D

7]B

8]C

9]BD

10]ABC

12] R2= 1-(SSr/SSt) where SSr sum of residual( Actual- PRDEICTED) , Sst - sum of total (actual -mean of predicted)

Now when the feature are added to the dataset , r2 tends to increase irrespective of the feature correlation with th target , Thus Adjusted R2 is instriduced so that to prevent the problem.

Adjusted r-square is a modified form of r-square whose value increases if new predictors tend to improve model’s performance and decreases

if new predictors does not improve performance as expected.

R2adj=1-[(1-r2)(n-1)/(n-k-1)]

Here, k is the no. of regressors and n is the sample size.

if the newly added variable is good enough to improve model’s performance, then it will overwhelm the decrease due to k. Otherwise, increase in k will decrease adjusted r-square value.

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. ... The difference between SST and SSR is remaining unexplained variability of

Y after adopting the regression model, which is called as sum of squares of errors (SSE)

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)