ASSIGNEMENT : MACHINE LEARNING

1. In which of the following you can say that the model is overfitting?

Ans. High R-squared value for train-set and Low R-squared value for test-set.

2. Which among the following is a disadvantage of decision trees?

Ans. Decision trees are highly prone to overfitting

3. Which of the following is an ensemble technique?

Ans. Random Forest

4. Suppose you are building a classification model for detection of a fatal disease where detection of the disease is most important. In this case which of the following metrics you would focus on?

Ans. Accuracy

5. The value of AUC (Area under Curve) value for ROC curve of model A is 0.70 and of model B is 0.85. Which of these two models is doing better job in classification?

Ans. Model B

6. Which of the following are the regularization technique in Linear Regression?

Ans. R-squared

7. Which of the following is not an example of boosting technique?

Ans. Random forest

8. Which of the techniques are used for regularization of Decision Trees?

Ans. None of the above

9.Which of the following statements is true regarding the Ada boost technique?

Ans. It is example of Bagging technique

10.Explain how does the adjusted R-squared penalize the presence of unnecessary predictors in the model?

Ans. In this case R- squared model are going to adjusting the preditors in the regression model by adding the variable increase the value count of the model by adding higher inputs model score become high

11. Differentiate between Ridge and Lasso Regression?

Ans. In Ridge regression alogorith is checking the high correlation among the models value of correlation increases least squared unbaised values data set can be very high in some bias values it having a chance of overfitting

In Laso regression:

Laso regression will give you the accurate values it is similar to the mean model having a high level of multicollinearity it particularly takes when more number of features.

13. Why do we need to scale the data before feeding it to the train the model?

Ans. As we using the scale In the machine learning to help the model to assign equal importance to all features and make predictions

14. What are the different metrics which are used to check the goodness of fit in linear regression?

Ans. METRICS that we are checking inside the linear regression

i. R-squared
ii. Overall F test
iii. Root mean squared Error