MACHINE LEARNING

1) C) High R-squared value for train-set and Low R-squared value for test-set.
2) B) Decision trees are highly prone to overfitting.
3) C) Random Forest
4) B) Sensitivity
5) B) Model B
6) A) Ridge D) Lasso
7) B) Decision Tree C) Random Forest
8) A) Pruning C) Restricting the max depth of the tree
9) A) We initialize the probabilities of the distribution as 1/n, where n is the number of data-points B) A tree in the ensemble focuses more on the data points on which the previous tree was not performing well
10) T
11) Ridge and lasso are two different regularization techniques. Lasso uses l1 whereas ridge uses l2 penalty functions. Ridge shrinks the value of coefficients but doesn't reach zero, which suggests no feature selection. Lasso shrinks coefficients to zero, which helps in feature selection.
12) VIF or Variance inflation factor is a measure of the amount of multicollinearity in a set of multiple regression variables. VIF for a regression model variable is equal to the ratio of the overall model variance to the variance of a model that includes only that single independent variable.
13) Scaling data before training is to normalize the data within a particular range. It helps in speeding up the calculation in an algorithm.
14) Mean absolute error: It takes the sum of absolute value of error. It is more of a direct representation of errors.
Root Mean squared error: It is calculated by the sum of squares of prediction error which is real output minus predicted output and then dividing it by the number of samples. T gives you an absolute number on how much your predicted results deviate from the actual number.
R square: It is the square of correlation coefficient. Measures how much variability in dependent variables can be explained by the model.
15) Precision = 0.80
Recall / Sensitivity = 0.95
Accuracy = 0.88
Specificity = 0.83