MACHINE LEARNING ASSIGNMENT

ANSWERS:

Objective type:

1. c. between -1 and 1
2. d. Ridge Regularisation
3. c. hyperplane
4. a. Logistic Regression
5. b. same as old coefficient of 'X'
6. c. decreases
7. c. Random Forests are easy to interpret
8. b and c
9. a, b and d
10. a, b and d

Subjective type:

11. **Outliers:** An outlier is an observation that lies an abnormal distance from other values in a random sample from a population.

    IQR:
    IQR is the range between the first and the third quartiles namely Q1 and Q3:
    IQR = Q3 – Q1.
    The data points which fall below Q1 – 1.5 IQR or above Q3 + 1.5 IQR are outliers.

12. In **bagging** weak learners' model learns from each other independently in parallel and combines them for determining the model average while in **boosting** weak learners learn sequentially and adaptively to improve model predictions of a learning algorithm.

13. Adjusted R^2 measures the proportion of variation explained by only those independent variables that really help in explaining the dependent variable.

    Adjusted R-squared value can be calculated based on value of r-squared, number of independent variables (predictors), total sample size.

    The formula for the same is below:
    Adjusted R-squared = $1 - (1 – R^2) * ((n - 1)/(n-p-1))$

    where:

R^2 = sample R-square
p = Number of predictors
N = Total sample size

14. In **normalization** minimum and maximum value of features are used for scaling while in **standardisation** mean and standard deviation is used for scaling.

15. **Cross-Validation:** It is a statistical method of evaluating and comparing learning algorithms by dividing data into two segments: one used to learn or train a model and the other used to validate the model.

    Advantage: An advantage of using this method is that we make use of all data points and hence it is low bias.

    Disadvantage: The disadvantage of this method is that the training algorithm has to be rerun from scratch k times, which means it takes k times as much computation to make an evaluation.