MACHINE LEARNING

Answers:-

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 and D) Lasso
7. B) Decision Tree
8. A) Pruning and C) Restricting the max depth of the tree
9. B) A tree in the ensemble focuses more on the data points on which the previous tree was not performing well
10. The adjusted R-squared takes into account the number of variables in a model and adjusts the R-squared value so that it only increases if the variables actually improve the model. This penalizes the presence of unnecessary variables, as including them will lower the adjusted R-squared.

11.

Ridge Regression	Lasso Regression
c) It adds a penalty term equal to the **square** of the magnitude of the coefficients to the loss function.	a) It adds a penalty term equal to the **absolute value** of the magnitude of the coefficients to the loss function.
b) Ridge Regression shrinks the coefficients to be small but doesn't force them to be exactly zero.	b) Lasso Regression can set some of the coefficients to zero, effectively eliminating them from the model.
a) More suited for models with many correlated features.	c) Better for models with a few important features.

12. VIF (Variance Inflation Factor) is a statistic that measures the impact of collinearity between variables in a multiple regression model. A high VIF means that a predictor is highly correlated with other predictors in the model, making the coefficients unstable and difficult to interpret. A suitable value of VIF for a feature to be included in a regression modeling is typically below 5 or 10.
13. Scaling the data is important because many machine learning algorithms are based on the distance between observations, and the scale of the features can greatly affect this distance. If the scale of different features is very different, this can lead to some features dominating others and biasing the model. Scaling helps to ensure that all features are on a similar scale and contribute equally to the model. This can improve the performance and stability of the model.

14. i. R-squared
    ii. Mean Squared Error (MSE)
    iii. Root Mean Squared Error (RMSE)
    iv. Mean Absolute Error (MAE)

v. Adjusted R-squared.

15. Sensitivity (True Positive Rate, or Recall) = 1000 / (1000 + 250) = 0.8
    Specificity (True Negative Rate) = 1200 / (1200 + 50) = 0.96
    Precision (Positive Predictive Value) = 1000 / (1000 + 50) = 0.95
    Accuracy = (1000 + 1200) / (1000 + 50 + 250 + 1200) = 0.89