1. Every time you add an independent variable to a model, the R-squared increases, even if the independent variable is insignificant. It never declines. Whereas Adjusted R-squared increases only when independent variable is significant and affects dependent variable. Because of this RSS method is better.

2. TSS tells you how much variation there is in the dependent variable.
   The Explained SS tells you how much of the variation in the dependent variable your model explained.
   The residual sum of squares (RSS) is a statistical technique used to measure the variance in a data set that is not explained by the regression model.
   TSS = ESS + RSS

3. Regularization refers to techniques that are used to calibrate machine learning models in order to minimize the adjusted loss function and prevent overfitting or underfitting. Using Regularization, we can fit our machine learning model appropriately on a given test set and hence reduce the errors in it.

4. Gini Impurity is a measurement used to build Decision Trees to determine how the features of a dataset should split nodes to form the tree.

5. Decision trees are prone to overfitting, especially when a tree is particularly deep. This is due to the amount of specificity we look at leading to smaller sample of events that meet the previous assumptions. This small sample could lead to unsound conclusions.

6. In statistics and machine learning, ensemble methods use multiple learning algorithms to obtain better predictive performance than could be obtained from any of the constituent learning algorithms alone.

7. Bagging is a method of merging the same type of predictions. Boosting is a method of merging different types of predictions. Bagging decreases variance, not bias, and solves over-fitting issues in a model. Boosting decreases bias, not variance.

8. The out-of-bag error is the average error for each calculated using predictions from the trees that do not contain in their respective bootstrap sample. This allows the Random Forest Classifier to be fit and validated whilst being trained.

9. K-fold Cross-Validation is when the dataset is split into a K number of folds and is used to evaluate the model's ability when given new data.

10. Hyperparameter tuning consists of finding a set of optimal hyperparameter values for a learning algorithm while applying this optimized algorithm to any data set. That combination of hyperparameters maximizes the model's performance, minimizing a predefined loss function to produce better results with fewer errors.

11. A learning rate that is too large can cause the model to converge too quickly to a suboptimal solution, whereas a learning rate that is too small can cause the process to get stuck.

12. No. because the target label has no linear correlation with the features. In such cases, logistic regression (or linear regression for regression problems) can't predict targets with good accuracy (even on the training data)

13. AdaBoost is the first designed boosting algorithm with a particular loss function. On the other hand, Gradient Boosting is a generic algorithm that assists in searching the approximate solutions to the additive modelling problem.

14. In statistics and machine learning, the bias–variance trade off is the property of a model that the variance of the parameter estimated across samples can be reduced by increasing the bias in the estimated parameters.

15. A) Linear – It is the most basic kernel type, usually one-dimensional in nature. It proves to be the best function when there are lots of features. Linear kernel functions are faster than other functions.

    B) RBF - Gaussian RBF kernel. It is one of the most preferred and used kernel functions in SVM. It is usually chosen for non-linear data. It helps to make proper separation when there is no prior knowledge of data.

    C) Polynomial Kernel - It is a more generalized representation of the linear kernel. It is not as preferred as other kernel functions as it is less efficient and accurate.