MACHINE LEARNING

1).R-squared is generally a better measure of the goodness of fit for a regression model than the residual sum of squares (RSS). , is a statistical measure that represents the proportion of the variance for the dependent variable that's explained by the independent variables in the model.

2).TSS = ESS + RSS, where TSS is Total Sum of Squares, ESS is Explained Sum of Squares and RSS is Residual Sum of Suqares. The aim of Regression Analysis is explain the variation of dependent variable Y.

3).Regularization plays several crucial roles in developing and performing machine learning models. Its main purposes revolve around managing model complexity, improving generalization to new data, and addressing specific issues like multicollinearity and feature selection.

4).The Gini Index is the additional approach to dividing a decision tree. Purity and impurity in a junction are the primary focus of the Entropy and Information Gain framework. The Gini Index, also known as Impurity, calculates the likelihood that somehow a randomly picked instance would be erroneously cataloged.

5).yes,unregularized decision trees are overfitting.If the training dataset is not representative, decision trees may overfit to the training data's idiosyncrasies, resulting in poor generalization. Lack of Early Stopping: Without proper stopping rules, decision trees may grow excessively, perfectly fitting the training data but failing to generalize well.

6).Ensemble learning is a machine learning technique that enhances accuracy and resilience in forecasting by merging predictions from multiple models. It aims to mitigate errors or biases that may exist in individual models by leveraging the collective intelligence of the ensemble.

7).Bagging is the simplest way of combining predictions that belong to the same type while Boosting is a way of combining predictions that belong to the different types. Bagging aims to decrease variance, not bias while Boosting aims to decrease bias, not variance.

8).In a Random Forest model, each tree within the ensemble calculates the Out-of-Bag (OOB) error using the data samples it did not select for training during the bootstrap sampling process. These samples, referred to as “out-of-bag” samples, are the ones left out for each tree.

9).In K-fold cross-validation, the data set is divided into a number of K-folds and used to assess the model's ability as new data become available. K represents the number of groups into which the data sample is divided. For example, if you find the k value to be 5, you can call it 5-fold cross-validation.

10).Hyperparameters directly control model structure, function, and performance. Hyperparameter tuning allows data scientists to tweak model performance for optimal results. This process is an essential part of machine learning, and choosing appropriate hyperparameter values is crucial for success.

11).If the learning rate is too high, the algorithm may overshoot the minimum, and if it is too low, the algorithm may take too long to converge. Overfitting: Gradient descent can overfit the training data if the model is too complex or the learning rate is too high.

12).Logistic regression is simple and easy to implement, but it also has some drawbacks. One of them is that it assumes a linear relationship between the input features and the output. This means that it cannot capture the complexity and non-linearity of the 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. This makes Gradient Boosting more flexible than AdaBoost.

14).The bias-variance tradeoff is about finding the right balance between simplicity and complexity in a machine learning model. High bias means the model is too simple and consistently misses the target, while high variance means the model is too complex and shoots all over the place.

15).linear kernel :The linear kernel is the simplest kernel function.it is used when the data is linearly separable.It maps the input data into a higher dimensional using a linear transfusion, making it suitable for linear classification problems.

RBF kernel: The RBF kernel also known as the Gaussian kernel, is a popular choice for non-linear classification problems. It maps the input data into an infinite-diamensional space, making it suitable for complex classification problems.It is sensitive to the gamma parameter, which controls the spread of the data.

Polynomial kernel: The polynomial kernel is used for non-linear classification problems and is sensitive to the degree of the polynomial.It maps the input data into a higher-dimensional space using a polynomial transformation. It is suitable for problems where the relationship between
features is polynomial.