

### 1. R-squared or Residual Sum of Squares (RSS) which one of these two is a better measure of goodness of fit model in regression and why?

R-squared and Residual Sum of Squares (RSS) are both measures of the goodness of fit of a regression model, but they serve slightly different purposes.

R-squared is a statistical measure that represents the proportion of the variance for a dependent variable that's explained by an independent variable or variables in a regression model

Residual Sum of Squares (RSS) measures the difference between the observed values of the dependent variable and the predicted values by the model

### 2. What are TSS (Total Sum of Squares), ESS (Explained Sum of Squares) and RSS (Residual Sum of Squares) in regression. Also mention the equation relating these three metrics with each other

(TSS) is the sum of squared differences between the observed dependent variables and the overall mean

(ESS) is the sum of the squares of the deviations of the predicted values from the mean value of a response variable, in a standard regression model'

(RSS) measures the difference between your observed data and the model's predictions

sum of squares (TSS) = explained sum of squares (ESS) + residual sum of squares (RSS)

### 3. What is the need of regularization in machine learning

It is a technique to prevent the model from overfitting by adding extra information to it.

### 4. What is Gini-impurity index?

Gini index, is a measure of how impure or mixed a dataset is. It's used in decision trees to help identify the best split in a dataset by finding features that create more homogeneous subsets.

### 5. Are unregularized decision-trees prone to overfitting? If yes, why?

Yes, they learn too much from the training data and fail to generalize well to new data.

## **6. What is an ensemble technique in machine learning?**

Ensemble learning refers to a machine learning approach where several models are trained to address a common problem, and their predictions are combined to enhance the overall performance.

## **7. What is the difference between Bagging and Boosting techniques?**

Bagging attempts to tackle the over-fitting issue. Boosting tries to reduce bias. If the classifier is unstable (high variance), then we need to apply bagging. If the classifier is steady and straightforward (high bias), then we need to apply boosting.

## **8. What is out-of-bag error in random forests?**

The out-of-bag error is an error estimation technique often used to evaluate the accuracy of a random forest and to select appropriate values for tuning parameters.

## **9. What is K-fold cross-validation?**

Fold Cross-Validation is a robust technique used to evaluate the performance of machine learning models.

## **10. What is hyper parameter tuning in machine learning and why it is done?**

Hyperparameter tuning is the process of finding the best set of hyperparameters for a machine learning model to optimize its performance.

## **11. What issues can occur if we have a large learning rate in Gradient Descent?**

A large learning rate in Gradient Descent can cause the algorithm to overshoot the optimal point, which can lead to poor model performance or failure to converge.

## **12. Can we use Logistic Regression for classification of Non-Linear Data? If not, why?**

No, because it assumes a linear relationship between the input features and the output.

## **13. Differentiate between Adaboost and Gradient Boosting.**

gradient boosting minimizes a loss function like MSE or log loss while AdaBoost focuses on instances with high error by adjusting their sample weights adaptively.

## **14. What is bias-variance trade off in machine learning?**

the bias–variance tradeoff describes the relationship between a model's complexity, the accuracy of its predictions, and how well it can make predictions on previously unseen data that were not used to train the model

**15. Give short description each of Linear, RBF, Polynomial kernels used in SVM.**

In machine learning, the radial basis function kernel, or RBF kernel, is a popular kernel function used in various kernelized learning algorithms.

Linear Kernel is used when the data is Linearly separable, that is, it can be separated using a single Line.

The polynomial kernel is a kernel function used in support vector machines (SVMs) to model non-linear relationships between input features