

# **MACHINE LEARNING**

**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?**

**Ans:** Residual Sum of Squares (RSS) is the better measure of goodness of fit, focusing on the magnitude of the prediction errors. If minimizing prediction errors is a primary concern, RSS is a more appropriate choice.

**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.**

**Ans:**  $TSS = ESS + RSS$

**3. What is the need of regularization in machine learning?**

**Ans:** Regularization techniques help reduce the chance of overfitting and help us get an optimal model.

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

**Ans:** Gini Impurity is a measure of how often a randomly chosen element from the set would be incorrectly labelled if it was randomly labelled according to the distribution of labels in the subset.

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

**Ans:** Decision trees have a tendency to overfit to the training set because they can keep growing deeper and more complex until they perfectly classify the training data. This can lead to the tree capturing noise in the data, rather than the underlying relationships, and thus performing poorly on new, unseen data.

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

**Ans:** Ensemble methods are techniques that create multiple models and then combine them to produce improved results. Two major types of ensemble techniques: Bagging and Boosting.

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

**Ans: Bagging** is the type of ensemble technique in which a single training algorithm is used on different subsets of the training data where the subset sampling is done with replacement. Once the algorithm is trained on all the subsets, then bagging makes the prediction by aggregating all the predictions made by the algorithm on different subsets.

**Boosting** is an ensemble approach that starts from a weaker decision and keeps on building the models such that the final prediction is the weighted sum of all the weaker decision-makers.

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

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

**Ans:** In K-Fold Cross Validation, we split the dataset into k number of subsets and then we perform training on the all the subsets but leave one (k-1) subset for the evaluation of the trained model. In this method, we iterate k times with a different subset reserved for testing purpose each time.

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

**Ans:** Hyperparameter tuning is settings that control the learning process of the model, such as the learning rate. The goal of hyperparameter tuning is to find the values that lead to the best performance on a given task.

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

**Ans:** If the learning rate is too high in Gradient Descent, the algorithm may overshoot the minimum.

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

**Ans:**

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

**Ans: Adaboost** is is a type of ensemble technique, where a number of weak learners are combined together to form a strong learner. Here, usually, each weak learner is developed as decision stumps (A stump is a tree with just a single split and two terminal nodes) that are used to classify the observations.

**Gradient Boost** also combines a no. of weak learners to form a strong learner. Here, the residual of the current classifier becomes the input for the next consecutive classifier on which the trees are built, and hence it is an additive model.

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

**Ans:** The goal of the machine learning is to strike balance between Bias and Variance. Bias-Variance tradeoff is about finding the right balance between simplicity and Complexity in a Machine Learning Model.

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

**Ans:**