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 - R-squared is generally considered a better measure of goodness of fit in regression models compared to Residual Sum of Squares (RSS). Here's why:

Interpretability: R-squared provides an intuitive interpretation as it represents the proportion of the variance in the dependent variable that is explained by the independent variables.

Normalized Scale: R-squared is normalized, meaning it ranges from 0 to 1. This allows for easier comparison between different models or different datasets.

Relationship to Correlation: R-squared is directly related to the correlation between the observed and predicted values of the dependent variable.

Overall, R-squared is a more comprehensive measure of goodness of fit in regression models because it not only quantifies the amount of variability explained by the model but also provides insights into the relative improvement over a baseline model and the strength of the relationship between the variables.

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 - The Total SS (TSS) tells us how much variation there is in the dependent variable.
Total SS = Σ(Yi – mean of Y)2.

The Explained SS tells us how much of the variation in the dependent variable your model explained.
Explained SS = Σ(Y-Hat – mean of Y)2.

The residual sum of squares tells us how much of the dependent variable's variation our model did not explain. It is the sum of the squared differences between the actual Y and the predicted Y:
Residual Sum of Squares = Σ e2

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

Ans – Its is required so we can fit model into our test

4. What is Gini–impurity index?

Ans - Gini impurity measures how often a randomly chosen element of a set would be incorrectly labeled if it were labeled randomly and independently according to the distribution of labels in the set

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

Ans – Yes they can be overfit because 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?

Ans - Ensemble methods are techniques that create multiple models and then combine them to produce improved results.

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

Ans - 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.

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

Ans - method of measuring the prediction error of random forests, boosted decision trees, and other machine learning models utilizing bootstrap aggregating (bagging).

9. What is K-fold cross-validation?

Ans- We use it for predictive models

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

Ans – Without this model make more errors so If we don't correctly tune our hyperparameters, our estimated model parameters produce suboptimal results, as they don't minimize the loss function.

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

Ans- It can result in exploding or oscillating performance over the training epochs and to a lower final performance

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

Ans- No, we can not because it will hamper end result

13. Differentiate between Adaboost and Gradient Boosting.

Ans - Adaboost is computed with a specific loss function and becomes more rigid when comes to few iterations. But in gradient boosting, it assists in finding the proper solution to additional iteration modeling problem as it is built with some generic features.

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

Ans- By this we can see how will our prediction come by considering non seen data

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

Ans- We use it for support vector machine classification