R-squared vs. RSS in Goodness of Fit:
- R-squared measures the proportion of variance explained by the model, while RSS measures the sum of squared residuals. R-squared is generally a better indicator of goodness of fit because it is normalized between 0 and 1, making it easier to interpret the proportion of variance explained by the model.

2. TSS, ESS, and RSS in Regression:
- TSS (Total Sum of Squares):* Measures the total variance in the dependent variable.
- ESS (Explained Sum of Squares):* Measures the variance explained by the model.
- RSS (Residual Sum of Squares):* Measures the variance not explained by the model (errors).
- Equation:* \( TSS = ESS + RSS \)

3. Need for Regularization:
- Regularization prevents overfitting by adding a penalty to the loss function, controlling model complexity, and ensuring that the model generalizes well to unseen data.

4. Gini Impurity Index:
- A measure of how often a randomly chosen element would be incorrectly classified. It is used in decision trees to determine the purity of a split.

5. Overfitting in Unregularized Decision Trees:
- Yes, unregularized decision trees are prone to overfitting because they can create complex models that perfectly fit the training data but do not generalize well to new data.

6. Ensemble Technique:
- Combining multiple models to improve the overall performance and robustness compared to individual models. Examples include Bagging and Boosting.

7. Bagging vs. Boosting:
- *Bagging:* Reduces variance by training multiple models independently and averaging their predictions.
- *Boosting:* Reduces bias by sequentially training models, where each model corrects the errors of the previous one.

8. Out-of-Bag Error in Random Forests:
- An estimate of the model error calculated using samples not included in the bootstrap sample (out-of-bag samples). It is used as a validation metric.

9. K-Fold Cross-Validation:
- A technique to assess model performance by dividing the data into ‘K’ subsets, training on ‘K-1’ subsets, and validating on the remaining subset. This is repeated ‘K’ times to ensure robust evaluation.

10. Hyperparameter Tuning:
- The process of selecting the best set of hyperparameters for a model to improve performance. It is done to optimize the model and avoid overfitting or underfitting.

11. Issues with Large Learning Rate in Gradient Descent:
- A large learning rate can cause the model to overshoot the optimal point, leading to divergence or failure to converge.

12. Logistic Regression for Non-Linear Data:
- Logistic regression cannot directly handle non-linear data. For non-linear relationships, feature transformations or a different model (e.g., SVM with non-linear kernels) is needed.

13. Adaboost vs. Gradient Boosting:
- Adaboost: Focuses on correcting misclassified points by adjusting their weights.
- Gradient Boosting:Builds models sequentially, optimizing the loss function gradient.

14. Bias-Variance Trade-off:
- A trade-off where a model with high bias is too simple (underfits), while a model with high variance is too complex (overfits). The goal is to find a balance to minimize total error.

15. SVM Kernels:
- Linear Kernel: Used for linearly separable data.
RBF Kernel: Captures non-linear relationships using a radial basis function.
Polynomial Kernel:*Handles non-linear data by considering polynomial combinations of features.