Ans 1 R-squared is preferred because it gives a standardized measure (proportion) of goodness of fit, allowing for easier interpretation and comparison of different models.

Ans 3 Regularization is employed to strike a balance between model complexity and model performance, aiding in the creation of models that generalize well to new, unseen data while reducing the risk of overfitting to the training data.

Ans 4 The Gini impurity index is a measure used in decision tree algorithms, especially in binary classification problems, to evaluate the impurity or randomness of a set of examples or data points within a specific node of the tree.

Ans 5 Decision Trees are prone to over-fitting. A decision tree will always overfit the training data if we allow it to grow to its max depth. Overfitting in decision trees occurs when the tree becomes too complex and captures the noise in the training data, rather than the underlying pattern.

Ans 6 Ensemble methods are techniques that aim at improving the accuracy of results in models by combining multiple models instead of using a single model.

Ans 7 bagging involves fitting many models on different samples of the dataset and averaging the predictions, whereas boosting involves adding ensemble members sequentially to correct the predictions made by prior models and outputs a weighted average of the predictions.

Ans 8 Out-of-bag error is a way of measuring the prediction accuracy of a random forest model, which is an ensemble of decision trees. It is calculated using the samples that are not used to train the individual trees, and it provides an unbiased estimate of the model's performance on unseen data.

Ans 9 K-fold cross-validation is a widely used technique in machine learning for assessing the performance of a predictive model. It involves partitioning the dataset into 'k' subsets (folds) of approximately equal size.

Ans 10 Hyperparameter tuning is the process of optimizing hyperparameters in a machine learning model to enhance its performance.

Ans 12 Logistic Regression is a linear classification algorithm designed for binary classification tasks. It separates classes using a linear decision boundary. While Logistic Regression is effective for linearly separable data, it may struggle when dealing with highly non-linear data distributions.

Ans 13 Adaboost and Gradient Boosting are both boosting algorithms used for ensemble learning, aiming to improve the performance of weak learners (typically decision trees) by combining them into a strong learner.

Ans 14 The bias-variance tradeoff is a fundamental concept in machine learning that describes the relationship between the complexity of a model, the accuracy of its predictions, and how well it can generalize to new data.

Ans 15 Linear kernel: This kernel is used when the data is linearly separable, meaning that a straight line can separate the classes. Meanwhile Polynominal kernel is used when the data is not linearly separable, but can be separated by a polynomial function of some degree.

And RBF kernel is also used when the data is not linearly separable, but can be separated by a nonlinear function that depends on the distance between the input vectors.