Que1.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 better measure of goodness of fit model in regression
Because RSS provides an absolute measure of fit but is not as interpretable or comparable across different models.R squared is relatively easy to interpret, the closer to 1.0, the better the model explains the variance in the data.

Que2.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 is calculated when we consider the line passing through the mean value of y to be the best fit line

TSS = $\sum_1^n$(yi − ymean)^2

The RSS measures the amount of error remaining between the regression function and the data set after the model has been run. A smaller RSS figure represents a regression function that is well-fit to the data.

$$\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

ESS is the sum of the differences between the predicted value and the mean of the dependent variable. In other words, it describes how well our line fits the data.

$$\mathrm{ESS} = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 .$$

Que3.What is the need of regularization in machine learning?

Ans – to avoide model over fitting and underfitting we need regularization, so that it reduce errors

Que4. What is Gini–impurity index?
Ans- Gini impurity index is a measure how much impurity our data have. Gini Impurity is a measurement used to build Decision Trees to determine how the features of a dataset should split nodes to form the tree

Que5.Are unregularized decision-trees prone to overfitting? If yes, why?

Ans-Decision trees, by their very nature, are prone to overfitting,because of its complexity especially when they are deep. A tree that is too complex might achieve a perfect accuracy score on the training data but perform poorly on new, unseen data. Such a tree has low bias but high variance, and its predictions can be unstable. So it need hyperparameter tuning

Que6. What is an ensemble technique in machine learning?

Ans- The underlying concept behind ensemble learning is to combine the outputs of diverse models to create a more precise prediction. By considering multiple perspectives and utilizing the strengths of different models, ensemble learning improves the overall performance of the learning system

Que7. What is the difference between Bagging and Boosting techniques?

Ans- The bagging technique combines multiple models trained on different subsets of data, whereas boosting trains the model sequentially, focusing on the error made by the previous model.

Que8. What is out-of-bag error in random forests?

Ans- OOB (out-of-bag) errors are an estimate of the performance of a random forest classifier or regressor on unseen data

Que 9. What is K-fold cross-validation?
Ans- into k subsets or folds. The model is trained and evaluated k times, using a different fold as the validation set each time.

Que10. What is hyper parameter tuning in machine learning and why it is done?

Ans- Hyperparameter tuning is the process of selecting the optimal values for model learning. Hyperparameters are settings that control the learning process of the model, such as the learning rate.
It is used to improve model accuracy and performance.

Que11. What issues can occur if we have a large learning rate in Gradient Descent?

Ans- if we have large learning rate the algoritham will overshoot globalminima .
also the learning rate is too large, gradient descent can suffer from divergence. This means that weights increase exponentially, resulting in exploding gradients which can cause problems like overfitting

Que12. Can we use Logistic Regression for classification of Non-Linear Data? If not, why?

Ans- we can not use Logistic Regression for classification of Non-Linear Data because it assumes a linear relationship between the input features and the output. This means that it cannot capture the complexity and non-linearity of the data.

Que13. Differentiate between Adaboost and Gradient Boosting

Ans- when it comes to Adaptive boosting the approach is done by up-lifting the weighted observation which is misclassified prior and used to train the model to give more efficacy. In gradient boosting, the complex observations are computed by large residues left on the previous iteration to increase the performance of the existing model.

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

Ans-In machine learning, as you try to minimize one component of the error (e.g., bias), the other component (e.g., variance) tends to increase, and vice versa. Finding the right balance

of bias and variance is key to creating an effective and accurate model. This is called the bias-variance tradeoff