MACHINE LEARNING

Q.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 (R^2) and Residual Sum of Squares (RSS) are both commonly used measures to assess the goodness of fit of a regression model, but they capture different aspects of model performance, and the choice between them depends on the context and what you want to evaluate.

Q.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: Goodness of fit: R2. TSS = ESS + RSS, where TSS is Total Sum of Squares, ESS is Explained Sum of Squares and RSS is Residual Sum of Squares. The aim of Regression Analysis is explaining the variation of dependent variable.

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

Ans: While training a machine learning model, the model can easily be overfitted or under fitted. To avoid this, we use regularization in machine learning to properly fit a model onto our test set. Regularization techniques help reduce the chance of overfitting and help us get an optimal model.

Q.4. What is Gini–impurity index?

Ans: 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.

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

Ans: Overfitting can be one problem that describes if your model no longer generalizes well. Overfitting happens when any learning processing overly optimizes training set error at the cost test error. While it's possible for training and testing to perform equality well in cross validation, it could be as the result of the data being very close in characteristics, which may not be a huge problem. In the case of decision tree's, they can learn a training set to a point of high granularity that makes them easily overfit. Allowing a decision tree to split to a granular degree, is the behaviour of this model that makes it prone to learning every point extremely well to the point of perfect classification.

Q. 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. Ensemble methods in machine learning usually produce more accurate solutions than a single model would.

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

Ans: Bagging is a learning approach that aids in enhancing the performance, execution, and precision of machine learning algorithms. Boosting is an approach that iteratively modifies the weight of observation based on the last classification. It is the easiest method of merging predictions that belong to the same type.

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

Ans: The out-of-bag (OOB) error is the average error for each calculated using predictions from the trees that do not contain in their respective bootstrap sample. This allows the RandomForestClassifier to be fit and validated while being trained.

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

Ans: K-fold cross-validation is a technique for evaluating predictive models. The dataset is divided into k subsets or folds. The model is trained and evaluated k times, using a different fold as the validation set each time. Performance metrics from each fold are averaged to estimate the model's generalization performance.

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

Ans: Hyperparameter tuning allows data scientists to tweak model performance for optimal results. This process is an essential part of machine learning, and choosing appropriate hyperparameter values is crucial for success. For example, assume you're using the learning rate of the model as a hyperparameter.

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

Ans: When 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 such as instabilities and overly high loss values.

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

Ans: Logistic regression has traditionally been used to come up with a hyperplane that separates the feature space into classes. But if we suspect that the decision boundary is nonlinear, we may get better results by attempting some nonlinear functional forms for the logit function.

Q. 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 modelling problem as it is built with some generic features.

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

Ans: In machine learning, the bias–variance trade-off describes the relationship between a model's complexity, the accuracy of its predictions, and how well it can make predictions on previously unseen data that were not used to train the model.

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

Ans: Linear Kernel: Decision Boundary: Form: The linear kernel produces a decision boundary that is a hyperplane in the feature space. This hyperplane separates data points from different classes in a linear fashion.

In machine learning, the radial basis function kernel, or RBF kernel, is a popular kernel function used in various kernelized learning algorithms. In particular, it is commonly used in support vector machine classification.

In machine learning, the polynomial kernel is a kernel function commonly used with support vector machines (SVMs) and other kernelized models, that represents the similarity of vectors (training samples) in a feature space over polynomials of the original variables, allowing learning of non-linear models.