**MACHINE LEARNING**

**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?**

R-squared and RSS serve various functions and provides different understandings. RSS is used to estimate the overall error of the model, R-squared is frequently used to evaluate a model's quality of fit. Thus, none is basically "better" than the other rather, their applicability is dependent upon the particular environment and the understandings we hope to derive from your model.

**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.**

-Total Sum of Squares (TSS): It is the sum of squared differences between the observed dependent variables and the overall mean.

-Explained Sum of Squares (ESS): it is the sum of the differences between the predicted value and the mean of the dependent variable

-Residual Sum of Squares (RSS): it is the difference between the observed and predicted values

TSS = ESS + RSS

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

Regularization plays a crucial role in machine learning as it enhances model accuracy, simplicity, and interpretability, and helps to address overfitting issues

**4. What is Gini–impurity index?**

The Gini Impurity Index is a measure used in decision trees and other machine learning algorithms to determine how the features of a dataset should split nodes to form the tree

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

Yes, Unregularized decision trees are prone to overfit. The tree has essentially learned the noise in the training data rather than the underlying pattern, leading to poor performance on new data.

**6. What is an ensemble technique in machine learning?**

ensemble technique refers to the process of merging many models to get predictions that are more accurate compared to those of any one model alone.

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

Bagging is a process that involves averaging the predictions from several models that have been fitted to various dataset samples. By reducing the variance and assisting in avoiding the risk of overfitting, it improves the stability and accuracy of machine learning algorithms.

Boosting means building a strong classifier from many weak classifiers.

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

The Out-of-Bag (OOB) error is a way of calculating the prediction error of machine learning models, particularly ensemble models like random forest. It is estimated using out-of-bag samples, which were not utilized in the model's training.

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

K-fold cross-validation is a statistical approach for estimating machine learning model ability. In applied machine learning, it is widely used to compare and select a model for a specific predictive Modeling task.

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

The process of identifying the ideal settings for a machine learning model's hyperparameters is known as hyperparameter tuning. Hyperparameters are model learning settings such as the learning rate, the number of neurons in a neural network, or the kernel size in a support vector machine.

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

In Gradient Descent, an enormous learning rate can lead the algorithm to diverge and overshoot the minimum. This may result in oscillations around the minimum or, in rare circumstances, in a complete divergence.

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

Logistic regression is a linear classification algorithm, meaning it assumes a linear relationship between the input features and the log-odds of the output. However, when dealing with non-linear data, where the decision boundary is not a straight line, logistic regression may not perform well.

**13. Differentiate between Adaboost and Gradient Boosting.**

Adaboost technique provides weak learners in a sequential manner. Each iteration involves training a new weak learner on the updated dataset and giving greater weights to the samples that were incorrectly identified in the previous iteration.

Gradient Boosting, on the other hand, trains weak learners in a stagewise fashion. It starts with an initial weak learner and then trains subsequent weak learners to minimize the residual errors made by the previous weak learners.

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

Bias-variance trade-off is a basic concept that refers to the choices between a model’s ability to minimize bias and variance. Bias is the difference between the expected value of the model’s predictions and the true values, while variance is the variability of the model’s predictions for different training sets.

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

**Linear Kernel**: The linear kernel is the simplest kernel function and is used when the data is linearly separable.

**RBF (Radial Basis Function):** It maps the input data into an infinite-dimensional space using Gaussian radial basis functions.

**Polynomial Kernel**: This function is used when the data is not linearly separable. It maps the input data into a higher-dimensional space using polynomials of the original variables.