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?

Ans- Both R^2 and RSS provide different perspective of the model. It depends on the context and the specific goals of our analysis. All in all, R^2 is considered a little better as it provides the answer in a single number which is easier to compare and more interpretable.

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- TSS- It tells the total variability in the dependent variable and also tells the deviation in each data point.

ESS- It tells the deviation of the predicted values from the dependent variable.

RSS- It tells the deviation of the observed values from the predicted values.

Equation relating all of these- TSS=ESS+RSS.

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

Ans- It is used to prevent over-fitting and under- fitting of the model. Regularization helps to get the perfect optimal model. Sometimes the model performs well with the training data but not that good with the test data. Therefore, regularization is important.

4. What is Gini–impurity index?

Ans- This measure is used in decision tree algorithms to check the impurity and disorder of the dataset. It gives a number between 0-0.5, which tells the likelihood of the random data being misclassified.

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

Ans- Yes, unregularized decision-trees are prone to over-fitting as they capture the noise during training along who the underlying patterns. Also, they might miss out on slight changes in the dataset

6. What is an ensemble technique in machine learning?

Ans- In ensemble technique, the prediction is collected from multiple models and they are combined to give a more precise and better result. It produces more accurate solutions.

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

Ans- Bagging and boosting are different forms of ensemble technique. In bagging the multiple models trained no different data are combined whereas in boosting the models are trained sequentially learning from the error made in the previous model and then combined.

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

Ans- It is a method used to measure the predicted error in the random forests. It is calculated by aggregating predictions over trees where each data bag is an out-of-bag sample.

9. What is K-fold cross-validation?

Ans- In this method the dataset is divided into k sets and then the mode is trained and evaluated k times, using a different fold for validation each time. It is used for evaluating predictive models.

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

Ans- Hyper parameter is tuned into machine learning to find the optimal configuration settings also known as hyperparameters. It is done to ensure that the model is able to adapt to different datasets and also to improve the efficiency and consistency of the model.

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

Ans- If the learning rate is too high in Gradient descent it can result into problems such as instability and overly high loss values.

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

Ans- We can use Logistics regression for classification of Non-Linear Data, but it is usually not preferred as it is not able to capture intricate patterns and relationships in the data, making it not very much preferred in non-linear classifications.

13. Differentiate between Adaboost and Gradient Boosting.

Ans- These both are ensemble techniques. Adaboost assigns weight to the data points to correct errors made by the previous models whereas Gradient boosting ensembles decision trees and it is less sensitive to the outliers.

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

Ans- It basically tells the balance between the accuracy and the ability of the model to make predictions also while representing the simplicity and the ability of the model to capture underlying effect.

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

Ans- Linear- This is the simplest kernel and works on linear data and it may face problems while dealing with the non-linear data.

RBF- These are the Radial Basis Functions that use supervised machine learning prepared to capture non-linear relationships.

Polynomial kernels- It is used for capturing polynomial relationships in a dataset.