

MACHINE LEARNING

Q1 to Q15 are subjective answer type questions, Answer them briefly.

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 - The residual sum of squares is the absolute amount of explained variation, whereas R-squared is the absolute amount of variation as a proportion of total variation. Because we feel that large negative residuals are as bad as large positive ones.

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-**Residual sum of squares** is statistical method that helps identify the level of discrepancy in a dataset not predicted by a regression model. **TSS** is the dispersion of observed variables around the mean, or the variance. **ESS** is the sum of the predicted values from the mean value of a response variable. **$TSS = ESS + RSS$**

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

Ans-**Regularization refers to techniques that are used to calibrate machine learning models in order to minimize the adjusted loss function and prevent overfitting or underfitting.**

4. What is Gini-impurity index?

Ans-**Gini – impurity calculates the amount of probability of a specific feature that is classified incorrectly when selected randomly.**

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

Ans-**Overfitting is a common problem, a data scientist needs to handle while training decision tree models. Comparing to other machine learning algorithms, decision trees can easily overfit.**

6. What is an ensemble technique in machine learning?

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

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

Ans-**Bagging is the simplest way of combining predictions that belong to the same type while Boosting is a way of combining predictions that belong to the different types.**

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

Ans-**The out-of-bag error is the average error for each calculated using predictions from the trees that do not contain in their respective bootstrap sample.**

9. What is K-fold cross-validation?

Ans-**K-fold cross-validation is when the dataset is split into a K number of folds and is used to evaluate the model's ability when given new data.**

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

Ans-**Hyperparameter tuning consists of finding a set of learning algorithm while applying this optimized algorithm to any data set. That combination of hyperparameters maximizes the model's performance, minimizing a predefined loss function to produce better results with fewer errors.**

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

Ans-**A learning rate that is too large can cause the model to converge too quickly to a suboptimal solution, whereas a learning rate that is too small can cause the process to get stuck.**

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

Ans-**Logistic Regression has traditionally been used as a linear classifier. Because when the classes can be separated in the feature space by linear boundaries. That can be remedied however if we happen to have a better idea as to the shape of the decision boundary.**

13. Differentiate between Adaboost and Gradient Boosting.

Ans-**Adaboost is the first designed boosting algorithm with a particular loss function. On the other hand, Gradient Boosting is a generic algorithm that assists in searching the approximate solutions to the additive modeling problem. This makes Gradient Boosting more flexible than Adaboost.**

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

Ans-**In statistics and machine learning, the bias-variance trade off is the property of a model that the variance of the parameter estimated across samples can be reduced by increasing the bias in the estimated parameters.**

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

Ans-**The kernel functions are used to map the original dataset into a higher dimensional space with view to making it linear dataset.**

