

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: - However, a smaller or lower value for the RSS is ideal in any model since it means there's less variation in the data set. In other words, the lower the sum of squared residuals, the better the regression model is at explaining the data.

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: -

1. TSS (Total Sum of Squares): -The total sum of squares (TSS) measures how much variation there is in the observed data.

2. RSS (Residual Sum of Squares) :- the residual sum of squares measures the variation in the error between the observed data and modeled values.

3. Explained sum of square (ESS) or Regression sum of squares or Model sum of squares is a statistical quantity used in modeling of a process.

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 is a measurement used to build Decision Trees 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?

Ans:- Yes, Decision trees are prone to overfitting, especially when a tree is particularly deep. This is due to the amount of specificity we look at leading to smaller sample of events that meet the previous assumptions. This small sample could lead to unsound conclusions.

6. What is an ensemble technique in machine learning?

Ans: - Ensemble methods is a machine learning technique that combines several base models in order to produce one optimal predictive model.

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

Ans: - Bagging is a technique for reducing prediction variance by producing additional data for training from a dataset by combining repetitions with combinations to create multi-sets of the original data. Boosting is an iterative strategy for adjusting an observation's weight based on the previous classification.

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.

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. K refers to the number of groups the data sample is split into.

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

Ans: - In machine learning, hyperparameter optimization or tuning is the problem of choosing a set of optimal hyperparameters for a learning algorithm. A hyperparameter is a parameter whose value is used to control the learning process. By contrast, the values of other parameters (typically node weights) are learned.

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

Ans:- If learning rate is too large, gradient descent can overshoot the minimum. It may fail to converge and even diverge.

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

Ans: - It can only be used to predict discrete functions. Hence, the dependent variable of Logistic Regression is bound to the discrete number set. It is very fast at classifying unknown records. Non-linear problems can't be solved with logistic regression because it has a linear decision surface.

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 modelling problem. This makes Gradient Boosting more flexible than AdaBoost.

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

Ans: - The bias-variance tradeoff 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: - SVM algorithms use a group of mathematical functions that are known as kernels. The function of a kernel is to require data as input and transform it into the desired form.

Different SVM algorithms use differing kinds of kernel functions. These functions are of different kinds—for instance, linear, nonlinear, polynomial, radial basis function (RBF), and sigmoid.

Linear Kernel

It is the most basic type of kernel, usually one dimensional in nature. It proves to be the best function when there are lots of features. The linear kernel is mostly

preferred for text-classification problems as most of these kinds of classification problems can be linearly separated. Linear kernel functions are faster than other functions.

RBF

It is one of the most preferred and used kernel functions in svm. It is usually chosen for non-linear data. It helps to make proper separation when there is no prior knowledge of data.

Polynomial kernels

It is a more generalized representation of the linear kernel. It is not as preferred as other kernel functions as it is less efficient and accurate.