

MACHINE LEARNING

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(Total sum of Squares) is the sum of squares of the differences between sample value and the mean of sample variable. It is a measure of total variance of dependent variable.

ESS(Explained Sum of Squares) is the sum of squares of the differences between predicted value and the mean of dependent variable.

RSS(Residual Sum of Squares) is the sum of squares of the residuals which are the differences between the observed values and the predicted values. It is a statistical technique used to measure the amount of variance in a data set that is not explained by a regression model.

Equation relating these three metrics:

$$TSS = ESS + RSS$$

9) What is K-fold cross-validation?

ANS: K-fold cross-validation is a model validation method in which the dataset is divided into K parts and each of these parts is used as a validation set and remaining part used as a training set. This method is helpful for avoiding overfitting.

6) What is an ensemble technique in machine learning?

ANS: Ensemble technique in machine is a method where multiple models are trained and work together to improve better prediction compared to a single model and helps to reduce the error rate. In machine learning there are mainly three types – 1) Bagging, 2) Boosting, 3) Random Forest. It is used to get better accuracy.

4) What is Gini–impurity index?

ANS: It is a criteria and it is used for building a decision tree by splitting with appropriate features of the dataset and it is used by CART algorithm (classification and regression).

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

ANS:

Feature	Bagging	Boosting
Main Goal	To reduce variance	To reduce variance and bias.
Dependency	Models are built independently.	Models are built sequentially.
Method	Use random sample of the dataset with replacement.	Uses the entire dataset.
Examples	Random forest, Bagging tree	Adaboost, Gradient decent boost, XGboost

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

ANS: Hyperparameter tuning is an advanced step in machine learning to find or measure better accuracy of a prediction. When we trained the model using proper dataset if we get the so much difference of accuracy between training data and testing data in prediction then it is overfitted. When we want to remove the overfitting in the model then we have to use hyperparameter tuning.

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 is a better measure of the goodness of fit for a regression model because it is generally more useful for directly comparing models across different datasets and R-squared also provides the ratio of the variability explained by the model to the total variability.

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

ANS: Yes, we can use Logistic Regression for Classification of Non-Linear Data.

13) Differentiate between Adaboost and Gradient Boosting.

ANS:

Feature	Adaboost	Gradient boost
Model	It is uses very simple model.	It is uses a complex model.
Sensitivity	It is more sensitive boosting technique.	It is less sensitive boosting technique.
Loss function	It uses an exponential loss function.	It uses a differential loss function.
Learning focus	Focuses on classifying training.	Focuses on reducing the residual error of the models.

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

ANS: Out-of-bag error is a method of measuring the performance of a random forest model by calculating the model's error rate on portions of the training data that are not used during the construction of decision tree. This fulfills the need of cross-validation.

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

ANS: Bias-variance trade off is refer to a relation between a model's complexity and the accuracy of its prediction on unseen data that are not used in training.

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

ANS: If gradient decent has a large learning rate, the model may fluctuate around the optimal point rather than reach the optimal solution causing divergence rather than convergence.

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

ANS: Regularization is required in machine learning to avoid overfitting the model. This prevents the model from overly dependent on training data and ensure better performance on new and unseen data.

