

Machine Learning

1. Typically, a smaller or lower RSS value is good in any model because it indicates less fluctuation in the data set. To put it another way, the lower the sum of squared residuals, the better the regression model explains the data.
2. The total sum of squares (TSS) measures the variance in the observed data, whereas the residual sum of squares (RSS) indicates the variation in the error between the observed data and the modelled values. ESS computes the difference between the estimated and mean values of the observed variable.
TSS = ESS + RSS, where TSS stands for Total Sum of Squares, ESS stands for Explained Sum of Squares, and RSS stands for Residual Sum of Squares. Regression Analysis is used to explain the variance of the dependent variable Y.
3. Regularization is a collection of strategies that can prevent overfitting in neural networks and, as a result, enhance the accuracy of a Deep Learning model when confronted with completely new data from the issue domain.
4. The Gini Index, also known as the Gini Impurity, is calculated by subtracting the total of each class's squared probability from one. It prefers larger sections that are easy to deploy. In layman's words, it calculates the likelihood of a randomly selected feature being erroneously labelled.
- 5.
6. The three main classes of ensemble learning methods are **bagging**, **stacking**, and **boosting**
Bagging Ensemble Learning : is an ensemble learning method that seeks a diverse group of ensemble members by varying the training data.
Stacking Ensemble Learning: is an ensemble method that seeks a diverse group of members by varying the model types fit on the training data and using a model to combine predictions.
Boosting Ensemble Learning : is an ensemble method that seeks to change the training data to focus attention on examples that previous fit models on the training dataset have gotten wrong.
7. In general, bagging will want to build an ensemble model with less variance than its components, whereas boosting and stacking will aim to produce strong models that are less biased than their components.
8. The out-of-bag (OOB) error is the average error for each determined using predictions from trees that did not appear in their respective bootstrap sample. This enables the Random Forest Classifier to be fitted and validated while being trained.
9. K-fold Cross-validation occurs when the dataset is divided into K folds and is used to assess the model's ability when presented with new data. K is the number of groups into which the data sample is divided.
10. A hyperparameter is a parameter in machine learning whose value is used to influence the learning process.
It is done because selecting appropriate hyperparameters is crucial to the effectiveness of neural network architecture and has a significant impact on the learned model. If the learning rate is too low, for example, the model will miss crucial patterns in the data. Collisions are possible if it is high.

11. Gradient Descent will not work unless we set the learning rate to a suitable value. This parameter dictates how quickly or slowly we will approach the optimal weights. We will omit the ideal solution if the learning rate is really high.
12. It can be concluded that linear regression is unsuitable for classification problems since it is unbounded and the projected value is continuous rather than probabilistic. Threshold is the parameter that determines whether or not to turn a predicted probability into a class label.
13. AdaBoost is the first designed boosting algorithm with a specific loss function. Gradient Boosting, on the other hand, is a generic technique that aids in the search for approximate solutions to the additive modelling problem. Gradient boosting is thus more adaptable than AdaBoost.
14. Bias refers to the model's simplifying assumptions that are used to make the target function easier to approximate. Variance is the extent by which the target function estimate changes given varied training data. The tension between the mistake generated by the bias and the variance is the trade-off.

Statistics

1. D
2. C
3. C
4. B
5. A
6. B
7. A
8. A
9. B
10. A