MACHINELEARNING

1. Whichofthefollowinginsk-learnlibrary isusedforhyper parametertuning?
   A) GridSearchCV() B)RandomizedCV()
   C)K-foldCrossValidation D)All oftheabove
2. Inwhichofthebelowensembletechniquestreesaretrainedinparallel?
   A) Randomforest B)Adaboost
   C)GradientBoosting D)All oftheabove
3. Inmachinelearning,ifinthebelowlineofcode:
   *sklearn.svm.**SVC**(C=1.0, kernel='rbf',degree=3)*
   weincreasingtheChyperparameter, what willhappen?
   A) Theregularizationwillincrease B) Theregularizationwilldecrease
   C)Noeffect onregularization D) kernelwillbechangedtolinear
4. Check the below line of code and answer the following questions:*sklearn.tree.**DecisionTreeClassifier**(*criterion='gini',splitter='best',max_depth=None, min_samples_split=2)*
   Whichofthefollowingis trueregardingmax_depthhyperparameter?
   A) Itregularizesthedecision treebylimitingthemaximum depthuptowhicha treecanbegrown.
   B) Itdenotesthenumberofchildrenanode canhave.
   C) bothA&B
   D) Noneoftheabove
5. Whichofthefollowingistrueregarding RandomForests?
   A) It'sanensembleofweaklearners.
   B) Thecomponenttreesaretrained inseries
   C) Incaseofclassificationproblem,thepredictionismadebytakingmodeoftheclasslabelspredictedbythe componenttrees.
   D) None of theabove
6. Whatcan bethedisadvantageifthelearningrateis veryhigh in gradientdescent?
   A) GradientDescentalgorithmcandivergefromthe optimalsolution.
   B) GradientDescentalgorithmcankeeposcillatingaroundtheoptimalsolutionandmaynotsettle.
   C) Bothofthem
   D) Noneofthem
7. As themodelcomplexityincreases,whatwill happen?
   A) Biaswillincrease,Variance decrease B) Bias will decrease, Variance increaseC)bothbiasandvarianceincrease D) Both biasandvariance decrease.
8. Suppose I have a linear regression model which is performing as follows:Trainaccuracy=0.95 andTestaccuracy=0.75
   Whichofthefollowingistrueregardingthemodel?
   A) modelisunderfitting B)modelisoverfitting
   C)modelis performinggood D)Noneoftheabove

Q9toQ15aresubjectiveanswertypequestions,Answerthembriefly.

9. SupposewehaveadatasetwhichhavetwoclassesAandB.ThepercentageofclassAis40%andpercentage ofclassB is60%.CalculatetheGiniindex andentropyofthedataset.
   Entropy is 0.97 bits
   Gini index = A/A+B = 0.4

MACHINELEARNING

10. Whataretheadvantages ofRandomForestsover Decision Tree?

Random forests consist of multiple single trees each based on a random sample of the training data. They are typically more accurate than single decision trees. Decision trees have a low bias and are non-parametric, they suffer from a high variance which makes them less useful for most practical applications.

By aggregating multiple decision trees, one can reduce the variance of the model output significantly, thus improving performance. While this could be archived by simple tree bagging, the fact that each tree is build on a bootstrap sample of the same data gives a lower bound on the variance reduction, due to correlation between the individual trees. Random Forest addresses this problem by sub-sampling features, thus de-correlating the trees to a certain extend and therefore allowing for a greater variance reduction / increase in performance.

11. Whatistheneedofscalingallnumericalfeaturesinadataset?Nameanytwotechniquesusedforscaling.

In many machine learning algorithms, to bring all features in the same standing, we need to do scaling so that one significant number doesn't impact the model just because of their large magnitude.

Feature scaling in machine learning is one of the most critical steps during the pre-processing of databefore creating a machine learning model. Scaling can make a difference between a weak machine learning model and a better one.

Two most common techniques of feature scaling are Normalization and Standardization.

MACHINELEARNING

12. Writedownsomeadvantageswhichscalingprovidesinoptimizationusinggradientdescentalgorithm.

    Gradient descent is an optimization algorithm used to minimize the cost function in machine learning algorithms like Logistic Regression, SVM, Neural Networks etc. If features are on different scale, certain weights are updated faster than others in Gradient Descent. However, feature scaling helps in causing Gradient Descent to converge much faster as standardizing all the variables on to the same scale.

13. Incaseofahighlyimbalanceddatasetforaclassificationproblem,isaccuracyagoodmetrictomeasurethepe rformanceofthemodel.Ifnot,why?

    Accuracy Metric is one the simplest and widely used metric to measure the performance of a classification predictive model. The reason for its wide use is because it is easy to calculate, easy to interpret, and is a single number to summarize the model's capability. However, accuracy metric fails to perform on an imbalanced dataset as it gives misleading conclusions. In an imbalanced dataset getting an accuracy score of 90 or 99 are trivial as model might have considered the less numbered observation as error or outliers and could have ignored them in the prediction.

14. Whatis "f-score"metric?Writeitsmathematicalformula.

    F1-score, is a measure of a model's accuracy on a dataset. It is used to evaluate binary classification systems, which classify examples into 'positive' or 'negative'.

    $F1=2*\frac{*precision*recall}{precision+recall}$

15. Whatisthedifferencebetweenfit(),transform()andfit_transform()?

    Fit() method is used to fit the transformer like MinMaxSCaler to the input data and perform the required computations to the specific transformer we apply.

    Transform() method of sklearn transformers, will transform the input data into some transformed spaced. The output is usually an array matrix with equal number of samples as the input data. The transformation will be performed based on the parameters that were computed during fit.

    Fit_transform() method is basically the combination of fit method and transform method, it is equivalent to fit().transform(). This method performs fit and transform on the input data at a single time and converts the data points. If we use fit and transform separate when we need both then it will decrease the efficiency of the model so we use fit_transform() which will do both the work.