MACHINE LEARNING ASSIGNMENT – 7

1. Which of the following in sk-learn library is used for hyper parameter tuning?

Ans=A) GridSearchCV()

1. In which of the below ensemble techniques trees are trained in parallel?

Ans=C) Gradient Boosting

1. In machine learning, if in the below line of code: sklearn.svm.SVC (C=1.0, kernel='rbf', degree=3) we increasing the C hyper parameter, what will happen?

Ans=A) The regularization will increase

1. 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)

Ans=A) It regularizes the decision tree by limiting the maximum depth up to which a tree can be grown.

1. Which of the following is true regarding Random Forests?

Ans=C) In case of classification problem, the prediction is made by taking mode of the class labels predicted by the component trees.

1. What can be the disadvantage if the learning rate is very high in gradient descent?

Ans=B) Gradient Descent algorithm can keep oscillating around the optimal solution and may not settle.

1. As the model complexity increases, what will happen?

Ans=A) Bias will increase, Variance decrease

1. Suppose I have a linear regression model which is performing as follows: Train accuracy=0.95 and Test accuracy=0.75 Which of the following is true regarding the model?

Ans=C) model is performing good

Q9 to Q15 are subjective answer type questions, Answer them briefly.

1. Suppose we have a dataset which have two classes A and B. The percentage of class A is 40% and percentage of class B is 60%. Calculate the Gini index and entropy of the dataset.

Ans=40+60/200=0.5 is gini index entropy of classes.

1. What are the advantages of Random Forests over Decision Tree?

Ans=A random forest is simply a collection of decision trees whose results are aggregated into one final result. Their ability to limit overfitting without substantially increasing error due to bias is why they are such powerful models. One way Random Forests reduce variance is by training on different samples of the data.

1. What is the need of scaling all numerical features in a dataset? Name any two techniques used for scaling.

Ans=Feature scaling is essential for machine learning algorithms that calculate distances between data. Therefore, the range of all features should be normalized so that each feature contributes approximately proportionately to the final distance.

Techniques in scaling:

1)Min-Max scaler=Transform features by scaling each feature to a given range. This estimator scales and translates each feature individually such that it is in the given range on the training set.

2)Standard scaler=The Standard Scaler assumes data is normally distributed within each feature and scales them such that the distribution centered around 0, with a standard deviation of 1.

12. Write down some advantages which scaling provides in optimization using gradient descent algorithm.

Ans=The main advantages:

We can use fixed learning rate during training without worrying about learning rate decay.

It has straight trajectory towards the minimum and it is guaranteed to converge in theory to the global minimum if the loss function is convex and to a local minimum if the loss function is not convex.

13.In case of a highly imbalanced dataset for a classification problem, is accuracy a good metric to measure the performance of the model. If not, why?

Ans=Classification accuracy is a metric that summarizes the performance of a classification model as the number of correct predictions divided by the total number of predictions.

It is easy to calculate and intuitive to understand, making it the most common metric used for evaluating classifier models. This intuition breaks down when the distribution of examples to classes is severely skewed.

Intuitions developed by practitioners on balanced datasets, such as 99 percent representing a skillful model, can be incorrect and dangerously misleading on imbalanced classification predictive modeling problems.

14. What is “f-score" metric? Write its mathematical formula.

Ans=The F-score, also called the F1-score, is a measure of a model’s accuracy on a dataset. It is used to evaluate binary classification systems, which [classify](https://deepai.org/machine-learning-glossary-and-terms/classifier) examples into ‘positive’ or ‘negative’.

The F-score is a way of combining the [precision and recall](https://deepai.org/machine-learning-glossary-and-terms/precision-and-recall) of the model, and it is defined as the [harmonic mean](https://deepai.org/machine-learning-glossary-and-terms/harmonic-mean) of the model’s precision and recall.

The F-score is commonly used for evaluating information retrieval systems such as search engines, and also for many kinds of [machine learning](https://deepai.org/machine-learning-glossary-and-terms/machine-learning) models, in particular in [natural language processing](https://deepai.org/machine-learning-glossary-and-terms/natural-language-processing).

Formula=F-score Formula

The formula for the standard F1-score is the harmonic mean of the precision and recall. A perfect model has an F-score of 1.

F1=2/1/recall+1/precision=2\*precision\*recall/precision+recall

Tp/tp+1/2(fp+fn)

15. What is the difference between fit(), transform() and fit\_transform()?

Ans=The fit() function calculates the values of these parameters. The transform function applies the values of the parameters on the actual data and gives the normalized value. The fit\_transform() function performs both in the same step. Note that the same value is got whether we perform in 2 steps or in a single step.