

Workset 7 Machine Learning

1. D) All of the above
2. A) Random forest
3. B) The regularization will decrease
4. C) both A & B
5. A) It's an ensemble of weak learners.
6. C) Both of them
7. B) Bias will decrease, Variance increase
8. C) model is performing good

Question 9. 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.

Answer

Formula for gini index

$$\text{Gini Index} = 1 - \sum_{i=1}^n (P_i)^2$$

$$\text{Gini} = 1 - ((0.4)^2 + (0.6)^2)$$

$$\text{Gini} = 1 - (.16 + .36)$$

$$\text{Gini} = 1 - .52$$

$$\text{Gini} = 0.48$$

Formula for entropy index

$$\text{Entropy} = -\sum_{i=1}^n p_i * \text{Log}_2(p_i)$$

$$\text{Entropy} = -((.4 * \log_2 .4) + (.6 * \log_2 .6))$$

$$\text{Entropy} = -((.4 * -.397) + (.6 * -.22))$$

$$\text{Entropy} = -((-0.1588) + (-0.132))$$

$$\text{Entropy} = -(-0.2908)$$

$$\text{Entropy} = 0.2908$$

Question 10. What are the advantages of Random Forests over Decision Tree?

Answer

Decision tree is simply supervised machine learning algorithm which uses tree like structure along with decision node which help to reach some specific result.

Random forest as name suggest is the forest of randomly created decision trees, each decision tree calculate its own output and random forest algorithm combine the decision of all trees and provide more accurate results for the problem.

The Random Forest Algorithm combines the output of multiple (randomly created) Decision Trees to generate the final output

Question 11. What is the need of scaling all numerical features in a dataset? Name any two techniques used for scaling.

Answer

Generally data is scaled before feeding it to the train the model. This is done to ensure that gradient descent moves smoothly towards its minima. When features are on the same scale, it helps gradient descent to move quickly to its minima.

For distance based algorithms also, it is advised to scale the data before training it because features in a data set have different magnitudes, so the model can give higher priority to highly magnitude features. This can be avoided by scaling the data so that all features may get equal importance.

Two techniques used for scaling are

Normalization

Standardization

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

Answer

- Scaling ensure that optimization process is more efficient.
- By scaling the features in the data set, machine consider all the features important and without scaling machine may give more importance to high magnitude features as compare to lower magnitude.
- Scaling also helps to prevent the algorithm to stuck to local minima.
- Scaling is done to ensure that gradient decent moves smother to it minima and move quickly to the minima

Question 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?

Answer

Accuracy is not good metric for any classification problem which have highly imbalanced dataset because while training the dataset machine learning algorithm learns more about the majority class and it learn very less about minority class, so most of the time it predict about majority class only. This leds to hide the true performance of the model .

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

Answer

The F-score is a metric used to evaluate the performance of a Machine Learning model. It is also known as F1 Score and F measure. It is the combination of precision and recall both

- Formula $F\text{-score} = 2 * (\text{precision} * \text{recall}) / (\text{precision} + \text{recall})$

Question 15. What is the difference between `fit()`, `transform()` and `fit_transform()`?

Answer

`Fit()` method is used to compute the mean and standard deviation for the given features, so that it may be used further for scaling the features

`Transform()` method is used to perform scaling and transform the data further after applying the fit method.

`fit_transform()` method is the combination of both the methods. Applying both the methods separately is preferred to apply `fit_transform()` method at once.