Machine Learning Worksheet-7

1. D)
2. A)
3. B)
4. C)
5. D)
6. C)
7. B)
8. B)
9. Gini Index: 0.28 Entropy: 0.97
10. The biggest advantage of Random forest is that it relies on collecting various decision trees to arrive at any solution. It reduces overfitting in decision trees and helps to improve the accuracy. Also, it is robust to outliers and automates missing values present in the data. Overall, it has better accuracy than Decision tree.
11. Feature scaling is essential for machine learning algorithms that calculate distances between data. If not scale, the feature with a higher value range starts dominating when calculating distances.
    Two techniques used for scaling:-
    1) Min Max Scaler
    2) Standard Scaler

12. Write down some advantages which scaling provides in optimization using gradient descent algorithm.
    - It makes the training faster.

- It prevents the optimization from getting stuck in local optima.
- It gives a better error surface shape.
- Weight decay and Bayes optimization can be done more conveniently.
- It's also important to apply feature scaling if regularization is used as part of the loss function so that coefficients are penalized appropriately.

13. Accuracy is not a proper measure for imbalanced datasets because the model which is trained has a very good experience of dominant class but not as good experience of the scarce class.

14. In statistical analysis of binary classification, the F-score or F-measure is a measure of a test's accuracy. It is calculated from the precision and recall of the test, where the precision is the number of true positive results divided by the number of all positive results, including those not identified correctly, and the recall is the number of true positive results divided by the number of all samples that should have been identified as positive.

15. fit() : In the fit() method, where we use the required formula and perform the calculation on the feature values of input data and fit this calculation to the transformer. For applying the fit() method we have to use .fit() in front of the transformer object.

    Suppose we initialize the StandardScaler object O and we do .fit() then what will it do that, it takes the feature F and it will just compute the mean (μ) and standard deviation (σ) of feature F. That has happened in the fit method.

    transform() : For changing the data we probably do transform, in the transform() method, where we apply the calculations that we have calculated in fit() to every data point in feature F. We have to use .transform() in front of a fit object because we transform the fit calculations.

    We use the example that is used above section when we create an object of the fit method then we just put it in front of the .transform and transform method uses those calculations to transform the scale of the data points, and the output will we get is always in the form of sparse matrix or array.

    fit_transform(): This 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.

    Suppose, we create the StandarScaler object, and then we perform .fit_transform() then it will calculate the mean(μ) and standard deviation(σ) of the feature F at a time it will transform the data points of the feature F