MACHINE LEARNING-
ASSISGMENT –7

1. D
2. B
3. B
4. C
5. C
6. C
7. B
8. A
9. Gini index (class a)=0.45
   Gni index (class b)=0
   GI (TOTAL)=0.27
10. DT---1. TREE like model of decision along with possible outcome in a day.
    2. There Is scope for over-fitting caused due to the presence of variance.
    3. The results are not accurate.
    4. It takes less time to implement and carrying low accuracy.
    5.it is easy too visualize the task.

    Random forest
    1. It is a classification algorithm consisting of many decision tress combined to get a more accurate result as compared to a single tree.
    2. It avoids and prevent overfitting .
    3. This gives accurate and precious results.
    4. This is time consuming.

11. Scaling is generally performed in the data pre processingstep when working in ML algorthim.

It improves and enhances the performance of the model. We scale the data to comparable ranges to get he proper model and improves the learning of he model.it brings the data to same range ,hence the efficiency of the model improves.
It can be done via Normalization or standardization.

12. Advantages of GD algorithm:
    1.it prevents the algorithm from overshooting the min pt.
    It also avoids oscillations/wild fluctuations around he min pts.
    It leads to more estimates of the gradient descent.

13. Accuracy is not good for unbalanced data set
    In an unbalanced data set and a badly performing model which always predict for the majority class.,it will receive good accuracy score for all the models, but it hides the true performance o the model,as it predict for an one class.
14. F score---is an alternative ML evaluation metrics that assesses the predictives skill of a model by elaboratingon its class wise performance rather than overall performance.it is a measure of test accuracy.
    F=2[p*r]/[p+r]
15. Fit()- method allows usto ge the paramters of the scaling function.

Transform()---it helps to transform thr data set to proceed withfurther data analysis step.
Fit_transform()--- method will determine the paramters and transform the dataset.

SQL---ASSISGMEN 7

1. B
2. B,D
3. C
4. A
5. B
6. C
7. C
8. DML
9. B
10. A
11. Join statement is used to combine data or rows from two or more tables based on a common field between them.
12. Inner join: return records that have matching values in both tables.
    Left join : return all records from left table and the matched records from the right rtable..
    Right join: return all records from the right table and the matched records from the left tables.
13. SQL: server is relational database mangament system developed and marketed by Microsoft.
14. A primary is a column or grp of columns in a table that uniquely identifies the rows of data in that table.
    For eg: customer no, display the Id no assigned to different customr is the primary key.
15. ETL: extract,transform and load are the three process that in combination more data from database, multipe database or other souces to a unified repository, typically a data warehouse.it is a data intregration process.

STATISTICS_. ASSISGMENT 7

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

14. A
15. B

STATISTICS____ASSISGMENT 8

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

MACHINE LEARNING –8

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