

WORKSHEET 7 SQL

1. B. Candidate keys
2. A. Primary keys can contain NULL values, C. A table can have only one primary key with single or multiple fields...
3. C. Insert
4. C. ORDERBY
5. C. SELECT
6. C. 3NF
7. C. All of the above can be done by SQL
8. B. DML
9. B. Table
10. B. 2 NF
11. JOIN is used to combine rows from two or more tables.
12. The different types of joins in SQL are:
 - a. Inner join
 - b. Left join
 - c. Right join
 - d. Full join
13. SQL server is a relational database management system.
14. A primary key is the column that contain the values that uniquely identify each row in a table. It must contain unique values and cannot contain nul values.
15. ELT stands for extract, transform, load. Are the three processes that, in combination, move data from one databases, multiple databases or other sources to a data warehouse.

MACHINE LEARNING

1. D) All of the above
2. A) Random forest
3. A) The regularization will increase
4. C) both A & B
5. B) The component trees are trained in series
6. C) Both of them
7. B) Bias will decrease, Variance increase
8. C) model is performing good
9. Gini index = 0.48
10. Random forest algorithm avoids and prevents overfitting by using multiple trees. Random forest gives accurate results. Random forest is more time consuming.
11. Feature Scaling is a method to transform the numeric features in a dataset to a standard range so that the performance of the machine learning algorithm improves. Normalization and Standardization are techniques used for scaling.
12. Gradient descent is an optimization algorithm which is commonly-used to train machine learning models and neural networks. Training data helps these models learn over time, and the cost function within gradient descent specifically acts as a barometer, gauging its accuracy with each iteration of parameter updates.
13. Accuracy is not a good metric for imbalanced datasets. The model would receive a very good accuracy score as it predicted correctly for the majority of observations, but hides the true performance of the model which is objectively not good as it only predicts for one class.
14. The F-score metric is a measure of a test's accuracy. $F\text{-score} = (2 * \text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})$.
15. The fit () method helps in fitting the data into a model, transform () method helps in transforming the data into a form that is more suitable for the model. Fit_ transform () method, on the other hand, combines the functionalities of both fit () and transform() methods in one step.

STATISTICS WORKSHEET-7

1. b) 0.135
2. d) 0.53
3. c) 0.745
4. b) 0.577
5. 0.3
6. c) 0.56
7. c) 0.33
8. b) 0.22
9. a) 0.66
10. a) 0.33
11. c) 0.5
12. c) 0.78
13. .
14. d) 0.06
15. b) $\frac{2}{3}$