WORKSHEET 2 ANSWER

- MACHINE LEARNING Assignment 2:

Q1. a) 2 Only

Q2. d) 1, 2 and 4

Q3. a) True

Q4. a) 1 only

Q5. b) 1

Q6. b) No

Q7. a) Yes

Q8. d) All of the above

Q9. a) K-means clustering

Q10. d) All of the above

Q11. d) All of the above

Q12. . **Is K sensitive to outliers?**

ANS: K-means can be used as outlier detection. BUT, more attention needs to be given for the definition of outliers. In K-means, using the symmetric distance measure is the key component to define the samples that belonging to the same cluster. symmetric distance measurement gives similar weight to each dimension (feature) this may not always be the case for defining outliers.

Q13. Why is K means better?

ANS: K-means has been around since the 1970s and fares better than other clustering algorithms like density-based, expectation-maximisation. It is one of the most robust methods, especially for image segmentation and image annotation projects. According to some users, K-means is very simple and easy to implement.

Q14. Is K means a deterministic algorithm?

ANS: K-Means is its **non-deterministic nature**. K-Means starts with a random set of data points as initial centroids. This random selection influences the quality of the resulting clusters. Besides, each run of the algorithm for the same dataset may yield a different output. K-means is undoubtedly the most widely used partitional clustering algorithm.

- SQL WORKSHEET Assignment 2

Q1. (D) Unique

Q2. (D) None of the above

Q3. (A) Each entry in the primary key uniquely identifies each entry or row in the table

Q4. (A) There should not be any duplicate entries

Q5. (B) Foreign Key

Q6. (C) 2

Q7. (A) One to many

Q8. (C) One to one

Q9. (A) delivery id

Q10. (D) 2

Q11. (B) Many to one

Q12. (C) Table

Q13. (A) insert into

Q14. (B) Unique (C) Primary Key

Q15. (A) A blood group can contain one of the following values- A, B, AB and O.

(B) A blood group can only contain characters.

- STATISTICS WORKSHEET Assignment-2

Q1. (c) Both

Q2. (c) 12

Q3. (d)

Q4. (c) Both

Q5. (c) Analyzing and interpreting

Q6. (b) Dataset

Q7. (a) 2 or more

Q8. (a), (b) Line graph and Scatterplot (You can choose any one)

Q9. (d) Analysis of variance

Q10. (a) Z-score

Q11. (c) Mean

Q12. (d) 400005.2

Q13. (d) Mean

Q14. (a) Descriptive and inferences

Q15. (d) H-L