QUIZ TOPIC - CLUSTERING

1. **The goal of clustering is to-**

- (●) A. Divide the data points into groups ✔
- ( ) B. Classify the data point into different classes
- ( ) C. Predict the output values of input data points
- ( ) D. All of the above

2. **Clustering is a-**

- ( ) A. Supervised learning
- (●) B. Unsupervised learning ✔
- ( ) C. Reinforcement learning
- ( ) D. None

3. **Which of the following clustering algorithms suffers from the problem of convergence at local optima?**

- ( ) A. K- Means clustering
- ( ) B. Hierarchical clustering
- ( ) C. Diverse clustering
- (●) D. All of the above ✔

4. **Which version of the clustering algorithm is most sensitive to outliers?**

- (●) A. K-means clustering algorithm ✔
- ( ) B. K-modes clustering algorithm
- ( ) C. K-medians clustering algorithm
- ( ) D. None

5. **Which of the following is a bad characteristic of a dataset for clustering analysis-**

- ( ) A. Data points with outliers
- ( ) B. Data points with different densities
- ( ) C. Data points with non-convex shapes
- (●) D. All of the above ✔

6. **For clustering, we do not require-**

- (●) A. Labeled data ✔
- ( ) B. Unlabeled data

C. Numerical data

D. Categorical data

7. **Which of the following is an application of clustering?**

A. Biological network analysis

B. Market trend prediction

C. Topic modeling

D. All of the above ✔

8. **On which data type, we can not perform cluster analysis?**

A. Time series data

B. Text data

C. Multimedia data

D. None ✔

9. **Netflix's movie recommendation system uses-**

A. Supervised learning

B. Unsupervised learning

C. Reinforcement learning ✔

D. All of the above

10. **The final output of Hierarchical clustering is-**

A. The number of cluster centroids

B. The tree representing how close the data points are to each other ✔

C. A map defining the similar data points into individual groups

D. All of the above

11. **Which of the step is not required for K-means clustering?**

A. a distance metric

B. initial number of clusters

C. initial guess as to cluster centroids

D. None ✔

12. **Which is the following is wrong?**

A. k-means clustering is a vector quantization method

B. k-means clustering tries to group n observations into k clusters

C. k-nearest neighbor is same as k-means ✔

D. None

13. **Which of the following uses merging approach?**

- (●) A. Hierarchical clustering ✔
- ( ) B. Partitional clustering
- ( ) C. Density-based clustering
- ( ) D. All of the above

14. **Which of the following is a method of choosing the optimal number of clusters for k-means?**

- ( ) A. cross-validation
- ( ) B. the silhouette method
- ( ) C. the elbow method
- (●) D. All of the above ✔

15. **When does k-means clustering stop creating or optimizing clusters?**

- ( ) A. After finding no new reassignment of data points
- ( ) B. After the algorithm reaches the defined number of iterations
- (●) C. Both A and B ✔
- ( ) D. None

16. **Which of the following clustering algorithm follows a top to bottom approach?**

- ( ) A. K-means
- (●) B. Divisible ✔
- ( ) C. Agglomerative
- ( ) D. None

17. **Which algorithm does not require a dendrogram?**

- ( ) A. K-means
- (●) B. Divisible ✖
- ( ) C. Agglomerative
- ( ) D. All of the above

18. **Which of the following clustering algorithms suffers from the problem of convergence at local optima?**

- ( ) A. Takes each data point as an individual cluster
- ( ) B. Goes on making clusters until it reaches to an optimal number of cluster
- ( ) C. Follows a top to bottom approach

◉ D. All of the above ✔

19. **For topic modeling what should we use?**

○ A. Random forest

○ B. Support vector machine

◉ C. K-means ✔

○ D. K-nearest neighbors

20. **What is a dendrogram?**

◉ A. A hierarchical structure ✔

○ B. A diagram structure

○ C. A graph structure ✖

○ D. None