FACULTY OF COMPUTERS AND INFORMATICS

FINAL EXAMINATION FOR (LEVEL 3)

COURSE TITLE: MACHINE LEARNING

DATE: 5/4/2023	TOTAL ASSESSMENT MARKS: 20	TIME ALLOWED: 60 MINITUES

Choose the correct answer:

1. Which of the following is finally produced by Hierarchical Clustering?
   a) final estimate of cluster centroids b) tree showing how close things are to each other
   c) assignment of each point to clusters d) all of the mentioned
2. Which of the following is required by K-means clustering?
   a) defined distance metric b) number of clusters
   c) initial guess as to cluster centroids d) all of the mentioned
3. Point out the wrong statement.
   a) k-means clustering is a method of vector quantization
   b) k-means clustering aims to partition n observations into k clusters
   c) k-nearest neighbor is same as k-means
   d) none of the mentioned
4. Hierarchical clustering should be primarily used for exploration.
   a) True b) False
5. Which of the following function is used for k-means clustering?
   a) k-means b) k-mean
   c) heatmap d) none of the mentioned
6. Which of the following clustering requires merging approach?
   a) Partitional b) Hierarchical
   c) Naive Bayes d) None of the mentioned
7. K-means is not deterministic and it also consists of number of iterations.
   a) True b) False
8. Unprocessed data or processed data are observations or measurements that can be expressed as text, numbers, or other types of media?
   A) True B) False
9. Amongst which of the following step is performed by data scientist after acquiring the data?
   A) Deletion B) Data Replication C) Data Integration D) Data Cleansing
10. Data integration (include data from multiple sources in your analysis) Integration of multiple
    A) databases B) data cubes C) files
11. 3, 4, 8, 18, 19 ...., 38
    A. 19 B. 22 C. 28 D. 20
12. 4, 7, 11, 18, 29, 47, ..., 123, 199
    A. 76 B. 70 C. 95 D. 105
13. What are some examples of data quality problems:
    A. Noise and outliers B. Duplicate data C. Missing values D. All of the Above
14. Why do we need feature transformation
    A. Converting non-numeric features into numeric B. Resizing inputs to a fixed size
    C. Both A and B D. None
15. **In standardization, the features will be rescaled with –**
    A. Mean 0 and Variance 0 B. Mean 0 and Variance 1
    C. Mean 1 and Variance 0 D. Mean 1 and Variance 1