**SECTION A**

**1. Among the following option identify the one which is not a type of learning**

a) Semi unsupervised learning b) Supervised learning

c) Reinforcement learning d) Unsupervised learning

**2. Identify the kind of learning algorithm for  “facial identities for facial expressions”.**

a) Prediction b) Recognition patterns

c) Recognizing anomalies d) Generating patterns

**3. Identify the model which is trained with data in only a single batch**

a) Offline learning b) Batch learning

c) Both A and B d) None

**4.Identify whether true or false:  In PCA the number of input dimensions is equal to principal components.**

a) True b) False

**5. Among the following identify the one in which dimensionality reduction reduces.**

a) Performance b) Entropy

c) Stochastics d) Collinearity

6. **Machine learning is a subset of which of the following.**

a) Artificial intelligence b) Deep learning

c) Data learning d) None of the above

**7.Which of the following machine learning techniques helps in detecting the outliers in data?**

a) Classification b) Clustering

c) Anomaly detection d) All of the above

**8. The father of machine learning is \_\_\_\_\_\_\_\_\_\_\_\_\_**

a) Geoffrey Everest Hinton b) Geoffrey Hill

c) Geoffrey Chaucer d) None of the above

**9. Choose the general limitations of the backpropagation rule among the following.**

a) Slow convergence b) Scaling

c) Local minima problem d) All of the above

**10. What is the primary objective of K-means clustering?**

a) To maximize the within-cluster similarity b) To minimize the within-cluster variance

c) To minimize the between-cluster variance d) To maximize the between-cluster similarity

**11. In K-means clustering, how is the initial cluster centroid usually chosen?**

a) Randomly b) Based on the mean of all data points

c) Based on the mode of the dataset d) Based on the median of the dataset

**12. Which of the following is a limitation of K-means clustering?**

a) It is sensitive to outliers b) It cannot handle non-linear data

c) It requires a predefined number of clusters d) All of the above

**13. What is the purpose of PCA (Principal Component Analysis)?**

a) To reduce the number of features while preserving most of the variance

b) To increase the number of features for better classification

c) To perform unsupervised clustering

d) To calculate the distance between data points

**14. In PCA, what does the first principal component represent?**

a) It represents the component with the least variance

b) It represents the component with the highest variance

c) It represents the component with the mean variance

d) It represents the component with the median variance

**15. Which of the following is a typical application of PCA?**

a) Image recognition b) Text classification

c) Dimensionality reduction d) Speech synthesis

**16. What is the purpose of an Artificial Neural Network (ANN)?**

a) To perform clustering analysis b) To simulate the human brain's learning process

c) To reduce the dimensionality of data d) To calculate the statistical moments of a dataset

**17. Which of the following is true about the activation function in ANN?**

a) It is used to determine the number of layers in the network

b) It transforms the input signal into an output signal

c) It is only applied to the output layer d) It is not necessary in ANN

**18. What is backpropagation in the context of ANN?**

a) It is the process of feeding input data to the network

b) It is the process of adjusting weights in the network based on prediction errors

c) It is the process of selecting the most important features

d) It is the process of evaluating the performance of the network

**19. Which of the following is a common activation function used in ANN?**

a) Linear function b) Sigmoid function c) Step function d) All of the above

**20. Which step is not involved in the K-means clustering algorithm?**

a) Initialization b) Cluster assignment

c) Dimensionality reduction d) Centroid update

**Section – B (2 marks 15 questions)**

**21. In K-means clustering, what metric is commonly used to calculate the distance between data points and cluster centroids?**

a) Manhattan distance b) Euclidean distance

c) Minkowski distance d) Mahalanobis distance

**22. What is a potential issue with the K-means clustering algorithm when dealing with high-dimensional data?**

a) It becomes computationally less efficient b) It becomes more sensitive to outliers

c) It requires a larger number of clusters d) It tends to overfit the data

**23. Which of the following is not a type of PCA technique commonly used?**

a) Incremental PCA b) Kernel PCA

c) Sparse PCA d) Hierarchical PCA

**24. What does the term "explained variance" refer to in the context of PCA?**

a) The variance explained by the model's predictions

b) The percentage of variance in the original data explained by each principal component

c) The variance of the original dataset

d) The variance of the residuals in the model

**25. In PCA, what is the purpose of eigen values?**

a) They represent the coefficients of the principal components

b) They represent the explained variance of each principal component

c) They represent the correlation between features

d) They represent the distance between data points

**26. Which of the following is a common application of Artificial Neural Networks (ANN)?**

a) Dimensionality reduction b) Image recognition

c) Feature selection d) Hypothesis testing

**27. What is the role of the learning rate in training an Artificial Neural Network?**

a) It determines the number of neurons in each layer

b) It controls the speed at which the model learns during training

c) It determines the number of layers in the network

d) It defines the activation function to be used

**28. Which layer in an Artificial Neural Network is responsible for making predictions?**

a) Input layer b) Hidden layer c) Output layer d) Activation layer

**29. Which of the following techniques is used to prevent overfitting in Artificial Neural Networks?**

a) Regularization b) Dimensionality reduction

c) Feature scaling d) Gradient descent

**30. What is the primary drawback of the K-means clustering algorithm?**

a) It requires a large number of iterations to converge

b) It does not work well with non-linear data

c) It always produces accurate results d) It is not suitable for high-dimensional data

**31. In PCA, what does the term "eigen vector" represent?**

a) The direction of maximum variance in the data

b) The weight assigned to each principal component

c) The mean of the dataset

d) The explained variance of each principal component

**32. Which of the following statements about Artificial Neural Networks is true?**

a) ANN is a supervised learning algorithm b) ANN requires labeled data for training

c) ANN can only have one hidden layer d) ANN does not require any parameter tuning

**33. What is the primary advantage of using PCA before applying K-means clustering?**

a) It reduces the dimensionality of the data

b) It improves the accuracy of cluster assignments

c) It speeds up the convergence of the algorithm

d) It eliminates the need for determining the number of clusters

**34. Which of the following techniques can be used to determine the optimal number of clusters in K-means clustering?**

a) Elbow method b) Silhouette analysis

c) Gap statistic d) All of the above

**35. What is the primary difference between PCA and Factor Analysis?**

a) PCA assumes that all variables are observed without error, while Factor Analysis does not make this assumption.

b) PCA can handle categorical variables, while Factor Analysis cannot.

c) PCA requires the specification of the number of factors, while Factor Analysis does not.

d) PCA produces uncorrelated components, while Factor Analysis produces correlated factors.