**Section – A (1 marks )**

**1. What is the primary function of an activation function in an Artificial Neural Network (ANN)?**

a) To initialize the network weights b) To introduce non-linearity to the network

c) To define the learning rate of the network d) To compute the loss function

**2. Which activation function is commonly used for binary classification tasks in the output layer of an ANN?**

a) ReLU (Rectified Linear Unit) b) Sigmoid

c) Tanh (Hyperbolic Tangent) d) Softmax

**3. What does backpropagation refer to in the context of ANN?**

a) Initializing the network weights b) Adjusting the network weights based on the error

c) Calculating the loss function d) Defining the learning rate of the network

**4. How does dropout regularization prevent overfitting in ANN?**

a) By increasing the number of training epochs b) By randomly removing neurons during training

c) By reducing the learning rate d) By increasing the number of hidden layers

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

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

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

**6. Which step of the K-means algorithm involves initializing cluster centroids?**

a) Assigning data points to clusters b) Computing the mean of data points in each cluster

c) Calculating the distance between data points and centroids d) Initializing cluster centroids

**7. How does the K-means algorithm update cluster centroids in each iteration?**

a) By randomly selecting new centroids from the data points

b) By computing the mean of data points in each cluster

c) By assigning data points to the nearest centroid

d) By adjusting centroids based on a predefined step size

**8. What does the term "K" represent in K-means clustering?**

a) Number of features in the dataset b) Number of iterations for convergence

c) Number of clusters to be formed d) Number of dimensions in the dataset

**9. What is the main objective of Principal Component Analysis (PCA)?**

a) To increase the interpretability of the dataset b) To increase the dimensionality of the dataset

c) To reduce the dimensionality of the dataset while preserving variance

d) To introduce non-linearity into the dataset

**10. How does PCA help in data preprocessing?**

a) By increasing the number of features in the dataset

b) By transforming the data to a lower-dimensional space

c) By normalizing the data to have zero mean and unit variance d) By adding noise to the dataset

**11. Which statement about PCA is true?**

a) PCA is a supervised learning algorithm

b) PCA performs feature extraction by combining original features

c) PCA is robust to outliers in the data d) PCA is used for classification tasks only

**12. What does the explained variance ratio represent in PCA?**

a) The cumulative variance explained by all principal components

b) The percentage of variance explained by each principal component

c) The total variance of the dataset before dimensionality reduction

d) The number of components retained after dimensionality reduction

**13. What role does the learning rate play in training an Artificial Neural Network (ANN)?**

a) It controls the number of hidden layers in the network b) It initializes the weights of the network

c) It adjusts the step size during weight updates

d) It defines the activation function used in the network

**14. How does batch normalization help in training deep neural networks?**

a) By reducing computational complexity b) By normalizing the inputs to each layer

c) By introducing non-linearity into the network d) By increasing the number of training epochs

**15. Which of the following is NOT a commonly used activation function in neural networks?**

a) ReLU (Rectified Linear Unit) b) Sigmoid c) Softmax d) Linear

**16. What does the term "vanishing gradient" refer to in the context of neural networks?**

a) When the learning rate is too high, causing weight updates to oscillate

b) When the gradients become extremely small during backpropagation

c) When the network fails to converge to a stable solution

d) When the network has too few hidden layers to learn complex patterns

**17. What problem can occur due to the sensitivity of K-means clustering to initial centroid placement?**

a) Overfitting b) Underfitting c) Local optima d) Global optima

**18. Which evaluation metric is commonly used to determine the optimal number of clusters in K-means clustering?**

a) Inertia b) Silhouette score c) Davies-Bouldin index d) Adjusted Rand index

**19. What does the "Elbow Method" aim to determine in K-means clustering?**

a) Optimal number of dimensions b) Optimal number of iterations

c) Optimal number of clusters d) Optimal learning rate

**20. How does the silhouette score measure the quality of clustering in K-means?**

a) It measures the compactness and separation of clusters b) It measures the within-cluster variance

c) It measures the between-cluster variance d) It measures the convergence of the algorithm

**21. What is the relationship between eigenvalues and explained variance in PCA?**

a) Eigen values represent the explained variance by each principal component

b) Eigenvalues represent the covariance between features in the dataset

c) Eigenvalues represent the total variance of the dataset

d) Eigenvalues represent the number of principal components

**22. How does PCA help in reducing overfitting in machine learning models?**

a) By increasing the number of features b) By reducing the number of samples

c) By projecting data onto a lower-dimensional space d) By adding noise to the dataset

**23. Which of the following is a limitation of PCA?**

a) It cannot handle categorical features b) It is computationally expensive for large datasets

c) It requires labeled data for training d) It is sensitive to outliers in the data

**24. What is the significance of the cumulative explained variance in PCA?**

a) It measures the proportion of variance explained by each principal component

b) It measures the total variance of the dataset

c) It measures the cumulative proportion of variance explained by the first k principal components

d) It measures the covariance between features in the dataset

**25. What does the term "epoch" refer to in the context of training an Artificial Neural Network (ANN)?**

a) The learning rate used during training b) The number of layers in the network

c) One complete pass of the training dataset through the network

d) The activation function used in the output layer

**26. Which optimization algorithm is commonly used to train deep neural networks?**

a) Gradient Descent b) Random Search

c) Genetic Algorithm d) Newton's Method

**27. What is the purpose of the bias term in a neural network?**

a) To control the learning rate during training b) To initialize the weights of the network

c) To introduce non-linearity into the network d) To shift the activation function

**28. Which of the following is a drawback of the K-means clustering algorithm?**

a) It is sensitive to outliers b) It requires labeled data for training

c) It can only handle numeric data d) It is computationally expensive for large datasets

**29. What is the effect of increasing the number of clusters (K) in K-means clustering?**

a) It increases the within-cluster variance b) It decreases the within-cluster variance

c) It has no effect on the clustering result d) It decreases the between-cluster variance

**30. Which initialization method for cluster centroids is prone to getting stuck in local optima?**

a) Random initialization b) K-means++ c) K-means|| d) K-means||

**31. What is the primary goal of PCA in dimensionality reduction?**

a) To maximize the number of features retained b) To minimize the number of features retained

c) To preserve as much variance as possible d) To increase computational efficiency

**32. In PCA, how are principal components ordered in terms of importance?**

a) In descending order of their eigen values b) In ascending order of their eigenvalues

c) In random order d) In order of their singular values

**33. How does PCA handle multicollinearity in a dataset?**

a) By removing one of the correlated features

b) By combining correlated features into a single component

c) By increasing the number of features

d) By introducing non-linear relationships between features

**34. What is the purpose of the activation function in a neural network?**

a) To initialize the weights of the network b) To introduce non-linearity into the network

c) To define the architecture of the network d) To compute the loss function

**35. Which evaluation metric is commonly used to determine the optimal number of clusters in K-means clustering?**

a) Inertia b) Silhouette score c) Davies-Bouldin index d) Adjusted Rand index