|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **P.V.P Siddhartha Institute of Technology(Autonomous)** | | | | | | **Signature of Invigilator with date:** | | **Marks Obtained:** | |
| **Department of Computer Science and Engineering** | | | | | |
| **Course: B. Tech** | | **Year: III** | **Semester -II** | | **Objective-II** |  | |  | |
| **Regulation: PVP20** | | **Maximum Marks: 10M** | | | **Session: F.N** |
| **A.Y: 2024-25** | | **Date:25-03-2025** | | | **Duration: 20 min** |
| **Subject Code: 20CS3602** | | **Subject Name: Machine Learning** | | | | | | | |
| **Registered Number:** | | | | | **Name:** | | | | |
| **Answer all the Questions. Each Question carries ½ Mark 20×½ M =10M** | | | | | | | | | |
| **S. No** | **Question** | | | | | | **CO** | **Level** | **Answer** |
| **1.** | **The error rate of hypothesis over a sample of data is called\_\_\_\_\_** | | | | | | **CO1** | **L2** | **a** |
| a) Sampling Error | | b) True Error | | | |
| c) Frame Error | | d) Probability Error | | | |
| **2.** | **The confusion matrix visualizes the \_\_\_\_ of a classifier by comparing the actual and predicted classes.** | | | | | | **CO1** | **L2** | **a** |
| a)Accuracy | | b) Connectivity | | | |
| c) Stability | | d) Comparativity | | | |
| **3.** | **During the Treatment of Cancer Patients, the doctor needs to be very careful about which patients needs to be given chemotherapy which metric should we use in order to decide the patients who should give chemotherapy?** | | | | | | **CO1** | **L2** | **b** |
| a) Recall | | b) Precision | | | |
| c) Call | | d) Score | | | |
| **4.** | **Boosting Means\_\_\_\_\_\_\_\_\_\_\_** | | | | | | **CO1** | **L2** | **a** |
| a) New models are affected by the performance of the previously developed  model | | | | | |
| b) Every Model is constructed independently | | | | | |
| c) Not dependent on the model | | | | | |
| d) None of the above | | | | | |
| **5.** | Which of the following is a widely used and effective machine learning algorithm based on the idea of bagging? | | | | | | **CO1** | **L2** | **d** |
| a) Decision Tree | | | b) Linear Regression | | |
| c) Classification | | | d) Random Forest | | |
| **6.** | **Choose the Instance based Learner** | | | | | | **CO1** | **L2** | **b** |
| a) Eager Learner | | | b) Lazy Learner | | |
| c) Both A & B | | | d) None of the Above | | |
| **7.** | In SVM, the dimension of the hyperplane depends upon which one? | | | | | | **CO1** | **L2** | **b** |
| a) No of Samples | | | b) No of Features | | |
| c) No of Target Variables | | | d) All of the Above | | |
| **8.** | In SVM, The Linear Separator Hyper plane formulae is | | | | | | **CO1** | **L2** | **c** |
| a) f(x)=sign(w/x+b) | | | b) f(x)=sign(w+x+b) | | |
| c) f(x)=sign(w.x+b) | | | d) f(x)=sign(w-x+b) | | |
| **9.** | **Which among the following is the most appropriate kernel the that can be used with SVM to separate the classes.** | | | | | | **CO1** | **L2** | **d** |
| a) Linear Kernel | | | b) Sigmoid Function | | |
| c) Polynomial Kernel | | | d) Radial Basis Kernel | | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **10.** | **In k-NN what will happen when you increase/decrease the value of k?** | | | | | | | **CO1** |
| a) The boundary becomes smoother with increasing value of K | | | | | | |
| b) The boundary becomes smoother with decreasing value of K | | | | | | |
| c) Smoothness of boundary doesn’t dependent on value of K | | | | | | |
| d) None of these | | | | | | |
| **11.** | **Which of the following is an Application of Case Based Reasoning** | | | | | | | **CO1** |
| a) Diagnosis | | b) Planning | | c) Design | | d) Analysis |
| **12.** | **Instances are represented by symbols not values are called\_\_\_\_\_\_** | | | | | | | **CO1** |
| a)Neural Network | | | | b) KNN Algorithm | | |
| c) Support Vector Machine | | | | d) Case Based Reasoning | | |
|  |  | | | |  | | |  |
| **1** | **Which of the following best describes unsupervised learning?** | | | | | | | **C** |
| a) Learning from labeled data | | | b) Learning a mapping from input to output | | | |
| c) Grouping data based on similarity | | | d) Predicting categorical outcomes | | | |
| **2** | **In K-Means, what does "K" represent?** | | | | | | | **B** |
| a) Number of iterations | | | b) Number of clusters | | | |
| c) Size of each cluster | | | d) Number of features | | | |
| **3** | **Which evaluation metric is typically used for unsupervised learning tasks like clustering?** | | | | | | | **C** |
| a) Accuracy | | | b) Precision | | | |
| c) Silhouette Score | | | d) Mean Absolute Error | | | |
| **4** | **The points with in the cluster are more close to each other is called\_\_\_\_\_\_\_\_\_** | | | | | | | **A** |
| a) Cohesion | | | b) Forward Propagation | | | |
| c) Separation | | | d) Coupling | | | |
| **5** | **\_\_\_\_\_\_\_\_\_\_consider the clusters as the dense region having some similarity and different from the lower dense region of the space** | | | | | | | **D** |
| a) Project Based | | | b) Hierarchical Based | | | |
| c) Grid Based | | | d) Density Based | | | |
| **6** | **\_\_\_\_\_\_\_ stands for Statistical Information Grid.** | | | | | | | **A** |
| a) STING | | | b) SNG | | | |
| c) STIG | | | d) SIG | | | |
| **7** | **Aggloramative Hierarchical Cluster follows the \_\_\_\_\_\_ Approach** | | | | | | | **A** |
| a) Top Down | | | b) Bottom Up | | | |
| c) Dawn Up | | | d) None | | | |
| **8** | \_\_\_\_\_\_\_\_\_\_clusters formed in this method forms a tree-type structure based on the hierarchy. | | | | | | | **B** |
| a) Project Based | | | b) Hierarchical Based | | | |
| c) Grid Based | | | d) Density Based | | | |
| **9** | **Principal Component Analysis (PCA) is used in unsupervised learning for:** | | | | | | | **C** |
| a) Classification | | | b) Regression | | | |
| c) Dimensionality Reduction | | | d) Overfitting Prevention | | | |
| **10** | **What is the primary goal of unsupervised learning?** | | | | | | | **C** |
| a) Predicting future values | | | b) Classifying labeled data | | | |
| c) Finding hidden patterns in data | | | d) Minimizing prediction error | | | |
| **11** | **The main objective of clustering algorithms in unsupervised learning is to:** | | | | | | | **B** |
| a) Predict continuous values | | | b) Determine the optimal number of clusters | | | |
| c) Assign input data to predefined categories or classes | | | d) Identify patterns in unlabeled data | | | |
| **12** | **Which unsupervised learning algorithm is based on the concept of "nearest neighbors"?** | | | | | | | **B** |
| a) K-means clustering | | | b) Hierarchical clustering | | | |
| c) Naive Bayes | | | d) K-nearest neighbors (KNN) | | | |
| **13** | **Which unsupervised learning algorithm is used for density estimation?** | | | | | | | **C** |
| a) Linear Regression | | | b) Decision Tree | | | |
| c) Gaussian Mixture Models (GMM) | | | d) Support Vector Machines (SVM) | | | |
| **14** | **Which unsupervised learning algorithm is used for data visualization?** | | | | | | | **A** |
| a) PCA | | | b) Decision Tree | | | |
| c) Naïve Bayes | | | d) SVM | | | |
| **15** | **Which of the following is a major limitation of K-Means clustering?** | | | | | | | **B** |
| a) High time complexity | | | b) Sensitivity to the choice of distance metric | | | |
| c) Inability to handle non-numeric data | | | d) Not scalable to large datasets | | | |
| **16** | **Which of the following algorithms can identify clusters of arbitrary shapes?** | | | | | | | **A** |
| a) DBSCAN | | | b) K-Means | | | |
| c) PCA | | | d) Mini Batch K-Menas | | | |
| **17** | The **Silhouette Score** ranges from: | | |  | | | | **A** |
| a) -1 to 1 | b) 0 to 1 | | c) -∞ to ∞ | | d) 0 to ∞ | |
| **18** | **In hierarchical clustering, which linkage method considers the distance between the two closest elements in each cluster?** | | | | | | | **C** |
| a) Complete Linkage | | | b) Average Linkage | | | |
| c) Single Linkage | | | d) Ward Linkage | | | |
| **19** | **What is the time complexity of DBSCAN in its naive form?** | | | | | | | **C** |
| a) O(n) | b) O(nlogn) | | c) O(n2) | | d) O(logn) | |
| **20** | **What is the primary objective of dimensionality reduction in unsupervised learning?** | | | | | | | **D** |
| a) Increase accuracy | | | b) Reduce model complexity | | | |
| c) Reduce overfitting | | | d) Preserve important structures while reducing features | | | |
| **21** | **Which Python library provides an efficient implementation of K-Means, DBSCAN, and PCA?** | | | | | | | **C** |
| a) NumPy | | | b) Pandas | | | |
| c) Scikit-learn | | | d) TensorFlow | | | |
| **22** | **Which of the following is a grid-based clustering algorithm?** | | | | | | | **C** |
| a) K-Means | b) DBSCAN | | c) STING | | d) PCA | |
| **23** | **Which of the following clustering approaches uses both grid partitioning and density concepts?** | | | | | | | **B** |
| a) K-Means | b) CLIQUE | | c) DBSCAN | | d) STING | |
| **24** | **Which parameter is critical in the CLIQUE algorithm to define dense regions?** | | | | | | | **C** |
| a) No of Clusters | | | b) Minimum Cluster Size | | | |
| c) Density Threshold | | | d) Distance Metric | | | |
| **25** | Inertia **in K-Means clustering refers to:** | | | | | | | **B** |
| a) Number of misclassified points | | | b) Total distance of samples to their nearest centroid | | | |
| c) Clustering accuracy | | | d) Sum of squared eigenvalues | | | |