|  |  |
| --- | --- |
| **UNIT IV** | |
| Support Vector Machines and Neural Networks-Support vector machines working principles-Maximum margin classifier- Support vector classifier- Support vector machines- Kernel functions- Artificial neural networks – ANN - Forward propagation and back propagation-Optimization of neural networks-Stochastic gradient descent – SGD- Introduction to deep learning-Solving methodology- Deep learning software | |
| **PART-A (Multiple Choice Questions)** | |
| **Q.**  **No** | **Questions** |
| **1** | Closest Point to the hyper plane are\_\_\_\_\_\_\_\_\_\_\_\_\_  **a. super vectors**  b. Linear vectors  c. Non linear vectors  d. None of these |
| **2** | SVM, which best segregates classes into how many classes?  a. 1  **b. 2**  c. 3  d. 4 |
| **3** | SVM is a supervised [Machine Learning](https://www.sas.com/en_in/insights/analytics/machine-learning.html#:~:text=Machine%20learning%20is%20a%20method,decisions%20with%20minimal%20human%20intervention.) can be used for  a. Regression  b. Classification  **c. Either a or b**  d. None of the above |
| **4** | Which is used to maximize the margin between the two classes?  a. minimum classifier  b. maximum separator  c. minimum separator  **d. maximum classifier** |
| **5** | The observations touching both the positive and negative hyperplanes called as  a. maximum classifier  **b. support vector**  c. minimum classifier  d. minimum vector |
| **6** | Which can be used to work with non-linearly separable data?  a. support vector classifier  b. minimum classifier  c. maximum classifier  **d. kernel trick** |
| **7** | In which function, given the original feature vectors, return the same value as the dot product of its corresponding mapped feature vectors?  **a. Kernel functions**  b. SVM functions  c. Hyperplane functions  d. Linear functions |
| **8** | Which kernels are popularly used, especially with degree 2?  a. Mononomial Kernel  **b. Polynomial Kernel**  c. Trinomial Kernel  d. Multinominal Kernel |
| **9** | The hyperplane with the maximum margin of separation width is called as \_\_\_\_\_\_\_\_\_\_  **a. Maximum margin classifier**  b. Minimum margin classifier  c. Kernel function  d. Super vectors |
| **10** | Which value of gamma in RBF Kernel will give you low bias and high variance solutions?  a. 0  b. 1  c. high  **d. low** |
| **11** | In RBF Kernel, which value of the gamma value gives you a pointed bump in the higher dimensions?  **a. larger**  b. smaller  c. medium  d. 0 |
| **12** | Which function takes a real valued number and squashes into a range between 0 and 1?  a. tanh  b. Relu  c. linear  **d. sigmoid** |
| **13** | Which model the relationship between a set of input signals and output signals using a model derived from a replica of the biological brain, which responds to stimuli from its sensory inputs?  **a. Artificial neural networks (ANNs)**  b Artificial Intelligence  c. Brain  d. Human interface |
| **14** | Which activation function is used in Linear regression?  a. tanh  b. Relu  **c. linear**  d. sigmoid |
| **15** | The number of neurons in the input layer is based on \_\_\_\_\_\_\_\_\_\_\_\_\_  a. dependent variables  **b. independent variables**  c. weight  d. all the above |
| **16** | Which deep learning software is Google'sdeeplearninglibraryrunningontopofPython/C++  a. Caffe  b. Keras  c. Theano  **d. TensorFlow** |
| **17** | In which technique, all training observations are utilized in each iteration for  optimizing the weights of neural networks?  a. Stochastic gradient descent  b. Adadelta  c. Adaptivemomentestimation  **d. Batch gradient descent** |
| **18** | Which utilizes neural networks for building models to solve both supervised and unsupervised problems on structured andunstructureddatasets?  **a. Deep Learning**  b. ANN  c. SVM  d. All the above |
| **19** | Which method is used to solve deep layers by calculating the error of the network atoutput units and propagate back through layers to update the weights to reduce errorterms?  a. Forward propogation  **b. Backwardpropogation**  c. Bipropogation  d. Tripropogation |
| **20** | In the thumb rule in designing Deep Neural networks, allhiddenlayersshouldhavethe\_\_\_\_\_\_\_\_\_numberofneuronsperlayer.  a. 1  b. 2  **3. same**  4. different |
| **21** | The number of neurons in the output layer is decided by the number of \_\_\_\_\_\_\_\_\_ the model needs to be predicted.  a. independent variables  b. weight  **c. classes**  d. dependent variables |
| **22** | Which is a Python-based deep learning library developed by the University ofMontreal?  **a. Theano**  b. Caffe  c. Keras  d. TensorFlow |
| **23** | Which thumb rule improves convergence, in addition totheuseof momentum and dropout in designing deep neural networks?  a. two hidden layers are good enough to solve the majority of problems  **b. Reduction in step size after each iteration**  c. Allhiddenlayersshouldhavethesamenumberofneuronsperlayer  d. Reduction in step size after each iteration in each layer |
| **24** | Which deep learning software is a deeplearninglibraryprimarilyusedfor  processingpictures?  a. Tensorflow  **b. Caffe**  c. Keras  d. Lasagne |
| **25** | In \_\_\_\_\_\_\_\_\_\_ technique,one observation periteration is considered for optimizing the weights of neural networks?  **a. Stochastic Gradient Descent**  b. Adadelta  c. Adaptivemomentestimation  d. Batch gradient descent |

|  |  |
| --- | --- |
| **UNIT V** | |
| K-means clustering-K-means working methodology from first principles- Optimal number of clusters and cluster evaluation - The elbow method- K-means clustering with the iris data example- Principal component analysis - PCA-PCA working methodology from first principles- PCA applied on handwritten digits using scikit-learn- Singular value decomposition – SVD- SVD applied on handwritten digits using scikit-learn | |
| **PART-A (Multiple Choice Questions)** | |
| **Q.**  **No** | **Questions** |
| **1** | 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 above** |
| **2** | \_\_\_\_\_\_\_\_\_\_is the task of grouping observations in such a way that members of the samecluster are more similar to each other and members of different clusters are very differentfrom each other.  a. Classification  b. Regression  **c. Clustering**  d. PCA |
| **3** | Which method is used to determine the optimal number of clusters in k-meansclustering?  **a. Elbow**  b. Euclidean distance  c. Manhattan distance  d. Centroid |
| **4** | What is the name of the measure ofthe compactness and separation of the clusters?  a. k-means coefficient  **b. silhouette coefficient**  c. k-medoid coefficient  d. k-centroid coefficient |
| **5** | Silhouette coefficient values ranges from  **a. -1 to +1**  b. 0 to 1  c.-1 to 0  d. -1 to +2 |
| **6** | Which reduces the dimensions of a dataset by projecting the data onto a  lower-dimensional subspace?  **a. PCA**  b. Regression  c. Clustering  d. Classification |
| **7** | A 2D dataset could be reduced by projecting thepoints onto a \_\_\_\_\_\_\_  a. point  b. plane  c. cube  **d. line** |
| **8** | What are called as the axes (directions) along which a lineartransformation acts simply by stretching/compressing and/or flipping?  a. Eigen values  b. Eigen points  **c. Eigen vectors**  d. Eigen array |
| **9** | What is the importance of using PCA before the clustering? a. Find good features to improve clustering  b. Avoid bad features  c. Find the explained variance  **d. Find which dimension of data maximize the features variance** |
| **10** | In PCA, it transforms the variables into a new set of variables called as \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.  a. linear components  b. scalar components  c. vector components  **d. principal components** |
| **11** | Why is so important to standardize the data in PCA?  **a. Find the features which can predict Y**  b. Make the training time more fast  c. Other people understand better  d. Use the best practices of data wrangling |
| **12** | A three dimensional dataset could be reduced by projecting thepoints onto a \_\_\_\_\_\_\_  a. point  **b. plane**  c. cube  d. line |
| **13** | Which is higher for compact clusters that are well separated andlower for overlapping clusters?  **a. silhouettecoefficient**  b. k-meanscoefficient  c. k-medoid coefficient  d. k-centroid coefficient |
| **14** | Why do you have to drop unimportant features in PCA?  a. find the correct clusters  b. standardize the data  c. To train the models faster  **d. Using the most important features will give better efficiency**  **predicting the target.** |
| **15** | K-means algorithm \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.  a. Requires the dimension of the feature space to be no bigger than the number of samples  b. has the smaller value of the objective function when k=1  **c. minimizes the within class variance for a given number of clusters**  d. converges to the global optimum if and only if the initial means are chosen as some of the samples themselves |
| **16** | What does K refers in the K-means algorithm?  a. fixed value  b. complexity  **c. number of clusters**  d. number of iterations |
| **17** | Which of the following cases will K-means clustering give poor results?  a. Data points with outliers  **b. Data points with different densities**  c. Data points with round shapes  b. Data points with convex shapes |
| **18** | The equivalent of eigenvalues obtained throughthe SVD method are called as \_\_\_\_\_\_.  **a. singular values**  b. singular points  c. singular vectors  d. singular angles |
| **19** | Which is a measure of how much two variables change together and it is a measure ofthe strength of the correlation between two sets of variables?  a. Precision  b. recall  **c. Covariance**  d. SVD |
| **20** | When can you conclude that there will not be any correlation between two sets ofthevariables?  **a. Covariance = 0**  b. Covariance =1  c. Covariance =2  d. Covariance =-1 |
| **21** | SVD can be applied even on which type of matrices?  a. Squared  **b. rectangular**  c. 1 dimensional  d. 3 dimensional |
| **22** | Orthogonal matrices in SVD have \_\_\_\_\_\_\_\_ dimensions.  a. same  b. 1  c. 3  **d. different** |
| **23** | What is the name of the vector obtained equivalent to eigenvectors in SVD?  a. singular values  b. singular points  **c. singular vectors**  d. singular angles |
| **24** | The cost function of k-means is determined by which method?  a. Elbow  **b. Euclidean distance**  c. Manhattan distance  d. Centroid |
| **25** | The eigenvalues are defined only for \_\_\_\_\_\_\_\_\_\_\_\_\_ matrices.  **a. Squared**  b. rectangular  c. 3 dimensional  d. 1 dimensional |