INSTITUTE OF AERONAUTICAL ENGINEERING

(Autonomous)
Dundigal, Hyderabad - 500 043

COMPUTER SCIENCE AND ENGINEERING

DEFINITIONS AND TERMINOLOGY

Course Name	:	**MACHINE LEARNING**
Course Code	:	**ACS014**
Program	:	**B.Tech**
Semester	:	**VIII**
Branch	:	**Computer Science and Engineering**
Section	:	**A, B**
Academic Year	:	**2019 – 2020**
Course Faculty	:	**Mrs. G Sulakshana, Assistant professor, CSE** **Mr. A Praveen, Assistant Professor, IT** **Mrs. B Anupama, Assistant professor, CSE**

OBJECTIVES:

I	Apply knowledge of computing and mathematics appropriate to the discipline.
II	Illustrate the concepts of machine learning and related algorithms.
III	Understand the dimensionality problems using linear discriminants.
IV	Study various statistical models for analyzing the data.
V	Learn clustering algorithms for unlabeled data.

DEFINITIONS AND TERMINOLOGY QUESTION BANK

S.NO	QUESTION	ANSWER	Blooms Taxonomy Level	CO	CLO	CLO Code
	UNIT-I					
1	Define cloud Machine learning	Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed.	Remember	CO 1	CLO 01	ACS014.01
2	Define machine	The advanced mathematics and complex programming at the heart of AI systems is challenging.	Remember	CO 1	CLO 01	ACS014.01
3	Define learning algor ithm	A learning algorithm is an algorithm used in machine learning to help the technology to imitate the human.	Remember	CO 1	CLO 01	ACS014.01
4	What is version space?	A version space is a hierarchical representation of knowledge that enables you to keep track of all the useful information supplied by a sequence of learning	Remember	CO 1	CLO 02	ACS014.01
5	What is Version space learning?	Version space learning is a logical approach to machine learning, specifically binary classification. Version space learning algorithms search a predefined space of hypotheses, viewed as	Remember	CO 1	CLO02	ACS014.02

6	What is Elimination Algorithm?	Elimination Algorithm will converge toward the hypothesis that correctly describes the target concept provided	Remember	CO 1	CLO02	ACS014.02
7	What is candidate elimination algorithm?	The candidate elimination algorithm incrementally builds the version space given a hypothesis space H and a set E of examples.	Remember	CO 1	CLO 02	ACS014.02
8	Define candidate-Elimination algorithm	The candidate-Elimination algorithm computes the version space containing all (and only those) hypotheses from H that are consistent with an observed sequence of training examples.	Remember	CO 1	CLO 02	ACS014.02
9	Define decision tree	Where the data is continuously split according to a certain parameter. The tree can be explained by two entities, namely decision nodes and leaves	Remember	CO 1	CLO 02	ACS014.02
10	What is tree in machine learning?	A tree has many analogies in real life, and turns out that it has influenced a wide area of machine learning, covering both classification	Remember	CO 1	CLO 02	ACS014.02
11	What is Decision Trees?	Decision Trees are a type of Supervised Machine Learning where the data is continuously split according to a certain parameter.	Remember	CO 1	CLO03	ACS014.04
12	What is classification ?	classification is a supervised learning approach in which the computer program learns from the data input given to it and then uses this learning to classify new observation	Remember	CO 1	CLO04	ACS014.04
13	Define Classification	Classification is a supervised learning approach in which the computer program learns from the data input given to it and then uses this learning to classify new observation..	Remember	CO 1	CLO04	ACS014.04
14	What is machine learning classifier?	Classification is the process of predicting the class of given data points. Classes are sometimes called as targets/ labels or categories	Remember	CO 1	CLO04	ACS014.04
	UNIT-II					
1	What is Perceptron?	A Perceptron is a simple model of a biological neuron in an artificial neural network. Perceptron is also the name of an early algorithm for supervised learning of binary classifiers.	Remember	CO 2	CLO 05	ACS014.05
2	What is Perceptron in machine learning?	Perceptron is the fundamental unit of a neural network which is linear in nature capable of doing binary classifications.	Remember	CO 2	CLO 05	ACS014.05
3	Define Perceptron	Perceptron is a machine learning algorithm that helps provide classified outcomes for computing.	Remember	CO 2	CLO 05	ACS014.05
4	What is feed forward neura l network?	The feed forward neural network was the first and simplest type of artificial neural network devised. In this network, the information moves in only one direction, forward, from the input nodes, through the hidden nodes (if any) and to the output nodes	Remember	CO 2	CLO 06	ACS014.05

5	What is Deep Learning: Feed forward Neural Nets?	Deep Learning: Feed forward Neural Nets and Convolutional. Neural Nets Composed of several Perceptron-like units arranged in multiple layers. Consists of an input, Back propagation.	Remember	CO 2	CLO 06	ACS014.06
6	What is the main job of Perceptron?	The Perceptron classifies inputs by finding the dot product of an input feature vector and weight vector and passing that number into a step function, which will return 1 for numbers greater than 0	Remember	CO 2	CLO 07	ACS014.06
7	What is binary classifier?	It is a type of supervised learning, a method of machine learning where the categories are predefined, and is used to categorize new probabilistic observations into said categories.	Remember	CO 2	CLO 07	ACS014.06
8	What is statistical Binary classification ?	When there are only two categories the problem is known as statistical binary classification.	Remember	CO 2	CLO 07	ACS014.07
9	What SVM ?	Support Vector Machine" (SVM) is a supervised machine learning algorithm which can be used for both classification or regression challenges	Remember	CO 2	CLO 07	ACS014.08
10	What is optimum separation hy per plane?	The optimum separation hyper plane (OSH) is the linear classifier with the maximum margin for a given finite set of learning patterns	Remember	CO 2	CLO 07	ACS014.08
11	What is kernel in machine learning?	In machine learning, kernel methods are a class of algorithms for pattern analysis, whose best known member is the support vector machine (SVM).	Remember	CO 2	CLO 07	ACS014.08
12	What is kernel?	Kernel is a way of computing the dot product of two vectors x and y in some (possibly very high dimensional) feature space.	Remember	CO 2	CLO 08	ACS014.08
13	Define kernel	A kernel is a shortcut that helps us do certain calculation faster which otherwise would involve computations in higher	Remember	CO 2	CLO 08	ACS014.08
14	What is kernel trick?	The Kernel Trick is a technique in machine learning to avoid some intensive computation in some algorithms.	Remember	CO 2	CLO 08	ACS014.08
	UNIT-III					
1	What is Variance?	Variance refers to the spread of a data set around its mean value.	Remember	CO 3	CLO09	ACS014.09
2	What is Covariance?	Covariance refers to the measure of the directional relationship between two random variables.	Remember	CO 3	CLO 09	ACS014.09
3	What is the Variance and covariance ?	Variance and covariance are mathematical terms frequently used in statistics and probability theory.	Remember	CO 3	CLO 09	ACS014.09
4	What is Gaussian?	A machine-learning algorithm that involves a Gaussian process uses lazy learning and a measure of the similarity between points (the kernel function) to predict the value for an unseen point from training data	Remember	CO 3	CLO 10	ACS014.09
5	What is Gaussian processes?	Gaussian processes are a powerful algorithm for both regression and classification.	Remember	CO 3	CLO 10	ACS014.10

6	What Gaussian distribution ?	Gaussian distribution model, often identified with its iconic bell shaped curve, also referred as Normal distribution.	Remember	CO 3	CLO 10	ACS014.10
7	What is a Gaussian model?	In probability theory and statistics, a Gaussian process is a stochastic process such that every finite collection of those random variables has a multivariate normal distribution, i.e. every finite linear combination of them is normally distributed.	Remember	CO 3	CLO 10	ACS014.10
8	What Gaussian signal?	In mathematics, a Gaussian function, often simply referred to as a Gaussian, is a function of the form: for arbitrary real constants a , b and non zero c	Remember	CO 3	CLO10	ACS014.10
9	What is the purpose of Bayes theorem in machine learning?	Essentially, the theorem allows a hypothesis to be updated each time new evidence is introduced	Remember	CO 3	CLO 11	ACS014.10
10	What is the benefit of naïve Bayes?	The Naive Bayes algorithm affords fast, highly scalable model building and scoring. It scales linearly with the number of predictors and rows.	Remember	CO 3	CLO 11	ACS014.11
11	How does Bayes classifier work	Naive Bayes is a kind of classifier which uses the Bayes Theorem. It predicts membership probabilities for each class such as the probability that given record or data point belongs to a particular class	Remember	CO 3	CLO 12	ACS014.12
12	What is the difference between hypervisor and virtual machine?	It is a classification technique based on Bayes' Theorem with an assumption of independence among predictors. In simple terms, a Naive Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature.	Remember	CO 3	CLO12	ACS014.12
	UNIT-IV					
1	What is genetic algorithm ?	A genetic algorithm is a search heuristic that is inspired by Charles Darwin's theory of natural evolution. This algorithm reflects the process of natural selection where the fittest individuals are selected for reproduction in order to produce offspring of the next generation	Remember	CO 4	CLO13	ACS014.13
2	What is genetic algorithm in neural network?	Genetic Algorithms (GAs) are search-based algorithms based on the concepts of natural selection and genetics. Gases are a subset of a much larger branch of computation known as Evolutionary Computation.	Remember	CO 4	CLO 13	ACS014.13
3	Why does genetic algorithm work?	A genetic algorithm solves optimization problems by creating a population or group of possible solutions to the problem. The genetic algorithm similarly occasionally causes mutations in its populations by randomly changing the value of a variable	Remember	CO 4	CLO 13	ACS014.14
4	What is genetic algorithm in optimization ?	A genetic algorithm (GA) is a method for solving both constrained and unconstrained optimization problems based on a natural selection process that mimics biological evolution.	Remember	CO 4	CLO 13	ACS014.14

5	Why mutation is important in genetic algorithm?	Mutation (genetic algorithm) Mutation is a genetic operator used to maintain genetic diversity from one generation of a population of genetic algorithm chromosomes to the next.	Remember	CO 4	CLO 13	ACS014.14
6	What is fitness in genetic algorithm?	The fitness function simply defined is a function which takes a candidate solution to the problem as input and produces as output how "fit" our how "good" the solution is with respect to the problem in consideration.	Remember	CO 4	CLO14	ACS014.15
7	What is simple genetic algorithm?	A genetic algorithm is an algorithm that imitates the process of natural selection. They help solve optimization and search problems. Genetic algorithms imitate natural biological processes, such as inheritance, mutation, selection and crossover.	Remember	CO 4	CLO 14	ACS014.15
8	Define Elitism	Elitism refers to a method for improving the GA performance; the basic idea is to transfer the best individuals of the current generation to the next generation.	Remember	CO 4	CLO 15	ACS014.15
9	What is convergence in genetic algorithm?	Convergence is a phenomenon in evolutionary computation. It causes evolution to halt because precisely every individual in the population is identical. Full convergence might be seen in genetic algorithms using only crossover	Remember	CO 4	CLO 13	ACS014.16
10	What is search space in genetic algorithm?	In this search space, lies a point or a set of points which gives the optimal solution. The aim of optimization is to find that point or set of points in the search space.	Remember	CO 4	CLO 15	ACS014.16
11	What is crossover and mutation in genetic algorithm?	The change of parts of one solution randomly, which increases the diversity of the population and provides a mechanism for escaping from a local optimum.	Remember	CO 4	CLO 16	ACS014.16
12	What is hybrid genetic algorithm?	Genetic algorithms (GAs) are iterative optimization procedures that repeatedly apply GA operators (such as selection, crossover, and mutation) to a group of solutions until some criterion of convergence has been satisfied	Remember	CO 4	CLO 16	ACS014.16
	UNIT-V					
1	Define clustering	Cluster analysis or clustering is the task of grouping a set of objects in such a way that objects in the same group (called a cluster) are more similar (in some sense) to each other than to those in other groups (clusters).	Remember	CO 5	CLO 17	ACS014.17
2	What are clustering and its use?	Clustering is the task of dividing the population or data points into a number of groups such that data points in the same groups are more similar to other data points in the same group than those in other groups	Remember	CO 5	CLO 17	ACS014.17
3	What is good clustering?	The quality of a clustering method is also measured by its ability to discover some or all of the hidden patterns	Remember	CO 5	CLO 17	ACS014.17
4	Why is clustering unsupervised	Clustering" is the process of grouping similar entities together. The goal of this unsupervised machine learning technique is	Remember	CO 5	CLO 18	ACS014.17

		to find similarities in the data point and group similar data points together. Grouping similar entities together help profile the attributes of different groups.				
5	What are clustering techniques?	Clustering methods are used to identify groups of similar objects in a multivariate data sets collected from fields such as marketing, bio-medical and geo-spatial	Remember	CO 5	CLO 18	ACS014.17
6	What is clustering in data analysis?	A cluster of data objects can be treated as one group. While doing cluster analysis, we first partition the set of data into groups based on data similarity and then assign the labels to the groups.	Remember	CO 5	CLO18	ACS014.17
7	Why K means clustering is used?	K-means clustering is a type of unsupervised learning, which is used when you have unlabeled data (i.e., data without defined categories or groups). The goal of this algorithm is to find groups in the data, with the number of groups represented by the variable K	Remember	CO 5	CLO 19	ACS014.17
8	What is difference between clustering and classification ?	The prior difference between classification and clustering is that classification is used in supervised learning technique where predefined labels are assigned to instances by properties, on the contrary, clustering is used in unsupervised learning where similar instances are grouped, based on their features	Remember	CO 5	CLO 19	ACS014.17
9	What is the purpose of cluster analysis?	Cluster Analysis has been used in marketing for various purposes. Segmentation of consumers in cluster analysis is used on the basis of benefits sought from the purchase of the product. It can be used to identify homogeneous groups of buyers	Remember	CO 5	CLO 19	ACS014.18
10	What is clustering in ML?	Clustering is the task of dividing the population or data points into a number of groups such that data points in the same groups are more similar to other data points in the same group and dissimilar to the data points in other groups	Remember	CO 5	CLO 19	ACS014.19
11	How does cluster analysis work?	Cluster analysis is also called segmentation analysis or taxonomy analysis. More specifically, it tries to identify homogenous groups of cases if the grouping is not previously known.	Remember	CO 5	CLO 20	ACS014.20
12	Why do we use clustering?	Clustering is the task of dividing the population or data points into a number of groups such that data points in the same groups are more similar to other data points in the same group than those in other groups.	Remember	CO 5	CLO20	ACS014.20

13	What is clustering in data structure?	Clustering is a process of partitioning a set of data (or objects) into a set of meaningful sub-classes, called clusters. Help users understand the natural grouping or structure in a data set. Used either as a stand-alone tool to get insight into data distribution or as a preprocessing step for other algorithms issues.	Remember	CO 5	CLO 20	ACS014.20

Signature of the Faculty **Signature of HOD**