Program: BE Computer Engineering

Curriculum Scheme: Revised 2016

Examination: Third Year Semester VI

Course Code: CSDLO6021 and Course Name:MACHINE LEARNING

Time: 1 hour Max. Marks: 50

==========

Note to the students:- All the Questions are compulsory and carry equal marks .

Q1.	Neural networks:
Option A:	Optimize a convex objective function
Option B:	Can use a mix of different activation functions
Option C:	are not suitable for learning.
Option D:	Can only be trained with stochastic gradient descent
Q2.	The amount of output of one unit received by another unit depends on what?
Option A:	output unit
Option B:	input unit
Option C:	activation value
Option D:	weight
Q3.	Negative sign of weight indicates?
Option A:	excitatory input
Option B:	inhibitory input
Option C:	excitatory output
Option D:	inhibitory output
Q4.	Let us implement a single neuron with threshold activation function to simulate working of logical AND gate.Give the correct values of weights and threshold.
Option A:	w1=1,w2=-1,T=-1
Option B:	w1=-1,w2=-1,T=-1
Option C:	w1=1,w2=1,T=2
Option D:	w1=-1,w2=1,T=-2
Q5.	Which of the following is not a clustering algorithm?
Option A:	EM-Algorithm
Option B:	K-means clustering
Option C:	Radial Basis Function
Option D:	Decision Tree

Q6.	A and B are two events. If P(A, B) decreases while P(A) increases, which of the following is true?
Option A:	P(A|B) decreases
Option B:	P(B|A) decreases
Option C:	P(B) decreases
Option D:	P(B|A) increases
Q7.	Given a large dataset of medical records from patients suffering from heart disease, try to learn whether there might be different clusters of such patients for which we might tailor separate treatments. What kind of learning problem is this?
Option A:	Supervised learning
Option B:	Unsupervised learning
Option C:	Reinforcement learning
Option D:	Semi-Supervised Learning
Q8.	You are given a labeled binary classification data set with N data points and D features. Suppose that N < D. In training an SVM on this data set, which of the following kernels is likely to be most appropriate?
Option A:	Linear kernel
Option B:	Quadratic kernel
Option C:	Higher-order polynomial kernel
Option D:	RBF kernel
Q9.	Which algorithm is State Transition Based Algorithm?
Option A:	K-Nearest neighbor
Option B:	Hidden markov model
Option C:	Bayes theorem
Option D:	Linear regression
Q10.	Below are the 8 actual values of the target variable in the train file.[0,0,0,1,1,1,1,1]What is the entropy of the target variable?
Option A:	-(5/8 log(5/8) + 3/8 log(3/8))
Option B:	5/8 log(5/8) + 3/8 log(3/8)
Option C:	3/8 log(5/8) + 5/8 log(3/8)
Option D:	5/8 log(3/8) – 3/8 log(5/8)
Q11.	What is the major component of PCA?
Option A:	all the eigen vectors for the projection space
Option B:	The average of eigen vectors for the projection space
Option C:	Value of the last among the eigen vectors for the projection space
Option D:	Value of the first among the eigen vectors for the projection space
Q12.	Principal component analysis is a technique for performing
Option A:	Dimensionality reduction
Option B:	Pruning

Option C:	Aggregation
Option D:	Sampling
Q13.	Which of the following option(s) is / are true? 1.You need to randomize parameters in PCA 2.You don't need to randomize parameters in PCA 3.PCA can be trapped into local maxima problem 4.PCA can't be trapped into local minima problem
Option A:	1 and 3
Option B:	1 and 4
Option C:	2 and 3
Option D:	2 and 4
Q14.	Which of the following techniques would perform worst for reducing dimensions of a data set?
Option A:	Removing columns which have high variance in data
Option B:	Removing columns which have too many missing values
Option C:	Removing columns with redundant data
Option D:	Removing columns with similar data trends
Q15.	Which of the following is not supervised learning algorithm
Option A:	PCA
Option B:	Decision Tree
Option C:	Bayes Theorem
Option D:	Linear regression
Q16.	A machine learning model gives 95% accuracy on an unbalanced dataset. What can be concluded about the classifier?
Option A:	Since accuracy is 95% the classifier will perform well in real life scenario
Option B:	Classifier will give good accuracy on the validation of the dataset.
Option C:	Unbalanced Dataset will not affect the performance of classifier
Option D:	Because of an unbalanced dataset the classifier will predict only one class of samples accurately.
Q17.	Predicting on whether it will rain or not tomorrow evening at a particular time is a type of _____ problem.
Option A:	Clustering
Option B:	Regression
Option C:	Unsupervised learning
Option D:	Supervised learning
Q18.	You ran gradient descent for 20 iterations with learning rate=0.2 and compute cost for each iteration. You observe that cost decreases after each iteration. Based on this which conclusion is more suitable.
Option A:	0.2 is an effective choice of learning rate.
Option B:	Try larger values of learning rate like 1.

Option C:	0.2 is not an effective choice of learning rate.
Option D:	The model is overfitting.
Q19.	The process of obtaining best result under given constraints is called as
Option A:	Optimization
Option B:	Generalization
Option C:	Summation
Option D:	Regularization
Q20.	Which of the following is not example of Derivative free optimization
Option A:	Random Search Method
Option B:	Downhill simplex method
Option C:	Genetic algorithm
Option D:	Steepest Descent
Q21.	Multiple regression model has
Option A:	Only one independent variable
Option B:	More than one dependent variables
Option C:	More than one independent variables
Option D:	Only one dependent variable
Q22.	Which statement is true about regression problems?
Option A:	Output attribute must be only categorical.
Option B:	Output attribute must be only numerical
Option C:	Output attribute can be either categorical or numerical.
Option D:	Output attribute can be neither categorical nor numerical.
Q23.	The average positive difference between computed and desired outcome values.
Option A:	root mean squared error
Option B:	mean squared error
Option C:	mean absolute error
Option D:	mean positive error
Q24.	Which of the following is a disadvantage of decision trees?
Option A:	Decision trees require less preprocessing.
Option B:	Decision trees are robust to outliers.
Option C:	Decision trees are prone to be overfit.
Option D:	Decision tree traces all possible alternatives.
Q25.	What is the approach of the basic algorithm for decision tree induction?
Option A:	Greedy
Option B:	Bottom up
Option C:	Procedural
Option D:	Step by Step