GURU NANAK INSTITUTE OF TECHNOLOGY
An Autonomous Institute under MAKAUT
2022
MACHINE LEARNING
IT701D

TIME ALLOTTED: 3 Hours **FULL MARKS: 70**

The figures in the margin indicate full marks.
Candidates are required to give their answers in their own words as far as practicable

GROUP – A
(Multiple Choice Type Questions)

Answer any *ten* from the following, choosing the correct alternative of each question: 10×1=10

		Marks	CO No.
1. (i)	What is machine learning? a. Machine learning is the science of getting computers to act without being explicitly programmed. b. Machine Learning is a Form of AI that Enables a System to Learn from Data. c. Both A and B d. None of the above	1	CO1
(ii)	Which of the following is true about Naive Bayes? a. Assumes that all the features in a dataset are equally important b. Assumes that all the features in a dataset are independent c. Both A and B d. None of the above	1	CO2
(iii)	Supervised learning and unsupervised clustering both require at least one a. hidden attribute b. output attribute c. input attribute d. categorical attribute	1	CO4
(iv)	Which machine learning models are trained to make a series of decisions based on the rewards and feedback they receive for their actions? a. Supervised learning b. Unsupervised learning c. Reinforcement learning d. All the above	1	CO3
(v)	Which of the following is a reasonable way to select the number of principal components "k"? a. Choose k to be the smallest value so that at least 99% of the variance is retained b. Choose k to be 99% of m (k = 0.99*m, rounded to the nearest integer) c. Choose k to be the largest value so that 99% of the variance is retained d. Use the elbow method	1	CO3

(vi)	Which of the following is a disadvantage of decision trees? a. Factor analysis b. Decision trees are robust to outliers c. Decision trees are prone to be overfit d. None of the above	1	CO3
(vii)	In which of the following cases will K-means clustering fail to give good results? 1) Data points with outliers 2) Data points with different densities 3) Data points with nonconvex shapes a. 1 and 2 b. 2 and 3 c. 1, 2, and 3 d. 1 and 3	1	CO1
(viii)	_____is a widely used and effective machine-learning algorithm based on the idea of bagging. a. Regression b. Classification c. Decision Tree d. Random Forest	1	CO3
(ix)	How can you prevent a clustering algorithm from getting stuck in bad local optima? a. Set the same seed value for each run b. Use multiple random initializations c. Both A and B d. None of these	1	CO3
(x)	PCA is a. forward feature selection b. backward feature selection c. feature extraction techniques d. All of the above	1	CO1
(xi)	What is the purpose of performing cross-validation? a. To assess the predictive performance of the models b. To judge how the trained model performs outside the sample on test data c. Both a & b d. None of the above	1	CO3
(xii)	Which of the following is a widely used and effective machine learning algorithm based on the idea of bagging? a. Decision Tree b. Regression c. Classification d. Random Forest	1	CO3

GROUP – B
(Short Answer Type Questions)
(Answer any three of the following) **3 x 5 = 15**

		Marks	CO No.
2.	What is Linear Regression? Explain	5	CO3
3. a)	Define Machine Learning?	2	CO2
b)	Explain different types of Machine Learning	3	CO2
4.	Explain Support Vector Machine.	5	CO2
5.	Compare Artificial Neuron with Biological Neuron with diagram.	5	CO3
6.	Mention How can you choose classifier base in training set size?	5	CO4

GROUP – C
(Long Answer Type Questions)
(Answer any three of the following) 3 x 15 = 45

		Marks	CO No.
7. a)	What is decision boundary? Mention the benefit of decision boundary in machine learning.	3	CO1
b)	What are Markov Models? What is Markov Chain? Explain it with diagram.	6	CO3
c)	What are Hidden Markov models (HMM)? Explain it with diagram.	6	CO1
8. a)	What is K-Means Clustering?	2	CO1
b)	Write down the algorithm of K-Means Clustering and explain it with proper example?	8	CO2
c)	Explain different types of clustering with proper examples.	5	CO3
9. a)	State and explain Bayes' theorem?	3	CO2
b)	Meningitis causes the patients to have neck stiffness 50% of the time. Find the probability of Meningitis given neck stiffness, provided probability of Meningitis is 0.002% and probability of neck stiffness is 5%.	5	CO3
c)	Explain Bayesian Network (BN). Explain joint probability with suitable example.	7	CO2
10. a)	State and explain k-mean clustering algorithm.	7	CO2
b)	BABUJI is deciding which courses he wants to take his next semester. The probability that he enrolls in an AI course is 40% and the probability that he enrolls in ML course is 60%. The probability that he will enroll in an AI course GIVEN that he enrolls in ML course is 30%.	8	CO3
	i. What is the probability he will enrol in AI course AND ML course.		
	ii. What is the probability he will enrol in AI course OR ML course.		
	iii. Are the two events independent?		
	iv. Are the two events mutually exclusive?		
11.	Write short notes on any three of the following:	3x5=15	
a)	K-N-N algorithm.	5	CO2
b)	Random Forest algorithm.	5	CO4
c)	Support Vector Machine.	5	CO2
d)	Over fitting problem	5	CO4
e)	logistic Regression	5	CO2