KIIT Deemed to be University
Online End Semester Examination(Autumn Semester-2021)

Subject Name & Code: Machine Learning CS 3035
Applicable to Courses: B.Tech CSSE

Full Marks=50
Time:2 Hours

SECTION-A(Answer All Questions. Each question carries 2 Marks)

Time:30 Minutes
(7×2=14 Marks)

Question No	Question Type (MCQ/SAT)	Question	CO Mapping	Answer Key (For MCQ Questions only)
Q.No:1	MCQ	A.____1. What is Machine learning? a) The set of techniques for extracting models from data b) The autonomous acquisition of knowledge through the use of manual programs c) The selective acquisition of knowledge through the use of computer programs d) The selective acquisition of knowledge through the use of manual programs	1	A
	MCQ	What is true about Machine Learning? A. Machine Learning (ML) is that field of computer science B. ML is a type of artificial intelligence that extract patterns out of raw data by using an algorithm or method. C. The main focus of ML is to allow computer systems learn from experience without being explicitly programmed or human intervention. D. All of the above	1	D

	MCQ	ML is a field of AI consisting of learning algorithms that? A) Does not much improve the performance B) Decreases the execution time C) does not requires past experience D) Improve their performance (P), At executing some task (T), Over time with experience (E).	1	D
	MCQ	Choose the options that are correct regarding machine learning (ML) and artificial intelligence (AI), (A) ML is an alternate way of programming intelligent machines. (B) ML is a set of techniques that turns a data-set into a software. (C) AI is a software that can emulate the human mind. (D) All the above	1	D
Q.No:2	**MCQ**	Which evaluation matric can be used for logistic regression? A. Mean squared error B. Root mean squared error C. Average mean square error D. AUC-ROC	5	D
		Variable selection is responsible in improving the computational efficiency. From the following which one can be used for variable selection. A. Ridge B. Lasso C. Both of the above D. None of the above	5	B
		Logistic Regression used mainly for A. Regression B. Feature extraction C. Classification D. Clustering	5	C

		Which of the following is used as the learning algorithm for Logistic Regression A. Maximum Liklihood B. Minimul Liklihood C. Gradient decent D. None of the above	5	A
Q.No:3	**MCQ**	Which of the following re-scales the observation values between 0 and 1. A. Euclidean distance B. Manhattan Distance C. Min-Max Normalization D. None of the above	3	C
	MCQ	Which of the following distance metric can not be used in k-NN? A) Manhattan B) Minkowski C) Tanimoto D) Jaccard E) Mahalanobis F) All can be used	3	F
	MCQ	Manhattan distance can be used for A. Continuous values B. Categorical values C. Both Continuous and categorical D. For constant values	3	A
	MCQ	When data set is having categorical values in K-NN which of the following distance measure is useful. A. Manhattan B. Hamming distance C. Euclidean D. Both Manhattan and Euclidean	3	A
Q.No:4	**MCQ**	K- NN is a classier ML algo. It has some disadvantages. Which of the following is	5	C

		correct. A. It has low accuracy B. It is insensitive to outlier C. It is computationally expansive D. Its needs very less memory		
		KNN is known to exhibit lazy learning behaviour. It does more computation during A. Test time B. Train time C. Finding k-values D. None of the above	5	A
		For imputing the missing values of both categorical and continuous variable which of the following is preferred most A. K-means B. K-NN C. Regression D. ANN	5	B
		Which of the following is true for K-NN A. Can be used for classifications B. Can be used for Regression C. Can be used for both classification and Regression D. Can be used for co-relation	5	C
Q.No:5	**MCQ**	Clustering is a class of unsupervised algorithm. How many minimum variables we need to perform clustering. A. 0 B. 1 C. 2 D. 3	5	B
		K- means can fall in local optima problem. Who from the following is the reason for such case. A. High number of features in the dataset B. Random initialization of centroids C. Large dataset	5	B

		D. Distance metrics		
		The central objective of the K-means is to A. Maximize the intra cluster distance B. Minimize the intra cluster distance C. Minimize the inter cluster distance D. None of the above	5	B
		Data scientists are very particular about detecting outliers. Which of the following can be helpful to detect outlier. A. K-median clustering B. K-means clustering C. K-modes Clustering D. None of the above	5	B
Q.No:6	**MCQ**	Input x is the input to the perceptron rule. The desired output is t and the actual output is y. If the learning rate is p then the weight updation formula shall be A. w = w + (t-y) B. w = w +p(t-y) C. w = w + p(t-y)x D. w = w+ p(t-y)	6	C
		Which of the following is not the promise of artificial neural network? A. It can explain result B. It can survive the failure of some nodes C. It has inherent parallelism D. It can handle noise	6	A
		There are two different types of weight updating Which of the followings is accurate A. Epoch B. Online C. None of the above	6	B
		Convergence of the ANN towards arriving its optimal values can be done by using A. learning parameter B. Momentum factor C. Back propagation	6	B

		learning algo D. None of the above		
Q.No:7	**MCQ**	We are dealing with samples x where x is a single value. Two regression models given below are being tested for sample data x which takes single value. 1) $b = ax + e$ 2) $b = ax + cx^2 + e$ Which of the two models is more likely to fit the test data better? A) model 1 B) model 2 C) both will equally fit D) impossible to decide	6	D
		When you compare the Decision tree with logistic regression, the serious weakness of decision tree is A. Decision tree is more likely to over-fit the date B. Decision tree is more likely to under-fit the date C. Decision trees do not assume independence of the input features D. None of the above	6	A
		Least Squares Estimation minimizes: A. summation of squares of errors B. summation of errors C. summation of absolute values of errors D. All	6	A
		K- fold cross-validation is (A) linear in K (B) quadratic in K (C) cubic in K (D) exponential in K	6	A

SECTION-B(Answer Any Three Questions. Each Question carries 12 Marks)

Time: 1 Hour and 30 Minutes
(3×12=36 Marks)

Question No	Question	CO Mapping (Each question should be from the same CO(s))
Q.No:8	A. Explain different types of learning.[2] B. How do we Handel non-linear separable data in SVM. Explain the process with suitable method. [6] C. List the issues in k-means[4]	3
	A. Write down the expression for RBF and Polynomial Kernel. [2] B. Explain one suitable algorithm for partitional clustering[6] C. Differentiate between classification and regression with suitable example [4]	
	A. What types of learning is used in computer game development Explain with example.[2] B. How will you measure that the cluster is a good cluster. Explain with an example with reference to K-means clustering?[6] C. What do you mean by Hyperplane. How the hyperplane is created. Explain with neat diagram [4]	
Q.No:9	A. Draw a comparison between Linear regression and Logistic Regression [4] B. Least square is not a suitable learning method in case of Logistic Regression . Prove or Disprove with suitable argument. [8]	4,5

	A. Derive the co-efficients of Liner regression model using Least Square estimation. [4] B. What are the purpose and requirements of Liner Regression?What are the metrics for judging accuracy of predictions? [8]	
	A. How does the Maximum Likelihood estimation help in computing the parameter of logit function? [4] B. What are the Evaluation of model estimators of Liner regression? Explain any tow with suitable example.[8]	
Q.No:10	A. Map the relationship between entropy, information gain and feature selection. [4] B. Explain L1 and L2 regularization with detailed mathematical equations and explain its importance.[8]	3,4,5
	A. What is the significance of Information gain?[4] B. Why is SVM more accurate than Logistic regression? Explain with suitable example. [8]	
	A. Using neat diagram show that basic AND,OR,NAND and NOR are linearly separable, while XOR is not . [4] B. What is bias and variance? Why a proper trade-off is necessary between the two? Take suitable ML algorithm and explain. [8]	
Q.No:11	A. Construct a NN with 6 inputs terminals , a hidden layer of 3 neurons and an output layer of 2 neuron. [2] B. Derive the mathematical model for the above ANN in (A) with sigmoid transfer function. [6] C. Explain back-propagation learning algo. [4]	6

	A.Explain the principle of batch training and case training protocols for training ANN [2] B.Draw an ANN architecture for 3 inputs, 2 hidden layer each with 2 neurons and one output. Derive the estimated output using sigmoid transfer function. [6] C. Explain under what circumstances the back-propagation learning algo.will be trapped in local optima and what can be done to address this problem. [4]	
	A. What do mean by feed-forward and feed- back in ANN. [2] B.Derive a mathematical model for ANN model. Choose suitable input, hidden layer and output neurons .[6] C. Whats are the demerits of back-propagation algorithm. How it can be addressed suitably. [4]	