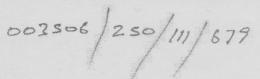


## December 2023

## B. Tech (CE/CE (Hindi Medium)) 5th Semester

	Machine Lear	rning (PEC-CS-D-501)	
Time: 3 H		Max. Ma	rks:75
Instructio	ns: 1. It is compulsory to answer all	the questions (1.5 marks each) of Part -A in short.	
	2. Answer any four questions fro		
	3. Different sub-parts of a questi	ion are to be attempted adjacent to each other.	
	4. Use of calculator is allowed		
	•	PART-A	- 10
	5), 85(6,4), 67(1,2), 65(6,9)		
Q1 (a)	What are different types of machi	ne learning paradigm?	(1.5)
	What is overfitting? How you can		(1.5)
(c)	What is training set and test set in	a machine learning model? How much data	(1.5)
2	will you allocate for training, valid		
(b)		spect to machine learning algorithm.	(1.5)
		machine learning, artificial intelligence and	(1.5)
(S1)(c)	deep leaning.	Surveilet site no estan rante estati	TO
(6)	What is 'naïve' in naïve bayes clas	reifiar?	(1.5)
(f)		y menent save ill	(1.5)
	(g) Define precision and recan.		
(h)	(h) What is pruning in decision trees? How it is done?		
(i) Why there is a need for regularizations?			(1.5)
(j)	Is a high variance in data good or	bad?	(1.5)
		PART-B	
Q2 (a)	Consider the following data (effects of hours of mixing on temperature of wood pulp), draw the best fit line using linear regression.		(7)
	HOURS OF MIXING (X)	TEMPERATURE OF WOOD PULP (Y)	
	2	21	
	4	27	
	6	29	
	8	64 86	
	10 12	92	
(b)	Consider the above data calculate	The state of the s	(4+4
(0)	I. the value of Y for X=16	the residuals in the predicted value of Y	
			-
Q3 (a)	How Random Forest works?		(5)
(b)	How prediction is done by time	series regression. What are the components	(10



		of time series? Can time series algorithms estimate the total sales in next 3 years of an insurance company?	
Q4		What is machine learning? Explain steps of implementation of machine learning algorithm? Why data preprocessing step is important? Discuss recent trends in various learning techniques of machine learning.	(15)
05	(a)	What is bias and variance in machine modeling?	(5)
		Explain distance measure formulas used in clustering. Using distance-based measure cluster the following eight points (with (x, y) representing locations) into three clusters:	(10)
		A1(2, 10), A2(2, 5), A3(8, 4), A4(5, 8), A5(7, 5), A6(6, 4), A7(1, 2), A8(4, 9)	
		Initial cluster centers are: A1(2, 10), A4(5, 8) and A7(1, 2).	(a) 10.
Q6	(a)	Differentiate between classification and clustering? How classification model is made to learn by machine learning algorithm? How do you design email	(10)
	(b)	spam filter? How is Amazon able to recommend other things to buy?	(5)
			o cam
Q7	7	Write short notes on the following:  I. Principal Component Analysis	(15)
		II. Artificial Neural Network  III. Bave's theorem	