Roll Number:_____	
Thapar Institute of Engineering and Technology, Patiala	
Computer Science and Engineering Department	
BE(3rd Year) September 30, 2022 *MST*	UML501: Machine Learning
Time: 2 Hours	Marks:25
Instructors: Dr. Singara Singh Kasana, Dr. Maninder Kaur, Dr. Jatin Bedi, Dr. Raman Goyal, Dr. Harpreet Singh, Ms. Swati	

Note: All questions are compulsory. All parts of a question must be answered in order.

Q1 Given the data in Table, reduce the dimensions from 2 to 1 using the Principal Component Analysis algorithm. [5]

X_1	4	8	13	7
X_2	11	4	5	14

Q2 Why regularization is needed in machine learning models? Derive the coefficients equation of ridge regularization for multiple linear regression using gradient descent optimization and discuss how the ridge regression can shrink the regression coefficients. [5]

Q3 The dataset of pass/fail in an exam for five students is given below: [5]

hours_study	Result (1= Pass, 0= Fail)
29	0
15	0
33	1
28	1
39	1

If we use logistic regression as a classifier and assume the model is given by following function for passing the exam.

$Y = (-64) + 2 \times \text{hours\_study}$

Assume that no pre-processing is required

(a) Calculate probability of pass for a student who studies 33 hours and compare the output with the actual Result.

(b) How many hours a student should study to ensure probability of passing is 95% or more.

P.T.O.

Q4 For the following confusion matrix of 4×4, [5]

		True (Actual)			
	..	*A*	*B*	*C*	*D*
	A	25	48	90	70
Predicted	*B*	12	14	16	30
	C	88	40	17	11
	D	33	24	13	18

Calculate the following

a) Precision with respect to each class (i.e. calculate separately for A, B, C and D respectively)
b) Sensitivity with respect to each class (i.e. calculate separately for A, B, C and D respectively)
c) Specificity for class A
d) False Positive rate for class A

Q5 Consider the following data give the height in inches (X) and the weight in kg (Y) of a random samples of 10 students from a large group of students of age 17 years: [5]

X	61	68	68	64	65	70	63	62	64	67
Y	112	123	130	115	110	125	100	113	116	125

(i) Estimate the regression equation using least square error method.
(ii) What will be the mostly likely estimate of the weight of the student when the height is 69 inches?