Ref. No.: Ex/PG/MEC/PE/T/123C/2025

M.E. MECHANICAL ENGINEERING
FIRST YEAR
SECOND SEMESTER EXAM 2025

SUBJECT: MACHINE LEARNING FOR MECHANICAL ENGINEERING
FULL MARKS 100

ANSWER ANY FIVE QUESTIONS. ALL QUESTIONS CARRY EQUAL MARKS
Assume appropriate value for any missing data

Question 1

a) What are regression and classification problems? Give examples.

b) What is supervised and unsupervised learning?

c) How do you find out the coefficients of a multivariate linear regression problem using least square method?

Question 2

a) What error term do you minimize in binary logistic regression?

b) Sketch a sigmoid function.

c) What are stochastic and batch gradient descent algorithms? Explain using flowcharts.

Question 3

a) Consider a linear support vector machine (SVM). Formulate the constrained optimization problem. How do you solve this?

b) What are the support vectors?

c) What is kernel trick for non-linear SVM?

[Turn over

Question 4

The following table (Table Q4) lists a dataset collected in an electronics shop showing details of customers and whether they responded to a special offer to buy a laptop

This dataset has been used to build a decision tree to predict which customers will respond to future special offers.

a) What is the Shannon's entropy for the dataset?

b) Among AGE, INCOME, STUDENT which should be root node?

c) What is Gini index?

ID	AGE	INCOME	STUDENT	CREDIT	BUYS
1	< 31	high	no	bad	no
2	< 31	high	no	good	no
3	$31 - 40$	high	no	bad	yes
4	> 40	med	no	bad	yes
5	> 40	low	yes	bad	yes
6	> 40	low	yes	good	no
7	$31 - 40$	low	yes	good	yes
8	< 31	med	no	bad	no
9	< 31	low	yes	good	yes
10	> 40	med	yes	bad	yes
11	< 31	med	yes	good	yes
12	$31 - 40$	med	no	good	yes
13	$31 - 40$	high	yes	bad	yes
14	> 40	med	no	good	no

Table Q4

Question 5

The following table (Table Q5) pertains to a waste water treatment plant. The explanation of the terms are as follows:-

SS-IN :- Measured value of solids coming into the plant
SED-IN:- Measured value of sediment coming into the plant
COND-IN:- Measured value of electrical conductivity at the inlet of the plant

SS-OUT, SED-OUT and COND-OUT are the corresponding quantities at the outlet

STATUS ok:- Everything is working correctly
STATUS settler:- There is a problem with the plant settler equipment
STATUS solids:- There is a problem with the amount of solids going through the plant

ID	SS -IN	SED -IN	COND -IN	SS -OUT	SED -OUT	COND -OUT	STATUS
1	168	3	1,814	15	0.001	1,879	ok
2	156	3	1,358	14	0.01	1,425	ok
3	176	3.5	2,200	16	0.005	2,140	ok
4	256	3	2,070	27	0.2	2,700	ok
5	230	5	1,410	131	3.5	1,575	settler
6	116	3	1,238	104	0.06	1,221	settler
7	242	7	1,315	104	0.01	1,434	settler
8	242	4.5	1,183	78	0.02	1,374	settler
9	174	2.5	1,110	73	1.5	1,256	settler
10	1,004	35	1,218	81	1,172	33.3	solids
11	1,228	46	1,889	82.4	1,932	43.1	solids
12	964	17	2,120	20	1,030	1,966	solids
13	2,008	32	1,257	13	1,038	1,289	solids

Table Q5

We want to formulate a naive Bayes model for the plant.

a) Compute P(SS-IN=250/STATUS=settler) assuming a normal probability distribution function.

b) Compute P(COND-IN=1500/STATUS=ok) assuming an exponential probability distribution function.

c) Describe the principle of k-nearest neighbor method using a simple example.

Question 6

a) What is underfitting and overfitting?

b) Illustrate standardiza on and range normalization of feature values.

c) Explain K-fold cross-validation