Nirma University
Institute of Technology
Sessional Examination, September 2025
M. Tech. in CSE/Data Science, Semester - I
6CS203CC22 – APPLIED MACHINE LEARNING

Roll / Exam No. | 25MCD005 | Supervisor's initial with date | |

Time: 2 Hours   Max. Marks: 50

Instructions:
1. Attempt all questions.
2. Show all calculations with a precision of four decimal places.
3. Assume necessary data wherever required and specify them.
4. Avoid attempting questions and their sub-questions in a random order.

Q.1
CO1,BL2

Why is feature scaling important when using Gradient Descent (GD) for regression, and what issues might arise if input features are not properly scaled? Provide an explanation in the context of model training and convergence. Also explain standard scaler method with a small example. **[5]**

Q.2
CO4,BL5

Consider following data of actual label and predicted label pairs: {(P, P), (P, P), (P, N), (N, P), (N, N), (P, P), (N, N), (N, P), (P, N), (N, N)}. Where each pair is actual label and predicted label for particular sample in the dataset of 10 samples. P stands for positive class and N stands for Negative class. Draw confusion matrix for the same. Also derive values for accuracy, precision, recall and F1 score. **[5]**

Q.3
CO2,BL3

Consider following dataset and identify appropriate Naive Bayes classification method for it: **[10]**

Win	Free	Hello	Class
1	1	0	Spam
1	0	0	Spam
0	1	0	Spam
0	0	1	Ham
0	0	1	Ham
1	0	1	Ham

Also calculate probability for the sample S with values (win = 1, free = 1, hello = 0) for being spam or ham. Use Laplace smoothing during calculation to avoid zero probabilities. At last mention the final class for the given sample S.

Q.4
CO3,BL4

Consider the following dataset with 6 samples, each having 2 **[15]** attributes (A1, A2) and a binary class label (0 or 1):

A1	A2	Label
2	4	0
4	2	1
4	4	1
6	2	0
6	4	1
5	3	0

A new sample Q = (5, 4) needs to be classified using the k-Nearest Neighbors (k=3) method. Use Cosine Similarity as a measure of similarity. Show all calculations along with final classification.

Q.5
CO3,BL4

To train a linear regression model using stochastic gradient descent **[15]** (SGD) with the loss function defined as ½(MSE). Given the following 5 training samples, and considering initial model parameter's values as $\theta = [\theta_0, \theta_1, \theta_2] = [0, 0, 0]$ with learning rate (η) as 0.1, calculate values of θ parameters after iteration 1, 2 and 3.

X_1	X_2	y
1	2	5
2	1	6
3	4	15
4	3	14
5	5	20

For the calculation, select samples in the same sequence as given in the table.