VII SEMESTER B.TECH. (CHEMICAL ENGINEERING)

END SEMESTER EXAMINATIONS, DEC- 2023

SUBJECT: Machine Learning in Chemical Engineering [CHE4073]

REVISED CREDIT SYSTEM

Date:07 /12/2023 Time duration: 180 minutes MAX. MARKS: 50

Instructions to Candidates:

- Answer **ALL** questions.
- Missing data may be suitably assumed.

1	Explain the following terms in reference to Machine learning: a. Labels b. Training set c. Testing set d. Loss function	2
2	Discuss the advantages of Principal Component analysis (PCA) along with the assumptions for PCA and its limitations	4
3	Distinguish between supervised learning and unsupervised learning algorithms of machine learning by giving an example.	4
4	Differentiate between the Regression and Classification algorithms with a minimum of 3 points	3
5	Summarize the importance of Regularization in the Machne learning algorithm and When the regularizations are used.	3
6	Suppose you run gradient descent to fit a logistic regression model for the data set below with parameter $\theta \in \Re^{n+1}$. The cost function is evaluated using the following equation: $J(\theta) = -\frac{1}{m}\sum_{i=1}^{m}\left[y^{(i)} \log h_\theta(x^{(i)}) + (1-y^{(i)}) \log(1-h_\theta(x^{(i)})) \right]$. Compute the cost of ***J(θ)*** with the parameter vector as [-10 0.1 0.1].	4
7	Justify the use of regularization in the Machine Learning algorithm. Explain the various methods used for regularization.	3
8	Justify the application of Machine learning in process systems engineering with *three* examples.	2

x_1	14	10	35	40	50
x_2	65	40	60	80	70
y	1	0	0	1	0

9	Given the data set below consider the linear regression model as $h_\theta(x) = \theta_0 + \theta_1 x$. Estimate the parameters i.e., θ_0 and θ_1 that you would obtain upon running gradient descent with a learning parameter equal to 0.1 after the ***second iteration*** on this model?. The initial guess of the parameter set i.e., $\theta 0$ and $\theta 1$ taken as $[1\ 1]^T$. X: 60, 50, 80, 70, 50, 80 y: 170, 90, 130, 110, 60, 110	5
10	Illustrate the Confusion matrix and differentiate between Precision and Recall	4
11	Examine the classification process of a Support Vector Machine with the help of two-dimensional space, decision boundaries and support vectors.	3
12	Assess the role of 'C' in the Support Vector Machine and the intuition of a large-margin classifier.	3
13	Differentiate between K-means and K-NN algorithms and develop a K-means algorithm	4
14	Formulate the structure of Artificial neural network configurations.	2
15	Formulate the principal component analysis to evaluate the first principal component (z_1), which captures the maximum variation in the data for the given data set below, and determine the percentage variance captured by each principal component. Evaluate the approximate value of original data using only the first principal component (i.e., z_1) and report the actual and approximate data error. X1: 2.5, 2.3, 2.2, 2, 1, 1.5, 1.1 X2: 2.4, 2.7, 2.9, 1.6, 1.1, 1.6, 0.9	4