Type: DES

Q1. Differentiate between Machine Learning and Deep Learning? (4)

Q2. Distinguish between supervised learning and unsupervised learning algorithms of machine learning by giving an example. (4)

Q3. Explain the following terms in reference to Machine Learning:
(i) Labels
(ii) Training set
(iii) Testing set
(iv) Loss function (2)

Q4. Summarize the importance of Regularization in the Machne learning algorithm and When the regularizations are used. (4)

Q5. Differentiate between regression and classification problems by giving two examples of the chemical engineering domain. When should you use classification over regression? (4)

Q6. Assess the trade-off between bias and variance in the machine learning algorithm (2)

Q7. It is decided to design a soft-sensor for a multicomponent distillation column to estimate the product composition and synthesize feedback control design. Develop a detailed step-by-step procedure to design the soft sensor using a suitable machine learning algorithm. (4)

Q8.

The data table below and a cost function for the linear regression of the form with a single variable. i.e.: $J(\theta_0, \theta_1) = \frac{1}{2m}\sum_{i=1}^{m}\left(h_\theta(x)^{(i)} - y^{(i)}\right)^2$. The gradient descent algorithm is used to estimate the parameter of the model. i.e. $h_\theta(x) = \theta_0 + \theta_1 x$ with learning rate (α) = 0.05. Obtain and report the parameter vector in every iteration till two iterations with initial parameter vector as [ 1 1]'.

| x | 7 | 6 | 7 | 5 | 6 | 10 |
|---|---|---|---|---|---|---|
| y | 4 | 1 | 3 | -2 | 1 | 4 |

(4)

Q9. Explain the best way to monitor whether gradient descent is working for regularized linear regression correctly? Justify your answer (2)

Q10. Discuss the advantages of Principal Component analysis (PCA) along with the assumptions for PCA and its limitations (4)

Q11. Distinguish between logistic regression and SVM (4)

Q12. Justify the application of Machine learning in process systems engineering with *two* examples. (2)

Q13. Differentiate between K-means and K-NN algorithms and develop a K-means algorithm. (4)

Q14.

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)

Q15. Explain the Confusion Matrix and its importance in assessing the machine learning models. (2)