AUTUMN MID SEMESTER EXAMINATION-2022

School of Computer Engineering
Kalinga Institute of Industrial Technology, Deemed to be University
Machine Learning
[CS 3035]

Time: 1 1/2 Hours **Full Mark: 20**

Answer any four Questions including Q.No.1 which is Compulsory.
The figures in the margin indicate full marks. Candidates are required to give theiranswers in their own words as far as practicable and all parts of a question should beanswered at one place only.

1. Answer all the questions. [1 x 5]

a) Out of the Following Examples, Which would you address using an supervised learning Algorithm?

(i) Given email labeled as spam or not spam, learn a spam filter
(ii) Given a set of news articles found on the web, group them into set of articles about the same story.
(iii) Given a database of customer data, automatically discover market segments and group customers into different market segments.
(iv) Find the patterns in market basket analysis

b) Which machine learning models are trained to make a series of decisions based on the rewards and feedback they receive for their actions?

(i)SupervisedLearning
(ii)UnsupervisedLearning
(iii)ReinforcementLearning
(iv)All of the above

c) Overfitting is attributed by high bias low variance.
(i) True
(ii)False

d) Why KNN is called as lazy learner algorithm?

e) State two advantages of Gradient Descent method over Least square method?

2. (a) Explain at least 5 different metrics with appropriate mathematical expressions for assessing regression performance of machine learning models. [3 Marks]

(b) Draw a comparison between Linear regression and Logistic Regression.[2 Marks]

3. (a) Give one real life application each of supervised and unsupervised algorithm. Differentiate between the two leanings. [3 Marks]

(b) We require low bias and low variance ate the same time for achieving good generalized performance.Explain the bias-variance trade-off using visual representations. [2 Marks]

4. (a) Using K-NN algorithm and the given data set, predict the class label of the test data point (8,5), where K=3 and Euclidean distance. [3 Marks]

X	Y	Label
4.2	3.8	0
6.5	7.7	1
7.3	8.6	1
5.7	5.9	0
8.0	8.1	1
10.0	6.5	1

(b) Discuss the important features of data partition in cluster analysis and discuss about K-means clustering. [2 Marks]

5) (a) Fit a straight line Y=a+bX to the data by the method of least squares. [3 Marks]

X	5	10	15	20	25
Y	16	19	23	20	30

(b) Explain the function of tuning parameter λ in Regularized Linear Regression? [2 Marks]

*** Best of Luck ***