PATTERN RECOGNITION
(CSEN 4233)

Time Allotted : 3 hrs   **Full Marks : 70**

Figures out of the right margin indicate full marks.

Candidates are required to answer Group A and
<u>any 5 (five)</u> from Group B to E, taking <u>at least one</u> from each group.

Candidates are required to give answer in their own words as far as practicable.

Group – A
(Multiple Choice Type Questions)

1. Choose the correct alternative for the following: **10 × 1 = 10**

  (i) Which of the following loss function should one minimize to fit the training data in logistic regression?
(a) Least Square Error   (b) Maximum Likelihood
(c) Negative log-likelihood   (d) None of the above.

  (ii) If the variance covariance matrix is diagonal matrix then the features are_____
(a) independent   (b) dependent
(c) zero mean value   (d) uniformly distributed.

  (iii) When two classes can be separated by a straight line, they are known as -
(a) Linearly separable classes   (b) Linearly inseparable classes
(c) May depend on system, can be separable/inseparable   (d) All of the above

  (iv) An artificial neuron receives n inputs x1, x2, ..., xn with weights $w_1, w_2, ..., w_n$ attached to the input links. The weighted sum ___ is computed to be passed on to a non-linear filter Φ called activation function to release the output.
(a) Σw_i (b) Σx_i
(c) $\Sigma w_i + \Sigma x_i$ (d) $\Sigma w_i * x_i$

  (v) What is the objective of perceptron learning?
(a) Class identification   (b) Weight adjustment
(c) Weight adjustment along with classification   (d) None of the above.

  (vi) A patient has been tested positive for a disease. Consider the following information and find the probability that the patient actually has the disease?
P(Test = +ve | Disease = True) = 0.95, P(Test = +ve | Disease = False) = 0.05
and P(Disease = True) = 0.01
(a) 0.116   (b) 0.516
(c) 0.95   (d) 0.161.

  (vii) Fuzzy c-means algorithm can be categorised as _____ technique.
(a) non-hierarchical clustering   (b) linear regression
(c) agglomerative clustering   (d) logistic regression

  (viii) Clustering algorithm usually employs
(a) supervised learning   (b) unsupervised learning
(c) reinforcement learning   (d) competitive learning.

  (ix) Consider the following two statements:
Statement 1: Principal Component Analysis is an unsupervised method
Statement 2: It searches for the directions that data have the largest variance.
(a) Only Statement 1 is correct   (b) Only Statement 2 is correct
(c) Both statements are incorrect   (d) Both statements are correct

  (x) Fisher's linear discriminant analysis is used in _____
(a) unsupervised classification   (b) supervised classification
(c) reinforcement classification   (d) competitive classification.

Group - B

2. (a) You are given the graph in Fig.1, which shows two types of data sets, one marked as circles, other as squares. Add at least one more data point for each data type (circle and square). Then add a point on the graph and mark it as X. (Read the next question before placing X on the graph.)

Categorize the point X using k-Nearest Neighbour (kNN) Classification Technique. Show two different values of k such that you get two different classifications for X each time. Show your working / justification clearly.
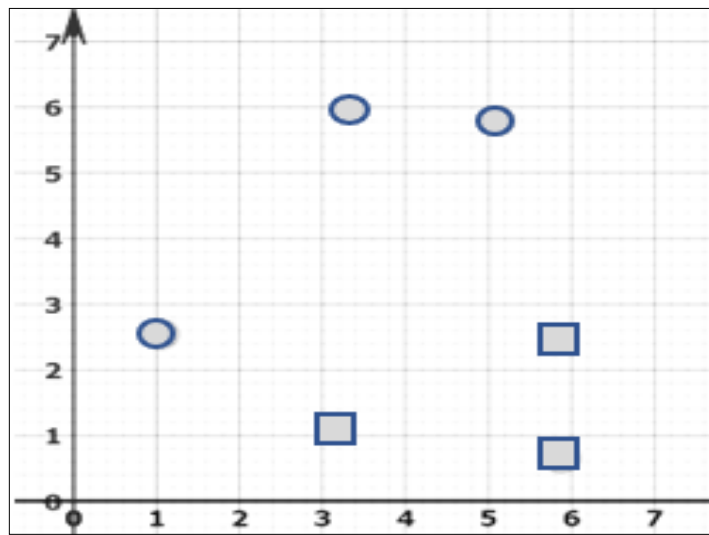
[(CO3)(Categorize/IOCQ)]

Fig.1

(b) What are the differences between Metric and Non-Metric Measures? Give one example of a measure which qualifies as a metric.
[(CO1)(Recall/LOCQ)]

6 + 6 = 12

3. (a) Explain the concept of Feature Selection using an example.
[(CO1)(Explain/LOCQ)]

(b) Compare Parametric and Non-Parametric learning approach. Use examples to indicate where they can be used.
[(CO1)(Compare/LOCQ)]

(c) A random sample of 395 people were surveyed and each person was asked to report the highest education level they obtained. The data that resulted from the survey is summarized in the following Table 1:

Gender	High_School	Bachelors	Masters	Ph.d.	Total
Female	60	54	46	41	**201**
Male	40	44	53	57	**194**
Total	**100**	**98**	**99**	**98**	**395**

Table 1

Are gender and education level dependent at 5% level of significance? In other words, given the data collected above, is there a relationship between the gender of an individual and the level of education that they have obtained?
If you require, a snapshot of the Chi-Square Probabilities is given below in Table 2:

df	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
1	---	---	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188

Table 2

[(CO5)(Judge/HOCQ)]

3 + 3 + 6 = 12

Group – C

4. (a) Derive the weight update equation of a feed forward Multi Layered Perceptron network using back propagation learning.
[Analyze/IOCQ/CO1, CO3, CO5, CO6]

(b) Write difference and similarity among batch (i.e., vanilla) gradient descent, stochastic gradient descent and mini-batch gradient decent.
[Understand/IOCQ/ CO1, CO3, CO5, CO6]

(c) Justify the statement with proper reasoning- "Stochastic gradient descent algorithm is superior than batch (i.e., vanilla) gradient descent"
[Evaluate/HOCQ/ CO1, CO3, CO5, CO6]

6 + 3 + 3 = 12

5. (a) Briefly explain Entropy and Gini Impurity in the context of decision tree learning.
[Understand/IOCQ/CO1, CO3, CO5]

(b) Consider the following data set for a binary class problem. Calculate the information gain when splitting on three attributes 'Over 170CM', 'Eye' and 'Hair length'. Which attribute would the decision tree induction algorithm choose?
[APPLY/IOCQ/CO1, CO2, CO3, CO6]

Sl No	Over 170CM	Eye	Hair length	Gender
1	No	Blue	Short	Male
2	Yes	Brown	Long	Female
3	No	Blue	Long	Female
4	No	Blue	Long	Female
5	Yes	Brown	Short	Male
6	No	Blue	Long	Female
7	Yes	Brown	Short	Female
8	Yes	Blue	Long	Male

[(CO1, CO4, CO5)[Remember/LOCQ][Understand/IOCQ] and [Apply/HOCQ]]
5 + 7 = 12

Group – D

6. (a) What is difference between soft and hard clustering techniques.   *[Understand/LOCQ/CO1,CO4,CO5]*
   (b) Explain the k means algorithm.   *[Understand/LOCQ/CO1,CO4,CO6]*
   (c) Using Fuzzy C means method cluster the group of 2D data objects, having level of fuzziness: 1.28, Number of clusters C = 2. Data Objects are ([0.1,0.3], [0.3,0.3], [0.7,0.2], [0.9,0.4], [0.5,0.5], [0.2,0.3], [0.8,0.6], [0.3,0.6])
   The initial membership value for the data points belonging to cluster 1 are (0.67, 0.51, 0.88, 0.30, 0.33, 0.44, 0.50, 0.18).
   [N.B. Show the calculations of first two epochs only.]   *[Apply/IOCQ/ CO1, CO4, CO6]*
   3 + 4 + 5 = 12

7. (a) Briefly explain the agglomerative clustering technique with suitable example.   *[Understand/LOCQ/CO1,CO4]*
   (b) Apply DBSCAN on the following set of data points, to determine the clusters. Show all steps in detail for first 2 epochs, assuming eps = 1 and minPts = 2. Data Points: (83,80), (93,91), (73,72), (4,1), (4,2), (210,211), (210,212), (220,221), (202,222), (121,132), (123,124).   *[Apply/IOCQ]/CO1,CO4,CO6]*
   5 + 7 = 12

Group – E

8. (a) What do you mean by Dimensionality Reduction? Mention the different approaches you can take to achieve Dimensionality Reduction.   *[(CO6)(Discuss/LOCQ)]*
   (b) Let the patterns (1, 2), (2, 3), (3, 3), (4, 5), (5, 5) belong to *C1* and (1, 0), (2, 1),(3, 1), (3, 2), (5, 3), (6, 5) belong to *C2*.
   Now reduce the dimension using Fisher's Linear Discriminant Analysis.   *[(CO6)(Design/HOCQ)]*
   4 + 8 = 12

9. (a) Explain the working principal of Principal Component Analysis (PCA) using an example.   *[(CO6)(Solve/IOCQ)]*
   (b) Write short notes on any two:
   (i) Linear Regression
   (ii) Logistic Regression
   (iii) Convolution Neural Network.   *[(CO5)(Recall/LOCQ)]*
   6 + 6 = 12

Cognition Level	LOCQ	IOCQ	HOCQ
Percentage distribution	31.25	39.58	29.17

Course Outcome (CO):

After the completion of the course students will be able to
CSEN4233.1. Learn and understand feature, pattern and the problem of pattern recognition.
CSEN4233.2. Understand and describe the difference between supervised and unsupervised learning.
CSEN4233.3. Understand and apply pattern recognition algorithm that utilizes supervised learning.
CSEN4233.4. Understand and apply pattern recognition algorithm that utilizes unsupervised learning.
CSEN4233.5. Analyze pattern recognition algorithms and techniques.
CSEN4233.6. Design simple pattern recognition systems.

LOCQ: Lower Order Cognitive Question; IOCQ: Intermediate Order Cognitive Question; HOCQ: Higher Order Cognitive Question.