Seat No:

# MIT ACADEMY OF ENGINEERING

**COURSE CODE: CS332T** **21 MAY 2019**

**TY BTECH SEMESTER - VI 2018 - 2019 EXAMINATION**

**DEPARTMENT OF COMPUTER ENGINEERING**

**END SEMESTER EXAMINATION**

**MACHINE LEARNING (OPEN ELECTIVE - MACHINE INTELLIGENCE)**

**TIME : 3 HOURS** **X MARKS : 100 MARKS**

**TOTAL NO OF QUESTIONS: 05** **TOTAL NO OF PRINTED PAGES:3**

**INSTRUCTIONS TO CANDIDATES:**

1. Assume suitable data wherever necessary
2. Non programmable scientific calculators are allowed
3. Black figures to the right indicate full marks

| | | | | | |
|---|---|---|---|---|---|
| **1** | **a)** | What is Bayes Theorem? How is it useful in a machine learning context?<br>(state & explain 2M, 2 application 4M) | **[06]** | **CO1** | **L1** |
| | **b)** | How we can reduce the data complexity in machine learning ?<br>Explain your answer with example/ real time application.<br>(Correct Ans. 2M, explain with application 2M) | **[04]** | **CO1** | **L1, L2** |
| | **c)** | Analyze your decision base on Decision tree by using entropy<br>and gain formula for any data set/ example.Find out the best<br>attribute in given dataset of Weather forcasting by calculationg an<br>entropy and gain value.<br>(Analyze derivation 3M, tree 2M,calculation with correct ans 5M | **[10]** | **CO4** | **L4** |

Table 1: Data set of weather forcasting

| Day | Outlook | Temp. | Humidity | Wind | Play Tennis |
|---|---|---|---|---|---|
| D1 | Sunny | Hot | High | Weak | No |
| D2 | Sunny | Hot | High | Strong | No |
| D3 | Overcast | Hot | High | Weak | Yes |
| D4 | Rain | Mild | High | Weak | Yes |
| D5 | Rain | Cool | Normal | Weak | Yes |
| D6 | Rain | Cool | Normal | Strong | No |
| D7 | Overcast | Cool | Normal | Weak | Yes |
| D8 | Sunny | Mild | High | Weak | No |
| D9 | Sunny | Cold | Normal | Weak | Yes |
| D10 | Rain | Mild | Normal | Strong | Yes |
| D11 | Sunny | Mild | Normal | Strong | Yes |
| D12 | Overcast | Mild | High | Strong | Yes |
| D13 | Overcast | Hot | Normal | Weak | Yes |
| D14 | Rain | Mild | High | Strong | No |

**Q.2 a)** Construct univariate and multivariate decision tree for approval of credit/loan application in bank. **[6] CO3 L3**

(Derivation 3M,explaination and tree 2 M,Graph 1M)

**b)** Assume x1: age,x2:Years in job,x3:Gender,and x4 :Job Type. Is it possible to use the decision tree for feature extraction to this assumption? If yes ,Simplify your answer. **[6] CO4 L4**

(Correct Ans 2M,Explanation 3M,Any diagram 1M)

**c)** Is Clustering an unsupervise learning method ? Justify user answer with application busness and retail and also solve problem in business with any one methode of clustering(Take your won Data set) . **[8] CO4 L5**

(Correct Ans 1M,Explanation 2M,Problem Solution: 5M)

**Q.3 a)** Solve example for Apriori algorithm to explain the concept of association rule in machine learning.Assume Support Count=2 **[10] CO4 L3**

| TID | List of Items |
|---|---|
| 1 | A,B,E |
| 2 | B,D |
| 3 | BC |
| 4 | A,B,D |
| 5 | A,C |
| 6 | B,C |
| 7 | A,C |
| 8 | A,B,C,E |
| 9 | A,B,C |

(Concept: 2M, Problem Solution: 8M)

**b)** Compare the Frequest Item Set and closed Item set with application of market basket analysis for Soft drinks and Diapers marketing **[6] CO4 L4**

(Comparison: 2M, Application: 4M)

**c)** Solve the utility theory example for gaming analytics. **[4] CO4 L3**

(Concept: 2M, Example with Solution: 2M)

**Q.4 a)** Is Hierarchical Clustering is Non-Linear method ? Justify answer involving raltime application.(1M for correct identification,3 marks for justification) **[12] CO5 L5**

Solve following example using Single link method of Hierarchical agglomerative clustering and also Draw correct Dendrogram.(Step by Step Soving-6M,Dendrogram 2M)

| | X1 | X2 |
|---|---|---|
| P1 | 0.4 | 0.53 |
| P2 | 0.22 | 0.38 |
| P3 | 0.35 | 0.32 |
| P4 | 0.26 | 0.19 |
| P5 | 0.08 | 0.41 |
| P6 | 0.45 | 0.3 |

**b)** Evaluate Confidence for SVM classification to get hyperplan.Also formulate Maximizing and Minimizing for find line. **[4] CO5 L5**

(Evaluation: 2M, Formulation: 2M)

**c)** Classify Binary and multiclass SVM classification with suitable example/application. **[6] CO5 L4**

(Classification: 4M, Example: 2 M)

**Q.5 a)** Compare FP-Growth algorithm with Apriori algorithm and which one is better algorithm in solving real time application to get more profit in buisness.(2M Comparision,1M application,Correct identification-1M) **[10] CO5 L4, L6**

Solve following with Frequest Pattern Growth algorithm using association rule mining(Solving with step by step 6M)

| TID | Item Set |
|---|---|
| 100 | {M,O,N,K,E,Y} |
| 200 | {D,O,N,K,E,Y} |
| 300 | {M,A,K,E} |
| 400 | {M,U,C,K,Y} |
| 500 | {C,O,O,K,I,E} |

**b)** Is Facebook Recommendation System uses concept of SVM.Then why it uses SVM ,what is need ,whether it is linear/Non linear classifier.Justify your answer. **[8] CO5 L5**

(Correct Ans:2M, , Need: 2M, justification with diagram 4M)