Jahangirnagar University
Department of Computer Science and Engineering
4th Year 2nd Semester B.Sc. (Hons.) Final Examination -2019

Assignment for Final Examination

Course Title: **Data Mining** Course No: **CSE-459**
Submission Dead Line: **29/12/2020, 11:59 PM** Full Marks: **10**

Problem Description

Data mining is a technique for the extraction of interesting and potentially useful information or patterns from data in large databases. Various organizations now-a-days apply data mining technique to mine their operational databases and data warehouses to extract patterns and knowledge about their customers and participants to predict their behaviour in choosing a product, taking an action, selecting a course etc.

Following is a description of the behavior of some customers of ABC electronics with a sample database. In this example, the following Table 1 presents a training data set of class-labelled tuples of ABC electronics of their customers.

Table 1: Sample Database.

sl. no.	age	income_level	employed	credit_rating	class: buys_computer
1	senior	low	no	fair	yes
2	youth	high	yes	fair	no
3	youth	high	yes	excellent	no
4	senior	medium	yes	fair	yes
5	senior	low	no	excellent	no
6	youth	medium	no	excellent	yes
7	youth	medium	yes	fair	no
8	middle_aged	high	yes	fair	yes
9	middle_aged	low	no	excellent	yes
10	middle_aged	medium	yes	excellent	yes

Write down the answers of the following questions by analyzing the database given in Table 1 considering it a data mining problem to help decision making by the Marketing manager in predicting customers behavior during a sale at ABC electronics.

1. Explain how does classification work?
2. How does classification data mining problem can be solved using decision tree induction?
3. Give the pseudocode for a decision tree induction algorithm.
4. Explain with example using Table 1:
   i) How should the training records be split?
   ii) How should the splitting procedure stop?
5. In Table 1, which attribute should be selected as the best attribute as the test condition for splitting the training records? Explain in brief.
6. Draw the decision tree that can be obtained from Table 1 by applying a classification data mining algorithm on the Table 1 data set.
7. Is it required to prune the decision tree obtained initially? Draw the pruned decision tree.
8. Discuss how classification data mining technique can help decision making authorities by mining knowledge from their operational database using Table 1, e.g., to help decision making by the Marketing manager in predicting customers behavior during a sale at ABC electronics.
