



FEDERAL UNIVERSITY OF TECHNOLOGY, MINNA  
SCHOOL OF INFORMATION AND COMMUNICATION TECHNOLOGY  
DEPARTMENT OF INFORMATION AND MEDIA TECHNOLOGY

FIRST SEMESTER 2019/2020 EXAMINATION

**COURSE CODE:** IMT 412  
**COURSE TITLE:** DATA MINING II  
**CREDIT UNITS:** 2  
**TIME ALLOWED:** 2 hours  
**COURSE LECTURER(S):** Dr. I. O. ALABI  
**NUMBER OF QUESTIONS:** 5  
**NUMBER OF PAGES:** 3 (INCLUDING THIS PAGE)

**INSTRUCTIONS**

- Answer ANY FOUR questions
- and Do NOT use red pen or other faint ink pens
- Please write **LEGIBLY** in a clear handwriting
- This exam is closed book, closed notes, closed laptop and closed cell phone
- Please use non-programmable calculators only



1. a. Define Web data mining. Then explain the major tasks in Web data mining? (5 marks)  
b. Briefly discuss whether or not each of the following activities is a data mining task: (2 marks each)
  - \* (i) Dividing the customers of a company according to their gender.
  - \* (ii) Dividing the customers of a company according to their profitability.
  - \* (iii) Computing the total sales of a company.
  - \* (iv) Sorting a student database based on student identification numbers.
  - ✓ (v) Predicting the outcomes of tossing a (fair) pair of dice.
2. a. Explain the following concepts: (10 marks)
  - i. Machine learning    ii. Supervised learning    iii. Unsupervised learning
  - iv. *Apriori* principle    v. Data exploration  
b. List any FIVE data mining classifiers you know. (5 marks)
3. a. Explain the concept of Clustering. (5 marks)  
b. Give any TWO examples of application of cluster analysis. (6 marks)  
c. A tourist is considering two locations, P and Q with regard to their similarities and dissimilarities. He itemized ten attributes for the two locations as shown in the table below:

Attribute	Location P	Location Q
1	1	1
2	0	0
3	1	1
4	0	0
5	0	1
6	1	1
7	1	1
8	1	1
9	1	0
10	1	1

Determine whether or not the two locations are similar. (4 marks)

4. 2. a. Explain the following, use sketches as necessary: (10 marks)
  - i) Binary attributes    ii) Nominal attributes    iii) Ordinal attributes
  - iv) Continuous attributes    v) Decision tree  
b) Sketch a lattice for all possible item-sets  $I = \{a, b, c, d, e\}$ . (5 marks).



5. a) What is a Confusion matrix? Illustrate with a simple 2-class classifier and use it to explain the following terms: i) True-Positive ii) False Positive iii) False Negative, and iv) True Negative. (10 marks).

b) State the formulae of coincidence matrix, then use it to calculate the Error rate and Accuracy rate of the 2-class classifier whose results are shown in the table 1. (5 marks)

Table 1

	<i>Class 1</i>	<i>Class 2</i>
<i>Class 1</i>	674	91
<i>Class 2</i>	89	146

March, 2021