



UNIVERSITY EXAMINATIONS

SECOND SEMESTER 2023/2024 ACADEMIC YEAR

**FOURTH YEAR EXAMINATION FOR THE DEGREE OF
BACHELOR OF COMPUTER SCIENCE AND
BACHELOR OF SCIENCE IN (ICT)**

**BICT 425: DATA ANALYTICS AND BUSINESS
INTELLIGENCE**

STREAM: R

TIME: 2 HRS

DAY: MONDAY [8.30 – 10.30 A.M]

DATE: 15/04/2024

THIS QUESTION PAPER CONSISTS OF THREE (3) PAGES

PLEASE DO NOT OPEN UNTIL THE INVIGILATOR SAYS SO.

INSTRUCTIONS: ANSWER QUESTION ONE AND ANY OTHER TWO QUESTIONS

SECTION A: (COMPULSORY) TOTAL MARKS FOR THIS SECTION IS 30 MARKS

QUESTION ONE (30 MARKS)

- a) Explain **SIX** reasons why business intelligence is important **(6 Marks)**
- b) Data Mining is the process of investigating hidden patterns of information to various perspectives for categorization into useful data, discuss **THREE** advantages and **THREE** disadvantages of business intelligence **(6 Marks)**
- c) Discuss any **FIVE** challenges of data mining **(10 Marks)**
- d) Using examples explain the following data mining techniques
- i. Classification **(2 Marks)**
 - ii. Clustering **(2 Marks)**
 - iii. Regression **(2 Marks)**
 - iv. Association rules **(2 Marks)**

SECTION B. TOTAL MARKS FOR THIS SECTION IS 40 MARKS

ANSWER ANY TWO QUESTIONS FROM THIS SECTION. EACH QUESTION IN THIS SECTION CARRIES 20 MARKS.

QUESTION TWO (20 MARKS)

- a) As a highly application-driven domain, data mining has incorporated many techniques from other domains such as statistics , machine learning , information retrieval and data visualization , Discuss how these domains have contributed to business intelligence. **(8 Marks)**
- b) Using a diagram describe Power BI architecture **(6 Marks)**
- c) Outlier detection (also known as *anomaly detection*) is the process of finding data objects with behaviors that are very different from expectation. In general, outliers can be classified into three categories,
- i. Global Outliers **(2 Marks)**
 - ii. Contextual Outliers **(2 Marks)**
 - iii. Collective Outliers **(2 Marks)**

QUESTION THREE (20 MARKS)

- a) Data Exploration is about describing the data by means of statistical and visualization techniques. Describe **FIVE** types of attributes. **(10 Marks)**
- b) After Data Exploration technique we use the data to process for improving the quality of patterns to mined, List **FIVE** factors of quality **(5 Marks)**
- c) Describe the following terms
- i. Confusion matrix **(2 Marks)**
 - ii. Clustering **(2 Marks)**
 - iii. Frequent pattern mining **(1 Mark)**

QUESTION FOUR (20 MARKS)

- a) List and explain the major steps involved in data Preprocessing **(10 Marks)**
- b) Design the Decision Tree Induction Algorithm **(4 Marks)**
- c) There are **three** popular attribute selection measures
- i. Gini index **(2 Marks)**
 - ii. Information gain **(2 Marks)**
 - iii. Gain ratio **(2 Marks)**

QUESTION FIVE (20 MARKS)

- a) Cross-industry Standard Process of Data Mining and business intelligence (CRISP-DM) comprises of **SIX** phases designed as a cyclical method, Discuss these steps **(12 Marks)**
- b) Compare and contrast machine learning and data mining **(4 Marks)**
- c) Data classification is a two-step process, explain them
- i. Learning step **(2 Marks)**
 - ii. Classification step **(2 Marks)**