



Kasarani Campus
Off Thika Road
Tel. 2042692 / 3
P. O. Box 49274, 00100
NAIROBI
Westlands Campus
Pamstech House
Woodvale Grove
Tel. 4442212
Fax: 4444175

KIRIRI WOMEN'S UNIVERSITY OF SCIENCE AND TECHNOLOGY
UNIVERSITY EXAMINATION, 2023/2024 ACADEMIC YEAR
FOURTH YEAR, FIRST SEMESTER EXAMINATION
FOR THE BACHELOR OF BUSINESS INFORMATION TECHNOLOGY
KBI 2404 – DATA WAREHOUSES & KNOWLEDGE MINING TECHNOLOGIES

Date: 13TH December 2023
Time: 2:30PM – 4:30PM

INSTRUCTIONS TO CANDIDATES

ANSWER QUESTION ONE (COMPULSORY) AND ANY OTHER TWO QUESTIONS

QUESTION ONE (30 MARKS)

Read the following excerpt carefully and answer the questions that follow

A DBMS built for Online Transaction Processing (OLTP) is generally regarded as unsuitable for data warehousing because each system is designed with a different set of requirements in mind. An organization will normally have a number of different OLTP systems for business processes such as inventory control, customer invoicing, and point of sale. These systems generate operational data that is detailed, current and subject to change. The OLTP systems are optimized for a high number of transactions that are predictable, repetitive and update intensive. The OLTP data is organized according to the requirements of the transactions associated with the business applications and supports the day to day decisions of concurrent operational users.

OLTP systems are not built to quickly answer ad hoc queries. They also tend not to store historical data, which is necessary to analyze trends. Basically, OLTP offers large amounts of raw data, which is not easily analyzed. The data warehouse allows more complex queries to be answered besides just simple aggregations such as, 'what is the average selling price for properties in the major towns in Kenya?' The types of queries that a data warehouse is expected to answer range from the relatively simple to the highly complex and are dependent on the type of end-user access tools used

- a) Discuss the following Data Mining Techniques
 - i) Regression
 - ii) Neural networks
 - iii) Genetic algorithm (6 Marks)
- b) Discuss Data mining principles used by practitioners to extract meaningful insights from vast data sets (6 Marks)
- c) Describe Common data mining analyses and their business applications (6 Marks)
- d) state and discuss types of OLAP as stated in data warehouse and data mining (6 Marks)
- e) Discuss the concept of Bayesian classifiers as a statistical classifier (6 Marks)

QUESTION TWO (20 MARKS)

- a) In detail discuss the main applications of association rule mining (4 Marks)
- b) State characteristics of data mining to support the management's decision-making process(8 Marks)
- c) Preprocessing of data is mainly to check the data quality. Discuss factors that are used to check quality of data. (8 Marks)

QUESTION THREE (20 MARKS)

- a) The business managers have realized the need to create a data warehouse as repository of historical data together with a current database for the organization
- i) Identify the application areas of data warehouse repository (8 Marks)
 - ii) Generate the benefits of using data warehouse to a business (8 Marks)
 - iii) Mention the techniques applied during data mining process (4 Marks)

QUESTION FOUR (20 MARKS)

- a) The implementation of a database may not be adequate to meet the demand of information needs of an organization, hence further processing of the data in the database is necessary
- i) Discuss two approaches to improving the quality of hierarchical clustering (6 Marks)
 - ii) State and explain typical Requirements of Clustering In Data Mining. (8 Marks)
 - iii) Discuss types of outlier detection used in data mining and data warehouse. (6 Marks)

QUESTION FIVE (20 MARKS)

The business managers have realized the need to create a data warehouse as repository of historical data together with a current database for the organization

- i) List the phases involved in the data warehouse delivery process. (8 Marks)
- ii) Generate the benefits of using a data warehouse to a business (8 Marks)
- iii) With appropriate examples state and explain the components of apriori algorithm (4 Marks)