

**Kiriri Women's University of Science and Technology**

**Data Warehousing and Knowledge Mining**

**2018/ 2019 Jan – April Exam**

**Section A: Compulsory**

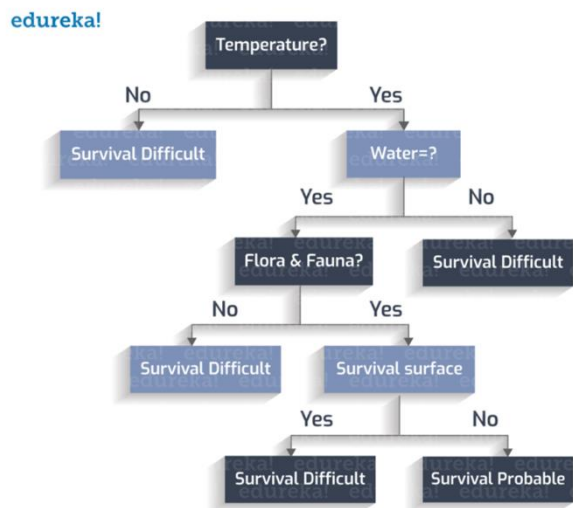
**Question One ( 30 marks )**

- A. Outline three reasons why an organization should engage in data warehousing and Knowledge mining. **( 6 marks )**
- B. Define the following: **( 5 marks )**
- i. Data Mart
  - ii. Knowledge discovery
  - iii. Information
  - iv. Data characterization
- C. Differentiate : **( 6 marks )**
- i. Descriptive functions and Predictive functions
  - ii. Entropy and information gain
  - iii. rule antecedent and rule consequent
- D. State any two classification methods that can be used in data mining. **( 2 marks )**
- E. Outline how data mining can be applied in the following industries: **( 4 marks )**
- i. Banking
  - ii. Retail
- F. In data mining Clustering is the process of making a group of abstract objects into classes of similar objects. State two requirements of clustering functions. **( 2 marks )**
- G. Explain two challenges faced in web mining. **( 4 marks )**
- H. What does the term “Vision-based page segmentation (VIPS) “refer to in web mining? **( 2 marks )**

**Section B: Optional (Attempt any two Questions)**

**Question Two (20 marks)**

- A. Define the following as used in decision trees: **( 3 marks )**
  - a. Root nodes
  - b. Branches
  - c. Leaf nodes
- B. State two advantages and two disadvantages of using decision trees in knowledge mining. **( 4 marks )**
- C. Derive four classification rules from the decision tree below: **( 4 marks )**



- D. Outline three different kinds of information that are collected for data mining. **( 6 marks )**
- E. State the six steps in the data mining life cycle. **( 3 marks )**

**Question Three (20 marks)**

- A. Explain three features of a data warehouse. **( 6 marks )**
- B. Outline the functions of the following types of data warehouse applications. **( 6 marks )**
  - a. Information processing
  - b. Analytical processing
  - c. Data mining
- C. Define the term “meta data.” **( 2 marks )**

- D. Explain the following components of the Meta data repository that is a part of a data warehouse. **( 6 marks )**
- a. Business metadata
  - b. Operational metadata
  - c. The algorithms for summarization

**Question Four (20 marks)**

- A. Explain the following types of data mining techniques: **( 10 marks )**
- a. Mining of frequent patterns
  - b. Mining of association
  - c. Mining of correlations
  - d. Evolution analysis
  - e. Outlier analysis
- B. Outline how the K – clustering technique works. **( 3 marks )**
- C. Explain what deidentification of a data set in data mining is and its possible advantages or disadvantages. **( 6 marks )**

**Question Five (20 marks)**

- A. The selection of a data mining system depends on a number of factors. Explain any three factors. **( 6 marks )**
- B. Outline the following statistical methods that are used in Knowledge mining. **( 10 marks )**
- a. Regression
  - b. Analysis of Variance
  - c. Factor Analysis
  - d. Time Series Analysis
  - e. Mixed-effect Models
- C. Explain any two trends in data mining. **( 4 marks )**