



NATIONAL OPEN UNIVERSITY OF NIGERIA
University Village, 91 Cadastral Zone, Nnamdi Azikwe Expressway, Jabi, Abuja
FACULTY OF SCIENCE
2020_2 EXAMINATIONS ...

COURSE CODE: DAM301

COURSE TITLE: Data Mining and Data Warehousing

CREDIT: 3 Units

TIME ALLOWED: 2 ½ Hours

INSTRUCTION: Answer Question 1 and any other FOUR (4) Questions

1. a. Outline the steps in the knowledge discovering in databases process. **(10½ marks)**
 - b. Compare the goals of data mining and that of online analytical processing. **(3 marks)**
 - c. Briefly explain the scope of data mining against the backdrop of the following:
 - i. Automated prediction of trends and behaviours.. **(1 mark)**
 - ii Automated discovery of previously unknown patterns.. **(1 mark)**
 - d. Briefly highlight the following data mining technologies.
 - i. Neural network . **(1½ marks)**
 - ii. K-Nearest Neighbour. **(1½ marks)**
 - iii. Genetic algorithms. **(1½ marks)**
 - e. Differentiate between a data warehouse and a data mart. **(2 marks)**
2. a. Outline in a tabular form, the steps in the evolution of data mining detailing the business question, the enabling technologies, product providers and the characteristics of each evolutionary step. **(9 marks)**
 - b. Distinguish between clustering and regression. **(3 marks)**
3. a. Outline any seven data mining challenges hindering the implementation of data mining. **(7 marks)**
 - b. Relate how multivariate adaptive regression splines (MARS) addressed the shortcomings of classification and regression trees (CART). **(5 marks)**
4. a. Highlight the following data reduction techniques:
 - i. dimensionality reduction **(4 marks)**

- ii. data cube aggregation(**3 marks**)
- iii. data compression.(**2 marks**)
- b. . Outline four techniques of noisy data removal.(**3 marks**)

5.a. Briefly explain the following categories of constraint-based data mining:

- i. knowledge type (**1 mark**)
 - ii. Rule (**1 mark**)
 - iii. Dimension/level (**1 mark**)
 - iv. Data (**1 mark**)
 - v. interestingness(**1 mark**)
- b. Outline the techniques that can be used in mining time series data.(**7 marks**)
- 6.a. Give a broad characterization of data warehouse architecture.(**2 marks**)
- b. Differentiate between OLAP and OLAP server.(**5marks**)
- c. Outline the benefits of OLAP technology.(**5 marks**)