

Total No. of Questions : 8]

[Total No. of Printed Pages : 4

Roll No

MCA-301

M.C.A. III Semester (Two Year Course)

Examination, December 2024

Data Mining

Time : Three Hours

Maximum Marks : 70

Note: i) Attempt any five questions.

किन्हीं पाँच प्रश्नों को हल कीजिए।

ii) All questions carry equal marks.

सभी प्रश्नों के समान अंक हैं।

iii) In case of any doubt or dispute the English version question should be treated as final.

किसी भी प्रकार के संदेह अथवा विवाद की स्थिति में अंग्रेजी भाषा के प्रश्न को अंतिम माना जायेगा।

1. a) Discuss major issues in data mining regarding mining methodology, user interaction, performance and diverse data types.

खनन पद्धति, उपयोगकर्ता इंटरैक्शन, प्रदर्शन और विविध डेटा प्रकारों से संबंधित डेटा खनन में प्रमुख मुद्दों पर चर्चा करें।

b) List out four data mining applications.

डेटा माइनिंग के चार अनुप्रयोगों की सूची बनाइये।

2. a) Explain the different challenges that motivated the development of the data mining technologies.

डेटा माइनिंग प्रौद्योगिकियों के विकास को प्रेरित करने वाली विभिन्न चुनौतियों को समझाइए।

[2]

b) Explain briefly the various stages of knowledge discovery process with the help of an example.

उदाहरण की सहायता से ज्ञान खोज प्रक्रिया के विभिन्न चरणों को संक्षेप में समझाइए।

3. a) Differentiate between operational databases and a data warehouse.

ऑपरेशनल डेटाबेस और डेटा वेयरहाउस के बीच अंतर करें।

b) Discuss briefly data cube computation technique for data warehouse implementation.

डेटा वेयरहाउस कार्यान्वयन के लिए डेटा क्यूब कम्प्यूटेशन तकनीक पर संक्षेप में चर्चा करें।

4. a) Draw a snowflake schema diagram for the data warehouse with suitable example.

उपयुक्त उदाहरण के साथ डेटा वेयरहाउस के लिए स्नोफ्लेक स्कीमा आरेख बनाइए।

b) What are the different types of OLAP server? Explain them.

OLAP सर्वर के विभिन्न प्रकार क्या हैं? उन्हें समझाइये।

5. a) How data cleaning is different from the data transformation?

डेटा क्लीनिंग डेटा रूपांतरण से किस प्रकार भिन्न है?

b) Discuss briefly the various data pre-processing techniques.

विभिन्न डेटा पूर्व-प्रसंस्करण तकनीकों पर संक्षेप में चर्चा करें।

[3]

6. a) A database has four transaction. Let min. support = 60% and min. conf = 80%.

TID	Item Bought
T ₁	{K, A, D, B}
T ₂	{D, A, C, E, B}
T ₃	{C, A, B, E}
T ₄	{B, A, D}

Find all frequent item sets using Apriori algorithm.

एक डेटाबेस में चार लेन-देन हैं। मान लें कि न्यूनतम समर्थन = 60% और न्यूनतम विश्वास = 80% है।

TID	Item Bought
T ₁	{K, A, D, B}
T ₂	{D, A, C, E, B}
T ₃	{C, A, B, E}
T ₄	{B, A, D}

Apriori एल्गोरिथम का उपयोग करके सभी लगातार आइटम सेट खोजें।

- b) Discuss mining of multilevel association rules and explain how to check redundant multilevel association rules?
बहुस्तरीय एसोसिएशन नियमों के खनन पर चर्चा करें और समझाइए कि अनावश्यक बहुस्तरीय एसोसिएशन नियमों की जाँच कैसे करें?
7. a) Specify the criteria for the evaluation of classification and prediction.
वर्गीकरण एवं पूर्वानुमान के मूल्यांकन के लिए मानदंड निर्दिष्ट करें।
- b) Explain K-means algorithm with help of example.
उदाहरण की सहायता से K-मीन्स एल्गोरिथम को समझाइए।

[4]

8. Write short notes on any three of the following:

- Constraint Based Association Mining
- Outlier Analysis
- Gini Index
- OLAP, MOLAP, HOLAP
- Data Integration

निम्नलिखित में से किन्हीं तीन पर संक्षिप्त टिप्पणी लिखें:

- बाधा आधारित एसोसिएशन माइनिंग
- आउटलायर विश्लेषण
- गिनी इंडेक्स
- OLAP, MOLAP, HOLAP
- डेटा एकीकरण
