

CS/IT/CI50031

Roll No. :

Nov. 2023

DATA SCIENCES : DATA WAREHOUSING AND DATA MINING

निर्धारित समय : 3 घंटे]

[अधिकतम अंक : 60

Time allowed : 3 Hours]

[Maximum Marks : 60

नोट : (i) प्रश्नपत्र में तीन सेक्शन ए, बी एवं सी हैं।

Note : There are **three** sections **A, B** and **C** in the paper.

(ii) सेक्शन ए में प्रश्न संख्या 1 के सभी 10 भागों के उत्तर दीजिए। प्रत्येक भाग एक अंक का है एवं सभी 10 भाग वस्तुनिष्ठ प्रकार के प्रश्नों के हैं।

Answer all the 10 parts of the question No. 1 in section A. Each part carries one mark and all 10 parts have objective type questions.

(iii) सेक्शन बी के 8 प्रश्नों में से किन्हीं 6 प्रश्नों के उत्तर दीजिए। प्रत्येक प्रश्न 3 अंक का है एवं इनका 5 लाइन/50 शब्दों में उत्तर दीजिए।

Answer any 6 questions out of the 8 questions in section B. Each question carries 3 marks and to be answered within 5 lines/50 words.

(iv) सेक्शन सी के 6 प्रश्नों में से किन्हीं 4 प्रश्नों के उत्तर दीजिए। प्रत्येक प्रश्न 8 अंक का है एवं इनका 15 लाइन/150 शब्दों में उत्तर दीजिए।

Answer any 4 questions out of the 6 questions in section C. Each question carries 8 marks and to be answered within 15 lines/150 words.

(v) प्रत्येक सेक्शन के सभी प्रश्नों को क्रमवार एक साथ हल कीजिये।

Solve all the questions of a section consecutively together.

(vi) दोनों भाषाओं में अन्तर होने की स्थिति में अंग्रेजी अनुवाद ही मान्य है।

Only English version is valid in case of difference in both the languages.

सेक्शन - ए

SECTION - A

1. (i) निम्न में से किसमें रिकॉर्ड को अपडेट नहीं किया जा सकता है ?

(a) OLTP

(b) फाइल्स

(c) RDBMS

(d) डाटा वेयरहाऊस

In which of the following, record cannot be updated ?

(a) OLTP

(b) Files

(c) RDBMS

(d) Data Warehouse



(ii) OLAP का पूरा नाम है

- (a) ऑनलाइन ऐप्लिकेशन प्रोसेसिंग (b) ऑनलाइन एनेलिटिकल प्रोसेसिंग
(c) ऑनलाइन ऐप्टिट्यूड प्रोसेसिंग (d) ऑनलाइन एग्रीगेट प्रोसेसिंग

OLAP stands for

- (a) Online Application Processing (b) Online Analytical Processing
(c) Online Aptitude Processing (d) Online Aggregate Processing

(iii) डेटा वेयरहाऊस में DSS का पूरा नाम है

- (a) डिसिजन सिंगल सिस्टम (b) डिसिजन सपोर्ट सिस्टम
(c) डिसिजन स्टोर सिस्टम (d) डेटा सपोर्ट सिस्टम

Expansion for DSS in Data warehouse is

- (a) Decision Single System (b) Decision Support System
(c) Decision Store System (d) Data Support System

(iv) डुप्लीकेट रिकॉर्ड्स को हटाने की प्रक्रिया कहलाती है

- (a) डेटा रिकवरी (b) डेटा प्रूनिंग
(c) डेटा क्लिनिंग (d) डेटा क्लस्टरिंग

Removing duplicate records is a process called

- (a) Data Recovery (b) Data Pruning
(c) Data Cleaning (d) Data Clustering

(v) विभिन्न स्रोतों से प्राप्त डाटा को कॉमन फॉर्मेट में बदलने की प्रक्रिया कहलाती है

- (a) सलेक्शन (b) प्रीप्रोसेसिंग
(c) इन्टरप्रिटेशन (d) ट्रांसफॉर्मेशन

Converting data from different sources into a common format is called

- (a) Selection (b) Pre-processing
(c) Interpretation (d) Transformation

(vi) निम्न में से किसमें डेटा को संगृहित, पुनर्प्राप्त और अद्यतन किया जा सकता है ?

- (a) OLTP (b) OLAP
(c) SMTP (d) FTP

In which of the following data can be stored, retrieved and updated ?

- (a) OLTP (b) OLAP
(c) SMTP (d) FTP

(vii) डाटा में से जानकारी खोजने की प्रक्रिया कहलाती है

- (a) डाटा माइनिंग (b) डाटा वेयरहाऊसिंग
(c) क्वेरी (d) उपरोक्त में से कोई नहीं

The process of knowledge discovery from data is called

- (a) Data mining (b) Data warehousing
(c) Query (d) None of the above

(viii) डेटा वेयरहाऊस में पाये जाने वाला डेटा निम्न प्रकार का होता है :

- (a) विषय उन्मुख (b) समयांतर
(c) एकीकृत (d) उपरोक्त सभी

Data found in the data warehouse are of following type :

- (a) Subject oriented (b) Time-variant
(c) Integrated (d) All of the above

(ix) डेटा वेयरहाऊस की त्रिस्तरीय वास्तुकला का द्वितीय स्तर है

- (a) डेटा वेयरहाऊस सर्वर (b) OLAP सर्वर
(c) फ्रंट एण्ड टूल्स (d) बैक एण्ड टूल्स

The second level of three tier architecture of data warehouse is

- (a) Data Warehouse Server (b) OLAP Server
(c) Front End Tools (d) Back End Tools

(x) निम्न में से कौन वेब माइनिंग का/के प्रकार है/हैं ?

- (a) वेब कन्टेन्ट माइनिंग (b) वेब स्ट्रक्चर माइनिंग
(c) वेब उपयोग माइनिंग (d) उपरोक्त सभी

Which of the following is/are the type(s) of web mining ?

- (a) Web content mining (b) Web structure mining
(c) Web Usage mining (d) All of the above

(1×10)

सेक्शन – बी

SECTION – B

2. वर्तमान परिदृश्य में डाटा माइनिंग का क्या महत्त्व है ?

What is the importance of data mining in present scenario ?

(3)

3. 'डाटा क्यूब' क्या है ?

What is 'data cube' ?

(3)

4. 'डाटा माइनिंग' और 'डाटा वेयरहाऊस' को परिभाषित कीजिए ।

Define 'data mining' and 'data warehouse'.

(3)

5. डाटा माइनिंग की एसोसिएशन तकनीक की मूल अवधारणा समझाइए ।

Describe the basic concepts of association technique of data mining.

(3)

6. डाटा माइनिंग की क्लासीफिकेशन तकनीक की मूल अवधारणा समझाइए ।

Describe the basic concepts of classification technique of data mining.

(3)

P.T.O.

7. क्लासीफिकेशन की शुद्धता को बढ़ाने वाली तकनीकों के नाम लिखिए ।
Write the names of techniques to improve the classification accuracy. (3)
8. 'वेब माइनिंग' की मूल अवधारणा को समझाइए ।
Describe the basic concepts of web mining. (3)
9. मॉडल-बेस्ड क्लस्टरिंग एप्रोच को समझाइए ।
Describe the model-based clustering approach. (3)

सेक्शन – सी**SECTION – C**

10. डाटा प्रीप्रोसेसिंग क्या होती है ? डाटा प्रीप्रोसेसिंग के प्रमुख चरणों को समझाइए ।
What is data pre-processing ? Explain the major steps of data pre-processing. (8)
11. 'डाटा' और 'सूचना' को परिभाषित कीजिए । डाटा माइनिंग फंक्शनेलीटिज को समझाइए ।
Define 'Data' and 'Information'. Explain the data mining functionalities. (8)
12. डाटा वेयरहाऊस डिजाइन और उपयोग को समझाइए ।
Explain Data warehouse design and usage. (8)
13. पैटर्न इवेल्यूशन विधियाँ क्या होती हैं ? छः पैटर्न इवेल्यूशन विधियों के नाम लिखिए और उनमें से किन्हीं दो विधियों को समझाइए ।
What are pattern evaluation methods ? Write the names of six pattern evaluation methods and explain any two of them. (8)
14. निर्णय वृक्ष को उदाहरण सहित समझाइए ।
Explain decision tree with suitable example. (8)
15. क्लस्टर विश्लेषण क्या होता है ? क्लस्टर विश्लेषण के पार्टीशनिंग मैथड को समझाइए ।
What is cluster analysis ? Explain the partitioning method of cluster analysis. (8)