

CS/IT50031

Roll No. : .....

Nov. 2022

**DATA SCIENCES : DATA WAREHOUSING AND DATA MINING**

निर्धारित समय : 3 घंटे]

[अधिकतम अंक : 60

Time allowed : 3 Hours]

[Maximum Marks : 60

नोट : (i) प्रश्न-पत्र में तीन सेक्शन-ए, बी एवं सी हैं।

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

(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 कीजिए।

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 कौन सी क्लस्टरिंग विधि नहीं है ?

(a) पार्टिशनिंग विधि

(b) हेरारकीकल विधि

(c) आयतन विधि

(d) घनत्व-आधारित विधि

Which of the following is not a clustering method ?

(a) Partitioning method

(b) Hierarchical method

(c) Volume method

(d) Density based method



(1 of 4)

P.T.O.

(ii) OLTP का पूरा नाम है :

- (a) ऑनलाइन टेक्नीकल प्रोसेसिंग (b) ऑनलाइन ट्रांजेक्शन प्रोसेसिंग  
(c) ऑनलाइन टाइम प्रोसेसिंग (d) उपरोक्त में से कोई नहीं

OLTP stands for

- (a) Online Technical Processing (b) Online Transaction Processing  
(c) Online Time Processing (d) None of the above

(iii) ROLAP का पूरा नाम है :

- (a) रिलेटिव ऑनलाइन एनेलेटिकल प्रोसेसिंग  
(b) रिलेशनल ऑनलाइन एनेलेटिकल प्रोसेसिंग  
(c) रिलेशनल ऑनलाइन एसेसमेंट प्रोसेसिंग  
(d) उपरोक्त में से कोई नहीं

ROLAP stands for

- (a) Relative Online Analytical Processing  
(b) Relational Online Analytical Processing  
(c) Relational Online Assessment Processing  
(d) None of the above

(iv) त्रुटिपूर्ण, गलत या लुप्त डाटा के साथ व्यवहार कहलाता है

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

Treating with erroneous, incorrect or missing data is called

- (a) Selection (b) Preprocessing  
(c) Transformation (d) Interpretation

(v) KDD बताता है

- (a) डाटा में से जानकारी का सार निकालने की संपूर्ण प्रक्रिया ।  
(b) डाटा का सार निकालना ।  
(c) सूचना का सार निकालना ।  
(d) नियमों का सार निकालना ।

KDD describes the

- (a) Whole process of extraction of knowledge from data  
(b) Extraction of data  
(c) Extraction of information  
(d) Extraction of rules

(vi) निम्न में से कौन वर्तमान तथा पुराने दोनों ट्रांजेक्शन्स को संभालता है ?

- (a) OLTP (b) OLAP  
(c) स्प्रेडशीट (d) उपरोक्त में से कोई नहीं

Which of the following manages both current and historic transactions ?

- (a) OLTP (b) OLAP  
(c) Spreadsheet (d) None of the above

(vii) निम्न में से कौन डेटा एट्रीब्यूट का प्रकार है ?

- (a) नॉमिनल (b) बाइनरी  
(c) ऑर्डिनल (d) उपरोक्त सभी

Which of the following is/are type(s) of data attribute ?

- (a) Nominal (b) Binary  
(c) Ordinal (d) All of the above

(viii) डाटा वेयरहाउस होता है :

- (a) केवल पढ़ने योग्य (b) केवल लिखने योग्य  
(c) दोनों पढ़ने-लिखने योग्य (d) उपरोक्त में से कोई नहीं

The data warehouse is

- (a) read only (b) write only  
(c) read-write both (d) None of the above

(ix) डाटा वेयरहाउस की त्रिस्तरीय वास्तुकला का तृतीय स्तर है :

- (a) डाटा वेयरहाउस सर्वर (b) OLAP सर्वर  
(c) फ्रंट एण्ड टूल्स (d) बैक एण्ड टूल्स

The third 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) क्लासीफिकेशन की शुद्धता बढ़ाने के लिए निम्न में से कौन सी तकनीक का उपयोग किया जाता है ?

- (a) एन्सेम्बल मैथड्स (b) सपोर्ट वेक्टर मशीन  
(c) पैटर्न-बैस्ड क्लासीफिकेशन (d) उपरोक्त सभी

Which of the following technique(s) is/are used to improve classification accuracy ?

- (a) Ensemble methods (b) Support vector machines  
(c) Pattern based classification (d) All of the above

(1×10)

### सेक्शन - बी

#### SECTION - B

2. डाटा माइनिंग टास्क प्रिमीटिव्स को समझाइए।

Describe Data mining task primitives.

(3)

3. OLAP क्या है ? समझाइए।

What is OLAP ? Describe.

(3)

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

Define 'Data mining' and 'Data warehouse'.

(3)

P.T.O.

5. डाटा माइनिंग की 'कोरिलेशन' तकनीक की मूल अवधारणा समझाइए ।  
Describe the basic concept of correlations technique in data mining. (3)
6. 'सपोर्ट वेक्टर मशीन' क्या है ? समझाइए ।  
What is 'support vector machine' ? Describe. (3)
7. 'पैटर्न-बेस्ड क्लासिफिकेशन' क्या होता है ? समझाइए ।  
What is 'pattern based classification' ? Describe. (3)
8. आउटलायर्स क्या होता है ? आउटलायर्स के प्रकार बताइए ।  
What are outliers ? Write the types of outliers. (3)
9. वेब माइनिंग के विभिन्न प्रकार बताइए ।  
Describe the different types of web mining. (3)

### सेक्शन – सी

### SECTION – C

10. डाटा वेयरहाउस के त्रिस्तरीय वास्तुकला का सचित्र वर्णन कीजिए ।  
Explain the three-tier architecture of a Data warehouse with diagram. (8)
11. फ्रीक्वेंट पैटर्न माइनिंग क्या है ? कॉन्स्ट्रैन्ट-बेस्ड फ्रीक्वेंट पैटर्न माइनिंग को समझाइए ।  
What is frequent pattern mining ? Explain constraint – based frequent pattern mining. (8)
12. रूल-बेस्ड क्लासिफिकेशन को समझाइए ।  
Explain rule based classification. (8)
13. न्यूरल नेटवर्क क्लासिफिकेशन को समझाइए ।  
Explain neural network classification. (8)
14. वेबमाइनिंग क्या है ? पेज रैंक एल्गोरिदम को समझाइए ।  
What is web mining ? Explain page rank algorithm. (8)
15. डाटा प्रीप्रोसेसिंग क्या होती है ? डाटा प्रीप्रोसेसिंग के प्रमुख चरणों को समझाइए ।  
What is Data preprocessing ? Explain the major steps of data preprocessing. (8)

