

- All questions are compulsory. (सभी प्रश्न अनिवार्य हैं।)
- Marks are mentioned on the right side of each question. (अंक सभी प्रश्न के दाईं ओर अंकित किये हैं।)

Group (A) (गुप-ए)

Q.1 Choose the most suitable answer for the following options. (सर्वाधिक उपयुक्त विकल्प को चुनकर लिखें।) :- (1*20=20)

- i. What is the primary motivation behind data mining? (डेटा माइनिंग के पीछे प्राथमिक प्रेरणा क्या है ?)
 (a) to collect data for research purposes (अनुसंधान उद्देश्यों के लिए डेटा एकत्र करना)
 (b) to identify and extract valuable pattern from data (डेटा से मूल्यवान पैटर्न की पहचान करना और निकालना)
 (c) to create visualization for reporting (रिपोर्टिंग के लिए विजुअलाइजेशन बनाना)
 (d) to store data in a data warehouse (डेटा वेयरहाउस में डेटा संग्रहीत करना)
- ii. Which data mining task primitive involves computing summary statistics for data? (किस डेटा माइनिंग कार्य आदिम में डेटा के लिए सारांश आंकड़ों की गणना शामिल है?)
 (a) pattern discovery (पैटर्न की खोज) (b) association mining (एसोसिएशन माइनिंग) (c) data summarization (डेटा सारांशिकरण) (d) data integration (डेटा एकीकरण)
- iii. What is the process of integrating a data mining system with a database or data warehouse called? (किसी डेटा माइनिंग सिस्टम को डेटाबेस या डेटा वेयरहाउस के साथ एकीकृत करने की प्रक्रिया को क्या कहा जाता है?)
 (a) data augmentation (डेटा संवर्धन) (b) data enrichment (डेटा इनरिचमेंट) (c) data mining integration (डेटा माइनिंग एकीकरण) (d) data warehousing (डेटा भण्डारण)
- iv. Which term refers to the basic unit of data in data mining? (डेटा माइनिंग में डेटा की मूल इकाई को कौन सा शब्द संदर्भित करता है?)
 (a) dataset (डेटासेट) (b) instance (उदाहरण) (c) observation (अवलोकन) (d) feature (विशेषता)
- v. Data visualization is used to : (डेटा विजुअलाइजेशन का उद्देश्य है :-)
 (a) hide patterns in data (डेटा में पैटर्न को छुपाना)
 (b) simplify data analysis (डेटा एनालिसिस को सरल बनाना)
 (c) remove outliers from data (डेटा से आउटलायर को हटाना)
 (d) present data in a visually meaningful way (डेटा को दृश्यमय रूप से दिखाना)
- vi. Data warehousing modeling involves the creation of : (डेटा वेयरहाउसिंग मॉडलिंग में किसका निर्माण शामिल है :-)
 (a) data cubes (डेटा क्यूब्स) (b) data charts (डेटा चार्ट) (c) data pipelines (डेटा पाइपलाइन) (d) data clusters (डेटा क्लस्टर)
- vii. What is the full form of OLAP? (OLAP का पूर्ण रूप क्या है ?)
 (a) online assessment processing (ऑनलाइन एसेसमेंट प्रोसेसिंग) (c) online analytical processing (ऑनलाइन एनालिटिकल प्रोसेसिंग)
 (b) online application program (ऑनलाइन एप्लीकेशन प्रोग्राम) (d) offline analysis processing (ऑफलाइन एनालिसिस प्रोसेसिंग)

- viii. OLAP is used for : (OLAP का उपयोग किसके लिए होता है :-)
 (a) real time data collection (रियल टाइम डेटा कलेक्शन) (c) analyzing historical data (ऐतिहासिक डेटा का विश्लेषण)
 (b) generating financial reports (वित्तीय रिपोर्ट उत्पन्न करने में) (d) creating data warehouses (डेटा वेयर हाउस का निर्माण)
- ix. Which method is used to evaluate the quality of discovered patterns? (खोजे गए पैटर्न की गुणवत्ता का मूल्यांकन करने के लिए किस विधि का उपयोग किया जाता है ?)
 (a) data visualization (डेटा विजुअलाइजेशन) (b) pattern evaluation (पैटर्न इवैल्यूएशन) (c) data summarization (डेटा सारांशिकरण) (d) data transformation (डेटा ट्रान्सफॉर्मेशन)
- x. Decision tree induction is a technique used for (डिसीजन ट्री इंडक्शन एक ऐसी तकनीक है जिसका उपयोग किया जाता है)
 (a) grouping similar data records (समान डेटा रिकॉर्ड को समूहीकृत करना)
 (b) predicting future data values (भविष्य के डेटा मूल्यों की भविष्यवाणी करे)
 (c) identifying frequent patterns (बारंबार पैटर्न की पहचान करना)
 (d) clustering data instances (क्लस्टरिंग डेटा उदाहरण)
- xi. Cluster analysis is used for (क्लस्टर एनालिसिस की उपयोगिता है -)
 (a) identifying outlier data points (बाहरी डेटा बिंदुओं को पहचानना)
 (b) group similar data records together (समान डेटा रिकॉर्ड को एक साथ समूहित करें)
 (c) predict future data values (भविष्य के डेटा मूल्यों की भविष्यवाणी करें।)
 (d) categorize data instances into classes (डेटा इन्स्टेंस को कक्षाओं में वर्गीकृत करें)
- xii. Which major clustering approaches focuses on partitioning data into distinct groups? (कौन सा प्रमुख क्लस्टरिंग दृष्टिकोण डेटा को अलग-अलग समूहों में विभाजित करने पर केंद्रित है?)
 (a) Hierarchical methods (हिरार्चिकल मेथड्स) (b) Density based methods (डेंसिटी बेस्ड मेथड्स) (c) Model based clustering (मॉडल बेस्ड क्लस्टरिंग) (d) Partitioning methods (पार्टिशनिंग मेथड्स)
- xiii. What is web mining in data science? (डेटा साइंस में वेब माइनिंग क्या है ?)
 (a) extracting precious metals from earth crust (पृथ्वी की पपड़ी से बहुमूल्य धातुएं निकालना)
 (b) analyzing patterns in weather data (मौसम डेटा में पैटर्न का विश्लेषण)
 (c) discovering insights from web related data (वेब संबंधित डेटा से अंतर्दृष्टि की खोज करना)
 (d) none of above (उपरोक्त में कोई नहीं)
- xiv. What does the PAGE RANK algorithm measure? (पेजरेक एल्गोरिथम क्या मापता है ?)
 (a) web page load times (वेब पेज लोड टाइम) (c) relevance and importance of web pages (वेबपेज की रिलेवेंस और जरूरत)
 (b) user engagement on websites (वेबसाइट पर यूजर एंगेजमेंट) (d) web page aesthetics (वेब पेज एस्टेटिक्स)
- xv. What is the purpose of data preprocessing in data science? (डेटा साइंस में डेटा प्रीप्रोसेसिंग का उद्देश्य क्या है ?)
 (a) generating large datasets (बड़े डेटासेट तैयार करना)
 (b) creating complex visualizations (जटिल विजुअलाइजेशन बनाना)
 (c) enhancing data security (डेटा सुरक्षा बढ़ाना)
 (d) cleaning and preparing data for analysis (विश्लेषण के लिए डेटा की सफाई और तैयारी)
- xvi. Data reduction techniques aim to (डेटा कटौती तकनीकों का लक्ष्य है?)
 (a) increase the size of the dataset (डेटासेट का आकार बढ़ाना)
 (b) expand the number of features (सुविधाओं की संख्या का विस्तार करे)
 (c) remove irrelevant or redundant data (अप्रसंगिक या अनावश्यक डेटा हटाएँ।)
 (d) replace missing values with random values (तुल्य मानों को यादृच्छिक मानों से बदलें)

- xviii. Data discretization is used to (डेटा विवेकीकरण का उपयोग इसके लिए किया जाता है)
 (a) replace numeric data with categorical labels (संख्यात्मक डेटा का श्रेणीबद्ध लेबल से बदलें)
 (b) aggregate data into a single value (डेटा को एक मान में एकत्रित करें)
 (c) remove noisy data points (शोर वाले डेटा बिंदु हटाएं)
 (d) split data into training and testing sets (डेटा को प्रशिक्षण और परीक्षण सेट में विभाजित करें)

xviii. What is the full form of KDD? (KDD का पूर्ण रूप क्या क्या है ?)

- (a) knowledge discovery database (ज्ञान डिस्कवरी डेटाबेस) (c) knowledge data definition (ज्ञान डेटा परिभाषा)
 (b) knowledge discovery data (ज्ञान डिस्कवरी डेटा) (d) knowledge data house (ज्ञान डेटा हाउस)

xix. ETL stands for (ETL का पूर्ण रूप है :-)

- (a) extract, transfer and load (एक्सट्रैक्ट, ट्रांसफर और लोड)
 (b) extract, transform and load (एक्सट्रैक्ट, ट्रांसफॉर्म और लोड)
 (c) explain, transfer and load (एक्सप्लेन, ट्रांसफर और लोड)
 (d) effect, transfer and load (एफेक्ट, ट्रांसफर और लोड)

xx. DSS in data warehouse stands for (डेटा वेयरहाउस में DSS का पूर्ण रूप है :-)

- (a) decision single system (डिसेजन सिंगल सिस्टम) (b) decision support system (डिसेजन सपोर्ट सिस्टम) (c) data support system (डेटा सपोर्ट सिस्टम) (d) data storage system (डेटा स्टोरेज सिस्टम)

Group (B) (ग्रुप - बी)

- Q.2 Explain importance of data mining in various fields. (विभिन्न क्षेत्रों में डेटा माइनिंग के महत्व की व्याख्या करें।) 4
 OR (अथवा)
 Define data mining and list its main functionalities. (डेटा माइनिंग को परिभाषित करें और इसकी मुख्य कार्यात्मकताओं को सूचीबद्ध करें।) 4
 Q.3 Differentiate between structured and unstructured data. Provide example of each. (संरचित और असंरचित डेटा के बीच अंतर करें। प्रत्येक के उदाहरण प्रदान करें।) 4
 OR (अथवा)
 Explain the different kinds of pattern that can be mined from data. (डेटा से निकाले जा सकने वाले विभिन्न प्रकार के पैटर्न की व्याख्या करें।) 4
 Q.4 Describe the classification of data mining system based on functionalities. (कार्यप्रणाली के आधार पर डेटा माइनिंग सिस्टम के वर्गीकरण का वर्णन करें।) 4
 OR (अथवा)
 What are data mining task primitives? Give example of commonly used primitives. (डेटा माइनिंग कार्य आदिम क्या है? आमतौर पर प्रयुक्त आदिमों के उदाहरण दीजिए।) 4
 Q.5 Discuss the integration of a data mining system with a database or database warehouse system. (डेटाबेस या डेटा वेयरहाउस सिस्टम के साथ डेटा माइनिंग सिस्टम के एकीकरण पर चर्चा करें।) 4
 OR (अथवा)
 Identify and elaborate on the major issues in data mining. (डेटा माइनिंग में प्रमुख मुद्दों को पहचानें और उनके बारे में विस्तार से बताएं।) 4

- Q.6 Define data sets and attribute values. Explain the concepts of attribute types. (डेटा सेट और विशेषता मान परिभाषित करें। विशेषता प्रकारों की अवधारणा को समझाएं।) 4

OR (अथवा)

Discuss the significance of data visualization in data analysis. (डेटा विश्लेषण में डेटा विजुअलाइजेशन के महत्व पर चर्चा करें।) 4

Group (C) (ग्रुप - सी)

- Q.7 Explain the concept of measuring data similarity. Provide examples of similarity measures. (डेटा समानता मापने की अवधारणा को स्पष्ट करें। समानता उपायों के उदाहरण प्रदान करें।) 6

OR (अथवा)

Describe the major tasks in data preprocessing. Why is data preprocessing essential in data mining? (डेटा प्रिप्रोसेसिंग में प्रमुख कार्यों का वर्णन करें। डेटा माइनिंग में डेटा प्रिप्रोसेसिंग क्यों आवश्यक है?) 6

- Q.8 Explain the concept of data cube and online analytical processing. (डेटा क्यूब और ऑनलाइन विश्लेषणात्मक प्रोसेसिंग (OLAP) की अवधारणा को समझाएं।) 6

OR (अथवा)

How is data generalization achieved through attribute oriented induction? Provide an example. (विशेषता उन्मुख प्रेरण के माध्यम से डेटा सामान्यीकरण कैसे प्राप्त किया जाता है? एक उदाहरण प्रदान करें।) 6

- Q.9 Describe pattern evaluation methods and their importance in association mining. (एसोसिएशन माइनिंग में पैटर्न मूल्यांकन विधियों और उनके महत्व का वर्णन करें।) 6

OR (अथवा)

- a) Define classification and its significance in data mining. (अ) डेटा माइनिंग में वर्गीकरण और उसके महत्व को परिभाषित करें।
 b) How do neural networks contribute to classification? Provide an example. (ब) तंत्रिका नेटवर्क वर्गीकरण में कैसे योगदान करते हैं? एक उदाहरण प्रदान करें।) 6

- Q.10 Discuss the concept of lazy learners in classification. How do they differ from eager learners? (वर्गीकरण में लेजी लर्नर की अवधारणा पर चर्चा करें। वे इंगर लर्नर से किस प्रकार भिन्न हैं?) 6

OR (अथवा)

Explain the different clustering structures and approaches in cluster analysis. (क्लस्टर विश्लेषण में विभिन्न क्लस्टरिंग संरचनाओं और दृष्टिकोणों की व्याख्या करें।) 6

- Q.11 a) Explain the page rank algorithm and its role in web mining. (अ) पेजरेक एल्गोरिथम और वेब माइनिंग में इसकी भूमिका की व्याख्या करें।
 b) Describe the HITS algorithm and its significance in web mining. (ब) हिट्स (हाइपरलिक प्रेरित विषय खोज) एल्गोरिथम और वेब माइनिंग में इसके महत्व का वर्णन करें।) 6

OR (अथवा)

Discuss Bayesian classification methods and their application in data analysis. (बेयसियन वर्गीकरण विधियों और डेटा विश्लेषण में उनके उपयोगों पर चर्चा करें।) 6

-----*****-----