To retrieve information from large voluminous database about current and update details is called as * 1 point

- A. On Line Transaction Processing(OLTP)
- B. On Line Analytical Processing(OLAP)
- C. On Line Data Processing(OLDP)
- D. On Line Privacy Processing(OLPP)

The comparison of the general features of the target class data objects against the general features of objects from one or multiple contrasting classes. * 1 point

- A) Data Characterization
- B) Association and Coorelation
- C) Data Discrimination
- D) Regression

The example of symmetric binary attribute is * 1 point

- a) Pass and Fail
- b) Positive and Negative
- c) Male and Female
- d) True and False

Suppose that the data for analysis include the attributes age. The age value for the data tuples aregiven:13,15,16,16,19,20,20,21,22,22,25,25,25,25,30,33,33,35,35,35,35,36,40 ,45,46 ,52,70 Find IQR and Outliers for the given data value. * 1 point

- (A) 18, 58
- (B)15, 70
- (C) 22, 52
- (D) 22, 70

Name *

Harsh Bardhan

---------- techniques can be applied to obtained a reduced representation of the dataset that is much smaller in volume, yet closely maintains the integrity of the original data. * 1 point

- a) Data Transformation
- b) Data Reduction
- c) Data Integration
- d) Attribute Transformation

Match the data preprocessing techniques: (i) Data cleaning - (a) Discretization (ii) Data Integration - (b) Regression (iii) Data Reduction - (c) Data value conflict detection and resolution (iv) Data Transformation - (d) Principal Component analysis * 1 point

- (A) i – c, ii – d, iii – a, iv-b
- (B) i – d,ii – b,iii – a,iv – c
- (C) i – b, ii – c, iii – d, iv – a
- (D) i – b , ii – d , iii – a , iv – c

The output of KDD is * 1 point

- a) Data
- b) Information
- c) Query
- d) Useful Information

---------- is a numeric attribute with an inherent zero-point * 1 point

- a) Interval-Scaled attribute
- b) Ratio-Scaled Attribute
- c) Time-Scaled Attribute
- d) Point-Scaled Attribute

---------- provides a multidimensional view of data and allows the precomputation and fast access of summarized data * 1 point

- A. Data Warehouse
- B. Data Mart
- C. Data Cube
- D. Business Intelligence

Use these method to normalize the following group of data :200, 300, 400, 600, 1000Use min-max normalization by setting min=0, and max=1 for the value 600 * 1 point

- (A) 0.4
- (B) 0.5
- (C) 0.6
- (D) 0.7

The data mining is a process of extracting and discovering patterns in large dataset. And it also known as * 1 point

- A. Pattern Recognition
- B. Knowledge discover from data
- C. Business Intelligence
- D. Process mining

The problem of finding abstracted patterns in the unlabelled data * 1 point

- A. Classification
- B. Clustering
- C. Outlier analysis
- D. Regression

------- allows the system to identify patterns within data sets on its own. * 1 point

- A) Supervised Learning
- B) Unsupervised Learning
- C) Regression
- D) Pattern Recognition

An attribute with possible values that have a meaningful order or ranking among them, but the magnitude between successive values is not known * 1 point

- (A) Binary attribute
- (B) Nominal attribute
- (C) Numeric attribute
- (D) Ordinal attribute

A data objects that do not comply with the general behavior or model is called--------- * 1 point

- (●) A. Outliers
- ( ) B. Prediction
- ( ) C. Evolution analysis
- ( ) D. Classification

The analysis tools pre-compute the summaries of the huge amount of data for what purpose? * 1 point

- ( ) a) In order to maintain consistency
- ( ) b) For authentication
- ( ) c) For data access
- (●) d) To obtain the queries response

Year and Section *

3rd F1

Compute and Approximate median value for the following data.200,450,300,1500,700,44 * 1 point

- ( ) a) 525
- ( ) b) 515
- (●) c) 350
- ( ) d) 500

Find covariance analysis of given numeric attribute values * 1 point

Time Point	All Electronics	High Tech
T1	6	20
T2	5	10
T3	4	14
T4	3	5
T5	2	5

- (●) a) 7
- ( ) b) 8
- ( ) c) 9
- ( ) d) 10

Register no *

RA1911003010358

The various way to represent attributes in different discipline i) Data Warehouse - (a)Variable ii) Machine Learning – (b)dimension iii) Statistics – (c)attribute iv) Data mining-(d)Features * 1 point

- [x] a) i)-b, ii)-d, iii)-a, iv)-c
- [ ] b) i)-a, ii)-d, iii)-c, iv)-d
- [ ] c) i)-a, ii)-b, iii)-c, iv)-d
- [ ] d) i)- c, ii) d, iii)-d, iv)-a

The Mode occurs at a value greater than the medium * 1 point

- [ ] a) Positively skewed data
- [x] b) Negatively skewed data
- [ ] c) Symmetric data
- [ ] d) Variance of data

Give the categories of data mining task * 1 point

- [ ] a) Characterization and Discrimination
- [x] b) Predictive and Descriptive
- [ ] c) Association and correlation
- [ ] d) Classification and Regression

Find Correlation analysis of given nominal attribute value using chi-square test. * 1 point

Qualification/Marital Status	Middle School	High School	Bachelor's	Master's	Ph.D	Total
Never Married	18	36	21	9	6	90
Married	12	36	45	36	21	150
Divorced	6	9	9	3	3	30
Widowed	3	9	9	6	3	30
Total	39	90	84	54	33	300

- a) 23.57
- b) 22.54
- c) 36.88
- d) 24

"Handling complex types of data" in data mining algorithms issues comes under * 1 point

- A. Diverse Data type
- B. Descriptive data type
- C. Mining Methodology and user interaction issue
- D. Performance issue

A copy of your responses will be emailed to hb6269@srmist.edu.in.

Submit