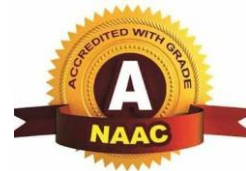




ZEAL EDUCATION SOCIETY'S  
**ZEAL COLLEGE OF ENGINEERING AND RESEARCH**  
NARHE | PUNE -41 | INDIA  
**DEPARTMENT OF COMPUTER ENGINEERING**



<b>Name of the Teacher: V. R. Vasekar</b>	
<b>Class: BE</b> <b>AY: 2020-21</b>	<b>Subject: Data Mining and Warehousing</b> <b>SEM: I</b>
<b>UNIT-1</b>	
1)	Binary attribute are
	a) This takes only two values. In general, these values will be 0 and 1 and .they can be coded as one bit b) The natural environment of a certain species c) Systems that can be used without knowledge of internal operations d) None of these
<b>Ans:</b>	<b>a</b>
<b>Explanation:</b>	All statement are true about Machine Learning.
2)	“Efficiency and scalability of data mining algorithms” issues come under?
	a) Mining Methodology and User Interaction Issues b) Performance Issues c) Diverse Data Types Issues d) None of the above
<b>Ans:</b>	<b>b</b>
<b>Explanation:</b>	<b>In order to effectively extract the information from huge amount of data in databases, data mining algorithm must be efficient and scalable.</b>
3)	—— is not a data mining functionality?
	a) Clustering and Analysis b) Selection and interpretation c) Classification and regression Characterization and Discrimination
<b>Ans:</b>	<b>b</b>
<b>Explanation:</b>	Selection and interpretation
4)	—— is the output of KDD
	a) Query b) Data c) Useful Information d) information
<b>Ans:</b>	<b>c</b>
<b>Explanation:</b>	Useful Information
5)	Which of the following is not belong to data mining? it is unsupervised learning ?
	a) Knowledge extraction b) Data archaeology c) Data exploration d) Data transformation

<b>Ans:</b>	<b>d</b>
<b>Explanation:</b>	Data transformation
6)	Which of the following is the right approach to Data Mining?
	e) Infrastructure, exploration, analysis, exploitation, interpretation f) Infrastructure, exploration, analysis, interpretation, exploitation g) Infrastructure, analysis, exploration, interpretation, exploitation None of these
<b>Ans:</b>	<b>b</b>
<b>Explanation:</b>	Infrastructure, exploration, analysis, interpretation, exploitation
7)	<b>Background knowledge referred to</b>
	a) Additional acquaintance used by a learning algorithm to facilitate the learning process b) A neural network that makes use of a hidden layer c) It is a form of automatic learning. d) None of these
<b>Ans:</b>	<b>a</b>
<b>Explanation:</b>	Additional acquaintance used by a learning algorithm to facilitate the learning process
8)	<b>Data mining is</b>
	a) The actual discovery phase of a knowledge discovery process b) The stage of selecting the right data for a KDD process c) A subject-oriented integrated time variant non-volatile collection of data in support of management d) None of these
<b>Ans:</b>	<b>a</b>
<b>Explanation:</b>	The actual discovery phase of a knowledge discovery process
09)	<b>Data selection is</b>
	a) The actual discovery phase of a knowledge discovery process b) The stage of selecting the right data for a KDD process c) A subject-oriented integrated time variant non-volatile collection of data in support of management d) None of these
<b>Ans:</b>	<b>b</b>
<b>Explanation:</b>	The stage of selecting the right data for a KDD process
10)	The Example of nominal attribute is
	a) Hair_color b) smoker c) temperature d) drink size

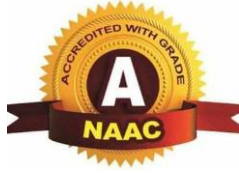
<b>Ans:</b>	<b>a</b>
<b>Explanation:</b>	Nominal means “relating to names.” The values of a nominal attribute are symbols or names of things
11)	The Example of binary attribute is
	<ul style="list-style-type: none"> <li>a) gender</li> <li>b) drink_size</li> <li>c) tempertaure</li> <li>d) professionl_rank</li> </ul>
<b>Ans</b>	<b>b</b>
<b>Explanation:</b>	A binary attribute is a nominal attribute with only two categories or states:0 or1
12)	The Example of ordinary attribute is
	<ul style="list-style-type: none"> <li>a) Years_of_experience</li> <li>b) age</li> <li>c) occupation</li> <li>d) customer_id</li> </ul>
<b>Ans:</b>	<b>b</b>
<b>Explanation:</b>	An ordinal attribute is an attribute with possible values that have a meaningful order or ranking among them
13)	Data cleaning includes_____
	<ul style="list-style-type: none"> <li>a. Handling missing values and noisy data</li> <li>b. Reduction of attributes</li> <li>c. Relevant attribute selection</li> <li>d. Sample data selection</li> </ul>
<b>Ans:</b>	<b>a</b>
<b>Explanation:</b>	Data cleaning (or data cleansing) routines attempt to fill in missing values, smooth out noise while identifying outliers, and correct inconsistencies in the data.
14)	To deal with missing values, the following strategy is used__
	<ul style="list-style-type: none"> <li>e. Use a measure of central tendency</li> <li>f. Reduction of attribute</li> <li>g. Sample data selection</li> <li>h. Data converted into other form</li> </ul>
<b>Ans:</b>	<b>a</b>
<b>Explanation:</b>	measures of central tendency, which indicate the “middle” value of a data distribution
15)	Noise is ____
	<ul style="list-style-type: none"> <li>a) Missing value from dataset</li> <li>b) Inaccurate data</li> <li>c) a random error or variance in a measured variable</li> <li>d) the data whose value known to user</li> </ul>
<b>Ans:</b>	<b>c</b>
<b>Explanation:</b>	

16)	At the time of data integration following problem occurs ____
	<ul style="list-style-type: none"> <li>a) Selection of proper values</li> <li>b) Raw data conversion</li> <li>c) Entity identification</li> <li>d) Attribute subset selection</li> </ul>
<b>Ans:</b>	<b>c</b>
<b>Explanation:</b>	Schema integration and object matching can be tricky.
17)	Which of the following is not example of data reduction strategy?
	<ul style="list-style-type: none"> <li>a) Outlier detection</li> <li>b) Principal Component Analysis</li> <li>c) Attribute subset selection</li> <li>d) Wavelet transforms</li> </ul>
<b>Ans:</b>	<b>a</b>
<b>Explanation:</b>	Outlier detection
18)	Data Transformation Strategies includes ____
	<ul style="list-style-type: none"> <li>a) smoothing</li> <li>b) Attribute construction</li> <li>c) Normalization</li> <li>d) All of the above</li> </ul>
<b>Ans:</b>	<b>d</b>
<b>Explanation:</b>	Smoothing, attribute construction and normalization includes in data transformation
19)	Data Discretization is used for ____
	<ul style="list-style-type: none"> <li>a) transforms numeric data by mapping values to interval or concept labels</li> <li>b) smoothing</li> <li>c) Attribute construction</li> <li>d) Normalization</li> </ul>
<b>Ans:</b>	<b>a</b>
<b>Explanation:</b>	transforms numeric data by mapping values to interval or concept labels
20)	KDD stands for
	<ul style="list-style-type: none"> <li>a) K data values</li> <li>b) Knowledge discovery from dataset</li> <li>c) K dataset</li> <li>d) None of the above</li> </ul>
<b>Ans.</b>	<b>b</b>
<b>Explanation</b>	Knowledge discovery from dataset
21)	<b>Data transformation includes:</b>

	<ul style="list-style-type: none"> <li>a) data are transformed and consolidated into forms appropriate for mining by performing summary or aggregation operations</li> <li>b) an essential process where intelligent methods are applied to extract data patterns</li> <li>c) data relevant to the analysis task are retrieved from the database</li> <li>d) it is used for knowledge representation.</li> </ul>
<b>Ans</b>	<b>a</b>
<b>Explanation</b>	data are transformed and consolidated into forms appropriate for mining by performing summary or aggregation operations
<b>22)</b>	<b>Pattern evaluation includes__</b>
	<ul style="list-style-type: none"> <li>a) data are transformed and consolidated into forms appropriate for mining by performing summary or aggregation operations</li> <li>b) an essential process where intelligent methods are applied to extract data patterns</li> <li>c) data relevant to the analysis task are retrieved from the database</li> <li>d) Identify the truly interesting patterns representing knowledge based on interestingness measures</li> </ul>
<b>Ans</b>	<b>d</b>
<b>Explanation</b>	To identify the truly interesting patterns representing knowledge based on interestingness measures
<b>23)</b>	<b>In KDD, the knowledge representation term used for__</b>
	<ul style="list-style-type: none"> <li>a) data are transformed and consolidated into forms appropriate for mining by performing summary or aggregation operations</li> <li>b) an essential process where intelligent methods are applied to extract data patterns</li> <li>c) visualization and knowledge representation techniques are used to present mined knowledge to users</li> <li>d) Identify the truly interesting patterns representing knowledge based on interestingness measures</li> </ul>
<b>Ans</b>	<b>c</b>
<b>Explanation</b>	visualization and knowledge representation techniques are used to present mined knowledge to users
<b>24)</b>	<b>Data mining functionalities are used to__</b>
	<ul style="list-style-type: none"> <li>a) to specify the kinds of patterns or knowledge to be found in data mining tasks</li> <li>b) to select data</li> <li>c) to find missing values</li> <li>d) to analyze the mining result</li> </ul>
<b>Ans</b>	<b>a</b>
<b>Explanation</b>	a) Data mining functionalities are used to specify the kinds of patterns or



ZEAL EDUCATION SOCIETY'S  
**ZEAL COLLEGE OF ENGINEERING AND RESEARCH**  
NARHE | PUNE -41 | INDIA  
**DEPARTMENT OF COMPUTER ENGINEERING**



	knowledge to be found in data mining tasks
<b>25)</b>	The challenging issues in data mining research_____
	a) efficiency and scalability b) dealing with diverse data types c) user interaction d) all of the above
<b>Ans</b>	d
<b>Explanation</b>	There are many challenging issues in data mining research. Areas include mining methodology, user interaction, efficiency and scalability, and dealing with diverse data types. Data mining research has strongly impacted society and will continue to do so in the future

Name and Sign of Subject Teacher

<b>Name of the Teacher: V. R. Vasekar</b>	
<b>Class: BE</b> <b>AY: 2020-21</b>	<b>Subject: Data Mining and Warehousing</b> <b>SEM: II</b>
<b>UNIT-2 Data Warehouse</b>	
1. ____ is a subject oriented, integrated, time variant, non-volatile collection of data in support of management decisions.	
	a) Data Mining b) Data Warehousing c) Web mining d) Text mining
<b>Ans:</b>	<b>b</b>
<b>Explanation:</b>	Data Warehousing
2. Data Warehouse is	
	a) Read only b) Write only c) Read and write only d) none
<b>Ans:</b>	<b>a</b>
<b>Explanation:</b>	Because of historical data storage
3. Expansion for DSS in DW is ____	
	a) Decision Single System b) Decision storable system c) Decision Support System d) Data Support System
<b>Ans:</b>	<b>c</b>
<b>Explanation:</b>	Decision support system
4. The important aspect of data warehouse environment is that data found within the data warehouse is ____	
	a) Subject oriented b) Time-variant c) Integrated d) All of the above
<b>Ans:</b>	<b>d</b>
<b>Explanation:</b>	All are correct
5. The time horizon in Data warehouse is usually ____	
	a) 1-2 year b) 3-4 year c) 5-6 years d) 5-10 years
<b>Ans:</b>	<b>d</b>
<b>Explanation:</b>	5 to 10 years
6. The data is stored , retrieved and updated in ____	

	a) OLAP b) OLTP c) SMTP d) FTP
<b>Ans:</b>	<b>b</b>
<b>Explanation:</b>	Online Analytical Transaction processing
7. ___describes the data oriented in the data warehouse	
	a) Relational data b) Operational data c) Metadata d) Informational data
<b>Ans:</b>	<b>c</b>
<b>Explanation:</b>	metadata
8. ___ predicts the future trends and behaviours, allowing business managers to make proactive knowledge-driven decisions	
	a) Data warehouse b) Data mining c) Datamarts d) metadata
<b>Ans:</b>	<b>b</b>
<b>Explanation:</b>	
9. ___ is the heart of Datawarehouse	
	a) Data mining database server b) Data warehouse database servers c) Data mart database servers d) Relational database servers
<b>Ans:</b>	<b>b</b>
<b>Explanation:</b>	Data warehouse database servers
10. ___ is the specialized data warehouse database	
	a) Oracle b) DBZ c) Informix d) Redbricks
<b>Ans:</b>	<b>d</b>
<b>Explanation:</b>	Redbricks
11.---defines the structure of the data held in operational databases and used by operational applications	
	a) User-level metadata b) Data warehouse metadata c) Operational metadata d) Data mining metadata
<b>Ans</b>	<b>c</b>
<b>Explanation:</b>	Operational metadata



12.----holds the catalog of the warehouse database system	
	a) Application level metadata b) Algorithmic level metadata c) Departmental level metadata d) Core warehouse metadata
<b>Ans:</b>	<b>b</b>
<b>Explanation:</b>	Algorithmic level metadata
13. ____maps the core warehouse metadata to business concepts, familiar and useful to end-users	
	a) Application level metadata. b) User level metadata.C. c) Enduser level metadata. d) Core level metadata
<b>Ans:</b>	<b>a</b>
<b>Explanation:</b>	
14. The star schema is composed of _____ fact table.	
	a) One b) Two c) Three d) four
<b>Ans:</b>	<b>a</b>
<b>Explanation:</b>	Only one fact table
15. The source of all data warehouse data is the__	
	a) operational environment b) informal environment c) formal environment. d) technology environmen
<b>Ans:</b>	<b>a</b>
<b>Explanation:</b>	
16.The @active data warehouse architecture includes which of the following?	
	a) At least one data mart b) Data that can extracted from numerous internal and external sources c) Near real-time updates d) All of the above.
<b>Ans:</b>	<b>d</b>
<b>Explanation:</b>	
17.An operational system is which of the following?	
	a) A system that is used to run the business in real time and is based on historical data. b) A system that is used to run the business in real time and is based on current data. c) A system that is used to support decision making and is based on

	current data. d) A system that is used to support decision making and is based on historical data.
<b>Ans:</b>	<b>b</b>
<b>Explanation:</b>	
18.A data warehouse is which of the following?	
	a) Can be updated by end users. b) Contains numerous naming conventions and formats. c) Organized around important subject areas d) Contains only current data.
<b>Ans:</b>	<b>c</b>
<b>Explanation:</b>	Data warehouse is subject oriented
19. Good performance can be achieved in a data mart environment by extensive use of	
	a) Indexes b) creating profile records c) volumes of data d) all of the above
<b>Ans:</b>	<b>d</b>
<b>Explanation:</b>	
<b>20.</b>	Warehouse administrator responsible for
	a) Administrator b) Maintenance c) both a and b d) none of the above
<b>Ans</b>	<b>c</b>
<b>Explanation</b>	
<b>21. What is data cube?</b>	
	a) allows data to be modeled and viewed in multiple dimensions b) data with dimensions c) data values d) description about data
<b>Ans.</b>	<b>a</b>
<b>23 .Which of the following is not a multidimensional data model?</b>	
	a) Star schema b) Fact constellation c) Snowflake schemas d) Entity-relationship model
<b>Ans</b>	<b>d</b>
<b>Explanation</b>	Three models of data warehouse: star, snowflake and fact constellation
<b>24. Snowflake schema consists of ___ fact tables</b>	
	a) One b) Two c) Three

	d) four
<b>Ans</b>	a
<b>Explanation</b>	Having only one fact table and many dimension tables
<b>25.Fact constellation consists of __ fact tables</b>	
	a) one b) two c) three d) many
<b>Ans</b>	d
<b>Explanation</b>	Many fact tables and many dimension tables

Name and Sign of Subject Teacher