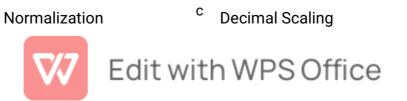
MULTIPLE CHOICE

1.	What is KDD in data mining?		
	^a Knowledge Discovery Database	C	Knowledge Discovery Data
	b Knowledge Data definition	d	Knowledge data house
	ANS: A PTS: 1		
2.	A is a repository of information	col	lected from multiple squares stored
	under a unified schema, and which usu	ally	resides at a single site.
	a Data mining	C	Data warehouse
	b Database	d	legacy databases
	ANS: C PTS: 1		
3.	Which of the following databases is us	ed t	to store time-related data?
	^a Spatial databases	С	Multimedia databases
	b Text databases	d	Temporal databases
	ANS: D PTS: 1		
4.	Which of the following is the most poprepo	ular	ly available and rich information
	sitories?		
	a Temporal databases	С	Transactional databases
	b Relational databases	d	spatial databases
	ANS: B PTS: 1		
5.	A data normalization technique for real nume	-val	lued attributes that divides each
	rical		



value by the same power of 10

Min-Max Normalization

	b Z-Score Normalization	d Decimal Smoothing				
	ANS: C PTS: 1					
6.	A data warehouse is which of the fo					
	Contains numerous naming cor C Organized around important sul C Contains only current data					
7.	ANS: C PTS: 1 The data Warehouse is					
	a read only b write only read and write only none					
	ANS: A PTS: 1					
8.	Data about data is called a Table b Database ANS: C PTS: 1	^C Metadata d Integration				
9.	is the heart of the warehouse.					
	a Data mining database servers.	Data warehouse databaseservers.				
	b Data mart database servers . ANS: C PTS: 1	d Relational data base servers				
10.	Data Mining is also referred to as					
	a Knowledge discovery in	C Data Extraction				



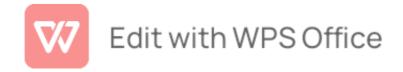
	Data Cleaning	d	Data Management
	ANS: A PTS: 1		
11.	A technique to reduce the reduct	nda	ncies in data representation in order
	o decrease data storage requirements a Data Removal b Data Mining ANS: D PTS: 1	c d	Data Cleaning Data Compression
12.	A goal of data mining includes which of a To explain some observed event or b To confirm that data exists To analyze data for expected relation To create a new data warehouse ANS: C PTS: 1	cor	ndition
13.	Data Warehouse deals with a Historical data b Live data ANS: A PTS: 1	c d	Future Data All of the above
14.	KDD has been described as the application of the Waterfall Model. b Object-Oriented Programming. ANS: C PTS: 1	etior c d	n of to Data Mining The Scientific Method Procedural Model
15.	Data transformation includes which of a A process to change data from a su b A process to change data from a de	ımn	nary level to a detailed level



	C Joining data from one source into various sources of data						
	d Separating data from one source into various sources of data						
	ANS: B PTS: 1						
16.	Data mining is best described as the process of						
	identifying patterns in data representing data						
	deducing relationships in data simulating trends in data						
	ANS: A PTS: 1						
17.	Reconciled data is which of the following?						
	 Data stored in the various operational systems throughout the organization 						
	 Current data intended to be the single source for all decision support systems 						
	C Data stored in one operational system						
	 Data that has been selected and formatted for end-user support applications 						
	ANS: B PTS: 1						
18.	The operation performs a selection on one dimension of the given cube, resul						
	ting in a sub cube						
	^a Pivot ^c Roll-up						
	b Slice d Drilldown						
	ANS: B PTS: 1						
19.	OLTP systems are used for						
	a Access to operational and historical data						
	b Providing various add-ons for quick decision making						
	C Supporting day-to-day business process						
	d Providing complicated computations						



	ANS: C PTS: 1						
20.	OLAP stands for						
	a On-Line Analytical Processing	С	On-Line Access Processing				
	b Operational Analytical Processing	d	On-Line Analytical Programming				
	ANS: A PTS: 1						
21.	This data transformation technique wo values fo	rks	well when minimum and maximum				
	r						
	a real-valued attribute are known						
	a Min-Max Normalization	C	Z-Score Normalization				
	b Decimal Scaling	d	Logarithmic Normalization				
	ANS: A PTS: 1						
22.	This process removes redundancies th	at r	may be present in a data model				
	a Abstraction	C	Standardization				
	b Granularization	d	Normalization				
	ANS: D PTS: 1						
23.	The set of values allowed to be entered	d int	o a column is known as its				
	a Data type	C	Domain				
	b Constraints	d	None of the above				
	ANS: C PTS: 1						
24.	The relational database model is desig	ned	Ito				
	^a Promote Data Redundancy						
	b Minimize Data Redundancy						
	c Eliminate the Need for Data Transfe	orm	ations				
	d Eliminate the Need for Data Preprocessing						
	ANS: B PTS: 1						

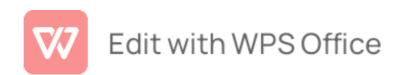


25.	Which of the following process includes data cleaning, data integration, data se					
	lection, data transformation, data mining, and pattern evolution and knowledge presentati					
	on? a KDD process b ETL process ANS: A PTS: 1					
26.	Which of the following Variable measured or their values defined, using one of t					
27.	kinds of Nonmetric scales a Numerical variable					
28.	This step of the KDD process model deals with noisy data a Creating a Target Dataset b Data Preprocessing d Data Mining ANS: B PTS: 1					
29	The derived model may be represented in the form of					

29. The derived model may be represented in the form of



	a ER model		Decision trees			
	b Flow chart	d	DFD			
	ANS: C PTS: 1					
30.	Which of the following is the classification	ation	of data mining systems?			
	a Summarization	C	Discrimination			
	b Visualization	d	Characterization			
	ANS: B PTS: 1					
31.	What are the functions of Data Mining	j??				
	Association and correctional analysis classification	C	Cluster analysis and Evolution analysis			
	b Prediction and characterization	d	All of the above			
	ANS: D PTS: 1					
32.	Which of the following refers to the st	eps (of the knowledge discovery process,			
	n which the several data sources are	coml	pined??			
	a Data selection	C	Data cleaning			
	b Data transformation	d	Data integration			
	ANS: D PTS: 1					
33.	The data is stored, retrieved & updated in					
	a OLAP					
	b OLTP					
	c SMTP					
	d FTP					
	ANS: B PTS: 1					
34.	Expansion for DSS in Data Warahouse	e is	?			
	a Decision Support system.					



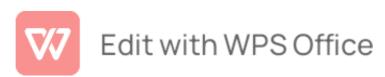
	b Decision Single System		
	C Data Storable System		
	d Data Support System		
	ANS: A PTS: 1		
35.	predicts future trends & to	beha	aviors, allowing business managers
	make proactive,knowledge-driven de	cisior	ns.
	a Data warehouse	С	Data mining
	b Datamarts	d	Metadata
	ANS: C PTS: 1		
36.	Fact tables are?		
	^a completely demoralized	С	partially demoralized
	b completely normalized.	d	partially normalized
	ANS: B PTS: 1		
37.	The star schema is composed of		fact table
	a one	C	three
	b two	d	four
	ANS: A PTS: 1		
38.	An system focuses mainly on the department, without referring to history		•
	ons		
	a On-Line Analytical Processing	C	On-Line Electronic Processing
	b On-Line Data Processing	d	On-Line Transaction Processing
	ANS: D PTS: 1		
39.	A DWH is a subject oriented, integrat collection of data	ed, tii	me- variant, and
	in support of management 's de	cisio	n-making process



ANS: 40. Data a da b da ANS: 41. Choo	can be upda ata warehous ata mining C	se PTS:	environme	c d	Object- oriented operational informational
40. Data a da b da . ANS: 41. Choo	can be update at a warehous at a mining C se the easures	ted in . se PTS:	environme	c d	informational
a da b da ANS: 41. Choo a M	ata warehous ata mining C se the	se PTS:	1	c d	informational
ANS: 41. Choo	ata mining C se the easures	PTS:		d	informational
ANS: 41. Choo a M	C se the easures				
41. Choo ^a M	se the easures			ead	ala f a a t t a la la mara a mel
a M	easures	th	at will populate	eac	-h foot to hio was a sad
b					en ract table record
b г:	mensions			C	Grain
ال .				d	Business Process
ANS:	Α	PTS:	1		
42. Meta	data reposit	ory co	ntains		
a O _l	perational m	eta da	ta		
b Da	ata irrelevant	t to sys	stem performar	nce	
c Th	ne mapping f	from th	ne DWH to the o	per	rational environment
d Su	ummarized d	lata			
ANS:	Α	PTS:	1		
43. A	contains	s a sub	set of corporat	:e-w	vide data that is of value to a specifi
С					
	of users.				
a Er	nterprise war	ehous	e	С	Data warehouse
b Vi	rtual wareho	use		d	Data mart
ANS:	D	PTS:	1		
44	_approach ru	ıles in	rules are desce	ndi	ng order of their priority
a Oı	rdered rules			c	Rule set
b Uı	nordered rule	es		d	None



	•		•	
	ANS: A	PTS: 1		
45.	are greedy	algorithms for the ind	lucti	on of decision trees
	a ID3		C	Both A and B
	b C4.5		d	Backpropagation
	ANS: C	PTS: 1		
46.	is a g	ood alternative to the	sta	r schema.
	^a Star schema		С	Star-snowflake schema
	b Snowflake sc	hema	d	Fact constellation
	ANS: D	PTS: 1		
47.	is a supe	rvised learning		
	^a Cluster		С	Classification
	b Association		d	prediction
	ANS: C	PTS: 1		
48.	the left-h	nand side of the rule is	s ca	lled
	a Rule consequ	ent	C	Rule set
	b Rule antecede	ent	d	None
	ANS: B	PTS: 1		
49.	the right	-hand side of the rule	is c	alled
	a Rule consequ	ent	C	Rule set
	b Rule antecede	ent	d	None
	ANS: A	PTS: 1		
50.	is a unsu	upervised learning alg	oritl	nm
	a Cluster		С	Classification
	b Association		d	prediction
	ANS: A	PTS: 1		



51.	The type of relation	onship in star schema	ais.	
	^a One-to-One		С	Many-to-One
	b One-to-Many		d	Many-to-Many
	ANS: A	PTS: 1		
52.	Association rule	Measured by		
	a Support only			
	b Confidence or	nly		
	c Support,Confi	dence only		
	d Support,Confi	dence and lift only		
	ANS: D	PTS: 1		
53.	In Hunt's algorith	m , a decision tree is	grov	vn in fashion
	^a Horizontal		С	Recursive
	b Vertical		d	None
	ANS: C	PTS: 1		
54.	Nearest neighbor information	classifier makes thei	ir pr	ediction based on
	^a Local		С	Temporary
	b Global		d	None
	ANS: A	PTS: 1		
55.	are resp	onsible for running qu	ierie	es and reports against data
	tables.			
	^a Hardware.		C	End users
	b Software		d	Middle ware
	ANS: C	PTS: 1		

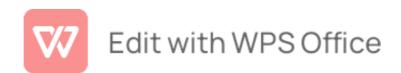
56. Prediction models



	a Classes	C Attributes					
	b Continuous-valued functions	d None					
	ANS: B PTS: 1	·					
57.	Classification is						
	a A subdivision of a set of examples	s into a number of classes					
	b A measure of the accuracy, of the	e classification of a concept that is giv	⁄en				
	c by a certain theory						
	d The task of assigning a classificat	tion to a set of examples					
	ANS: A PTS: 1						
58.	classification model can also be used to predict the class label of unk						
	nown records						
	a Descriptive modeling	C Regression					
	b Predictive modelling	d None					
	ANS: B PTS: 1						
59.	classification model can se	erve as an explanatory toll to distingui	ish				
	etween objects of different classes						
	a Descriptive modeling	^C Regression					
	b Predictive modelling	d None					
	ANS: A PTS: 1	·					
60.	In approach the rules are in de	escending order of their priority					
	a Ordered rules	^C Rule set					
	b Unordered rules	d None					
	ANS: A PTS: 1						
61.	Query tool is meant for						



	a data acquisition	C	information exchange						
	b information delivery	d	communication						
	ANS: A PTS: 1								
62.	Classification belongs to which type of learning?								
	^a supervised	С	rote						
	b unsupervised	d	machine						
	ANS: A PTS: 1								
63.	is a flow-chart like tree struct	ture							
	a Bayesion classifier	C	Backpropagaton						
	b Decision tree	d	None						
	ANS: B PTS: 1								
64.	is a eager learner classifier								
	^a Bayesion	С	Backpropagaton						
	b Decision tree	d	Nearest neighbor						
	ANS: B PTS: 1								
65.	Accuracy can be defined as								
	a Number of Correct predictions/To	tal n	umber of predeictions						
	b Number of wrong predictions/Total number of predeictions								
	C Total number of predeictions/Total records								
	d None								
	ANS: A PTS: 1								
66.	Classification predicts								
	^a Classes	С	Attributes						
	b Continuous-valued functions	d	None						
	ANS: A PTS: 1								



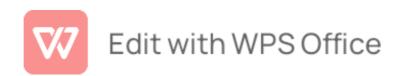
67.	not predicts the categorical labels							
	a Cluster b Association		c d	Classification Prediction				
	ANS: A	PTS: 1	٠					
58.	node ha	as no incomming edg	es					
	a Internal node		C	root node				
	b Leaf node		d	terminal node				
	ANS: C	PTS: 1						
69.	predicts	s the categorical labe	ls					
	a Cluster		C	Classification				
	b Association		d	Prediction				
	ANS: C	PTS: 1						
70.	is a	unsupervised learning	g alg	jorithm				
	a Cluster		C	Classification				
	b Association		d	Prediction				
	ANS: A	PTS: 1						
71.	Data classification	on is proce	ss					
	a One-step		C	Three-step				
	b Two-step		d	None				
	ANS: B	PTS: 1						
72.	Classification rule	es are extracted from	l					
	a root node.							
	b siblings							
	c decision tree							
	d branches							
	ANS: C	PTS: 1						



73.	Which of the following is constructed where the enterprise warehouse is the sole custodian of all warehouse datA. Which is then distributed to the various depend	<u>)</u>
	ent data marts. a Enterprise DWH c Multi-tier DWH	
	b Two- tier DWH d Virtual warehouse	
	ANS: C PTS: 1	
74.	Dimensionality reduction reduces the data set size by removing	
	a relevant attributes C derived attributes.	
	b irrelevant attributes d composite attributes.	
	ANS: B PTS: 1	
75.	A concept hierarchy that is a total or partial order among attributes in databa	as
	e schema is called a hierarchy. a Set-grouping	
	ANS: D PTS: 1	
76.	Classification accuracy is A subdivision of a set of examples into a number of classes No of examples correctly classified/total no of Examples The task of assigning a classification to a set of examples None of these ANS: B PTS: 1	
77.	How many components are there in a data warehouse? a two c four .	



	b three			d	five
	ANS: D	PTS:	1		
78.	The dimension table a entities	es des	cribe the	_ C	facts.
	b keys.			d	units of measures.
	ANS: C	PTS:	1		
79.	In Fact Table unique a primary key	e-id is	refered as?	С	foreign key
	b candidate key			d	None
	ANS: C	PTS:	1		
80.	In Dimesion table ur a primary key	nique-i	d refered as?	С	foriegn key
	b candidate key			d	none
	ANS: A	PTS:	1		
81.	Which of the following Exploratory data	•			ta mining? Deductive learning.
	b Data driven disc	overy		d	All of the above.
	ANS: D	PTS:	1		
82.	Converting data from called as	m diffe	erent sources into	оас	common format for processing is
	a selection.			С	interpretation
	b preprocessing.			d	transformation.
	ANS: D	PTS:	1		
83.	Various visualizationa selection.	n tech	niques are used i		step of KDD data mining
	b transformaion.			d	interpretation.
	ANS: D	PTS:	1		
84.	The proportion of tra	ansac	tion supporting X	in ⁻	



		b :	support.			d d	All of the above
		ANS	S: B	PTS:	1	·	
;	85.		absolute numb confidence	er of tr	ansactions supp	ortii c	ng X in T is called support count.
		b	support			d	None of the above
		ANS	S: C	PTS:	1		
1	86.	tran		i jam,10	0000 transaction		nsaction contain bread, 30000 Itain both bread and jam. Then the
		a		ia jairi		С	20%
		b :	3%			d	30%
		ANS	S: A	PTS:	1		
;	87.	If T consist of 500000 transactions, 20000 transaction contain bread, 30000 transaction contain jam,10000 transaction contain both bread and jam. Then the confidence of buying bread with jam is					
		a :	33.33%		·	С	45%
		b	66.66%			d	50%
		ANS	S: D	PTS:	1		
:	88.		ch of the follow Apriori	ing is r	not a frequent pa	tteri c	n mining algorithm? Decission Tree
		b	Eclat			d	FP-Growth
		ANS	S: C	PTS:	1		
:	89.		ch algorithm red FP-Growth	quires	fewer scans of d	ata? c	Apriori
		b	Eclat			d	Decission Tree
		ANS	S: A	PTS:	1		
•	90.	1. I1 2. I7 3. I1 4. I1	the question giv , 2, 3, 4, 5, 6 7, 2, 3, 4, 5, 6 , 8, 4, 5 , 9, 10, 4, 6		ow consider the	data	a Transactions :



	a <11>, <12>, <14>, <15>, <16>, <11, 14>, . <12, 14>, <12, 15>, <14, 15>, <14, 16>,	С	< 111>, < 4>, < 5>, < 6>, < 1, 4>, < 5
	< 2, 4, 5> b < 2>, < 4>, < 5>, < 2, 4>, < 2, 5>, < 4, . 5>, < 2, 4, 5>	d	< 2, 4, 5> < 1>, < 4>, < 5>, < 6>
	ANS: A PTS: 1		
91.	What techniques can be used to improve to a Hash-based techniques	he e c	
	b Transaction Increases	d	prediction
	ANS: A PTS: 1		
92.	The first phase of Apriori algorithm is a Candidate generation	 C	Pruning
	b Itemset generation	d	Partitioning
	ANS: A PTS: 1	•	
93.	The step eliminates the extensions	of	(k-1)-itemsets which are not found
	to be frequent, from being considered for counti a Candidate Generation	ng s c	support Pruning
	b Itemset Generation	d	Prediction
	ANS: C PTS: 1		
94.	The generalization of cross-tab which is re	pres	sented visually is wh
	ich is also called as data cube. a Two dimensional cube	С	N-dimensional cube
	b Multidimensional cube	d	Cuboid
	ANS: A PTS: 1	•	
95.	The operation of moving from finer-granula	arity	data to a coarser granularity (by
	means of aggregation) is called aa	- С	Dicing
	b Drill down	d	Pivoting
		_	



ANS: A PTS: 1

96. _____ uses array-based multidimensional storage engines for multidimensional views of data.

a KOLAP c ZOLAP

b MOLAP d ROLAP

ANS: B PTS: 1

97. The pivot operation is also known as?
a Aggregation c dice
b Rotation d none

.

ANS: B PTS: 1

