Department of Information Technology

Internal Assessment-I Second Half Y. 2020-21

Semester: VI **Class / Branch: TE- IT**
Subject: Data Mining and Business Intelligence (DMBI) **Date of Exam: 26/03/2021**

SN	Questions	Course Outcome	Bloom Taxonomy Level
1	**----------may be defined as the data objects that do not comply with the general behavior or model of the data available.** a) Outlier Analysis b) Evolution Analysis c) Prediction d) Classification	CO1	L1
2	**Which statement is true about prediction problems**? a) The output attribute must be categorical. b) The output attribute must be numeric. c)The resultant model is designed to determine future outcomes. d)The resultant model is designed to classify current behavior.	CO1	L3
3	**Classification problems are distinguished from estimation problems in that** a)classification problems require the output attribute to be numeric. b)classification problems require the output attribute to be categorical. c)classification problems do not allow an output attribute. d)classification problems are designed to predict future outcome	CO1	L2
4	**Which data maps into predefined groups**. a. Regression. b. Time series analysis c. Prediction. d. Classification.	CO1	L3
5	**Prediction can be viewed as forecasting a_____ value.** a. non-continuous. b. constant. c. continuous. d. variable.	CO1	L2
6	**Classification rules are extracted from ___** a. root node. b. decision tree.	CO1	L2

	c. siblings. d. branches.		
7	**Which of the following is a predictive model?** a) Clustering. b) Regression. c) Summarization. d) Association rules.	CO2	L3
8	**Bayesian classifiers is** a. A class of learning algorithm that tries to find an optimum classification of a set of examples using the probabilistic theory. b. Any mechanism employed by a learning system to constrain the search space of a hypothesis. c. An approach to the design of learning algorithms that is inspired by the fact that when people encounter new situations, they often explain them by reference to familiar experiences, adapting the explanations to fit the new situation. d) A measure of the accuracy, of the classification of a concept that is given by a certain theory	CO2	L1
9	**What is noise?** a. component of a network b. context of KDD and data mining c. aspects of a data warehouse d. None of these	CO2	L3
10	**_____ may be defined as the data objects that do not comply with the general behavior or model of the data available.** a. Outlier Analysis b. Evolution Analysis c. Prediction d. Classification	CO2	L2
11	**Data Visualization in mining cannot be done using Select one** a. Photos b. Graphs Charts c. Information d. Graphics	CO2	L1
12	**Dimensionality reduction reduces the data set size by removing _____ Select one:** a. composite attributes b. derived attributes c. relevant attributes d. irrelevant attributes	CO2	L1
13	**To detect fraudulent usage of credit cards, the following**	CO2	L2

	data mining task should be used Select one: a. Outlier analysis b. prediction c. association analysis d. feature selection		
14	**Nominal and ordinal attributes can be collectively referred to as_____ attributes Select one** a. perfect b. qualitative c. consistent d. optimized	CO2	L2
15	**Which of the following statement is correct** a. The standard deviation of constant is equal to unity b. The sum of absolute deviations is minimum if these deviations are taken from mean c. The second moment about origin equals variance d. The variance is positive quantity and is expressed in the square of units of the observation	CO3	L3
16	**In Binning, we first sort data and partition into (equal-frequency) bins and then which of the following is not a valid step Select one** a) Smoothing by Bin Boundaries b) Smoothing by Bin Medians c) Smoothing by Bin Means d) Smoothing by Bin Values	CO3	L1
17	**What is the use of data cleaning?** a. to remove the noisy data b. correct the inconsistencies in data c. transformations to correct the wrong data. d. Smoothing by Bin Values	CO3	L3
18	**Knowledge discovery in database refers to _____.** a. whole process of extraction of knowledge from data. b. selection of data. c. coding. d. cleaning the data.	CO3	L2
19	**Two fundamental goals of Data Mining are ____& ____.** a. Analysis and Description b. Data cleaning and organizing the data c. Prediction and Description d. Data cleaning and organizing the data	CO3	L2
20	**Spatial databases are used to store----------** a. Geographical information like map etc. b. Images, audio, video etc. c. Transactional data d. binary format files	CO3	L2