

Total No. of Questions: 6

Total No. of Printed Pages:2

Enrollment No.....



Faculty of Engineering
End Sem Examination May-2024
CB3EL08 Data Mining & Analytics

Programme: B.Tech.

Branch/Specialisation: CSBS

Duration: 3 Hrs.

Maximum Marks: 60

Note: All questions are compulsory. Internal choices, if any, are indicated. Answers of Q.1 (MCQs) should be written in full instead of only a, b, c or d. Assume suitable data if necessary. Notations and symbols have their usual meaning.

- Q i. Capability of data mining is to build _____ predictive models. 1
(a) Imperative (b) Predictive
(c) Interrogative (d) Retrospective
- ii. Strategic value of data mining is _____. 1
(a) Cost-sensitive (b) Work-sensitive
(c) Time-sensitive (d) Technical-sensitive
- iii. To remove noise and inconsistent data _____ is needed. 1
(a) Data cleaning (b) Data reduction
(c) Data transformation (d) Data integration
- iv. Data can be visualized using- 1
(a) Graphs (b) Charts (c) Maps (d) All of these
- v. What is association rule mining? 1
(a) Same as frequent itemset mining
(b) Finding of strong association rules using frequent itemsets
(c) Using association to analyze correlation rules
(d) Finding Itemsets for future trends
- vi. Which of the following is not used in supervised learning algorithms? 1
(a) Training data (b) Validation data
(c) Knowledge data (d) Test data
- vii. Which of the following methods do we use to best fit the data in logistic regression? 1
(a) Least square error (b) Maximum likelihood
(c) Jaccard distance (d) None of these

[2]

- viii. The Newton Raphson method is also called as _____. 1
(a) Tangent method (b) Secant method
(c) Chord method (d) Diameter method
- ix. Which of the following cross validation technique is better suited for time series data? 1
(a) K-fold cross validation
(b) Leave-one-out cross validation
(c) Stratified shuffle split cross validation
(d) Forward chaining cross validation
- x. Which of the following is not an example of a time series model? 1
(a) Naïve approach (b) Exponential smoothing
(c) Moving average (d) None of these
- Q.2 i. What is data mining? 2
ii. What are the stages of data mining process? 3
iii. Write a note on Online Analytical Processing (OLAP). 5
OR iv. Explain any two data mining applications. 5
- Q.3 i. What is data transformation? 3
ii. Write a note on data pre-processing. 7
OR iii. Explain the visualization techniques in detail. 7
- Q.4 i. What is association rule mining? Write steps for association rule mining generation. 4
ii. Explain any one decision tree algorithm with example. 6
OR iii. Explain the Bayesian classification with example. 6
- Q.5 i. Brief the linear regression. 4
ii. Explain the logistic regression with example. 6
OR iii. Write a note on semiparametric and nonparametric regression models. 6
- Q.6 Attempt any two: 5
i. Explain exploratory time series analysis. 5
ii. Write a note on autoregressive moving average. 5
iii. Write a note on prescriptive analysis. 5

Marking Scheme

Data Mining & Analytics(T)- CB3EL03(T)

Q.1	i)	Capability of Data Mining is to build _____ predictive models		1
		b) predictive		
	ii)	Strategic value of Data Mining is _____		1
		c) time-sensitive		
	iii)	To remove noise and inconsistent data _____ is needed		1
		a) data Cleaning		
	iv)	Data can be visualized using?		1
		d) All of these		
	v)	What is association rule mining?		1
		b) Finding of strong association rules using frequent itemsets		
	vi)	Which of the following is not used in Supervised Learning algorithms?		1
		c) Knowledge data		
	vii)	Which of the following methods do we use to best fit the data in Logistic Regression?		1
		b) Maximum Likelihood		
	viii)	The Newton Raphson method is also called as _____		1
		a)Tangent Method		
	ix)	Which of the following cross validation technique is better suited for time series data?		1
		d) Forward Chaining Cross Validation		
	x)	Which of the following is not an example of a time series model?		1
		d) None of these		
Q.2	i.	What is Data Mining?		
		Definition of Data Mining	1 Mark	2
		Example of Data Mining	1 Mark	
	ii.	What are the stages of Data Mining Process?		3
		Stages (Each stage carry equal mark)	3 Marks	
	iii.	Write a note on Online Analytical Processing (OLAP)		
		What is OLAP	1 Marks	
		Example of OLAP	1 Marks	5
		Type of Operations	3 Marks	
OR	iv.	Explain any two Data Mining applications		5
		Any two applications with necessary briefing	5 Marks	
Q.3	i.	What is Data Transformation?		2
		Data Transformation	1Marks	

	ii.	Example	1 Marks	
		Write a note on Data pre-processing		
		What is pre-processing	2 Marks	8
OR	iii.	Briefing about different types of preprocessing	6 Marks	
		Explain the visualization Techniques in detail		8
Q.4	i.	What is Association Rule Mining? Steps in Association Rule Mining generation		
		Association Rule definition	1 Marks	3
		Example	1 Marks	
	ii.	Steps find association rule	1 Marks	
		Explain any one Decision Tree algorithm with example		7
OR	iii.	Decision Tree	2 Marks	
		Explanation with example	5 Marks	
		Explain the Bayesian Classification with example		7
	iii.	Bayesian Classification	2 Marks	
		Explanation with example	5 Marks	
Q.5	i.	Brief the Linear Regression		
		Linear Regression	1 Marks	4
		How linear regression works	3 Marks	
	ii.	Explain the Logistic regression with example		6
		Logistic Regression	1 Marks	
OR	iii.	Explanation with example	5 Marks	
		Write a note on semiparametric and nonparametric regression models.		6
		Explanation about semiparametric regression model	3 Marks	
	iii.	Explanation about nonparametric regression model	3 Marks	
Q.6	Attempt any two:			
	i.	Explain Exploratory Time Series Analysis		5
		What is Time Series Analysis	1 Marks	
	ii.	Explanation	4 Marks	
		Write a note on Autoregressive Moving Average		5
	iii.	Autoregressive Moving Average	1 Marks	
		Explanation	4 Marks	
	iii.	Write a note on Prescriptive Analysis		5
		Prescriptive Analysis	1 Mark	
	iii.	Explanation	4 Marks	
