Register Number:……..

KUMARAGURU
college of technology
character is life

B.E DEGREE EXAMINATIONS: APRIL / MAY 2023

(Regulation 2018)

Sixth Semester

COMPUTER SCIENCE AND ENGINEERING

U18CSI6203**:** Data Warehousing and Data Mining

COURSE OUTCOMES

CO1: Demonstrate data warehouse schema and process of data retrieval for real time applications.

CO2: Identify necessity of data pre-processing and apply the appropriate procedure.

CO3: Design and deploy appropriate Classification/ Clustering techniques for various problems with high dimensional data using modern tools.

CO4: Apply the association rules for real life mining applications.

CO5: Synthesize various mining techniques and work in teams to develop project on complex data objects.

Time: Three Hours **Maximum Marks: 100**

Answer all the Questions:-

PART A (10 x 2 = 20 Marks)
(Answer not more than 40 words)

1.	Compare discrete and continuous attributes in data objects.	CO1	[K1]
2.	List the various process carried out in data preprocessing.	CO1	[K_2]
3.	How is a data warehouse different from a database? How are they similar?	CO2	[K_2]
4.	Mention the approaches for building a data warehouse?	CO2	[K_1]
5.	Define rule-based classification. Give Example	CO3	[K_2]
6.	How are association rules mined from large databases?.	CO3	[K_2]
7.	Classify hierarchical clustering methods.	CO4	[K_2]
8.	Define outliers. List various outlier detection approaches.	CO4	[K_1]
9.	What is spatial mining in web mining?	CO5	[K_1]
10.	List the most common application of data mining in real life.	CO5	[K_2]

Answer any FIVE Questions:-
PART B (5 x 4 = 20 Marks)
(Answer not more than 80 words)

11.	Explain major requirements and challenges in data mining	CO1	[K_3]
12.	Mention the role of meta data in a data warehouse.	CO2	[K_3]
13.	Illustrates the fundamental differences of OLTP and OLAP in a practical comparison.	CO2	[K_4]
14.	Explain Naïve Bayesian classification with an example.	CO3	[K_4]
15.	Explain density based clustering methods.	CO4	[K_4]
16.	Explain the various types of web mining.	CO5	[K_4]

Answer any FIVE Questions:-
PART C (5 x 12 = 60 Marks)
(Answer not more than 300 words)

17.		Explain the steps involved in the data mining process.	12	CO1	[K2]
18.		Explain the multi-tier architecture suitable for evolving a data warehouse with a suitable diagram.	12	CO2	[K3]
19.	a)	What is classification? With an example explain how a support vector machine can be used for the classification.	6	CO3	[K5]
	b)	Discuss the classification by decision tree induction.	6	CO3	[K3]
20.		Explain the Hierarchical method clustering of classification with examples.	12	CO4	[K4]
21.	a)	Explain graph mining and time series mining.	6	CO5	[K4]
	b)	Explain in detail about outlier analysis.	6	CO5	[K4]
22.	a)	Discuss any two case studies in data mining applications.	8	CO4	[K5]
	b)	Explain in detail about grid based methods.	4	CO4	[K3]

****