

Reg. No:

--	--	--	--	--	--	--	--	--

SIDDHARTH INSTITUTE OF ENGINEERING & TECHNOLOGY:: PUTTUR
(AUTONOMOUS)

B.Tech IV Year I Semester Supplementary Examinations July-2022

DATA WAREHOUSING AND DATA MINING

(Computer Science Engineering)

Time: 3 hours

Max. Marks: 60

(Answer all Five Units **5 x 12 = 60** Marks)

UNIT-I

- 1 a What is KDD? Explain about data mining as a step in the process of knowledge discovery. **6M**
- b How to classify data mining systems? Explain. **6M**

OR

- 2 a Explain Data Integration. **6M**
- b Data Transformation methods. **6M**

UNIT-II

- 3 a Define Data warehouse? Explain Design principles. **8M**
- b Write in brief about schemas in multidimensional data model. **4M**

OR

- 4 a Briefly explain about the types of OLAP servers. **7M**
- b Explain about Indexing of OLAP data. **5M**

UNIT-III

- 5 a Explain about basic concepts of frequent item set mining. **6M**
- b Write the Apriori Algorithm. **6M**

OR

- 6 a Write about basic concept in Association Rule Mining. **5M**
- b Can we overcome the draw backs of Apriori algorithm? Explain. **7M**

UNIT-IV

- 7 Describe the data classification process with a neat diagram. How does the Naive classification works? Explain. **12M**

OR

- 8 a What is Bayesian belief network? Explain in detail. **6M**
- b Write a note attribute selection measures. **6M**

UNIT-V

- 9 a Differentiate between AGNES and DIANA algorithms. **6M**
- b How to access the cluster quality? **6M**

OR

- 10 a What is the drawback of k-means algorithm? How can we modify the algorithm to diminish that problem? **6M**
- b Give a brief note on PAM Algorithm. **6M**

*** END ***