

**Reg. No:**

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

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

**B.Tech III Year I Semester Supplementary Examinations November-2020**

**DATA WAREHOUSING AND DATA MINING**

(Computer Science & Information Technology)

Time: 3 hours

Max. Marks: 60

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

**UNIT-I**

- 1 a Why do we preprocess the data? Discuss? 6M  
b Write in brief about Data cleaning. 6M

**OR**

- 2 a Describe about Data discretization. 6M  
b Write about Dimensionality reduction methods. 6M

**UNIT-II**

- 3 a Define Data warehouse? Discuss Design principles. 6M  
b Write in brief about schemas in multidimensional data model. 6M

**OR**

- 4 a Briefly explain the German super market EDEKA's Data warehouse. 6M  
b Write about case study of Data Warehousing in the Tamilnadu Government 6M

**UNIT-III**

- 5 a What are the advantages of FP-Growth algorithm? 6M  
b Discuss the applications of association analysis. 6M

**OR**

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

**UNIT-IV**

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

**OR**

- 8 Write and explain about Classification by Back propagation Algorithm. 12M

**UNIT-V**

- 9 a What is outlier detection? Explain distance based outlier detection. 6M  
b Write partitioning around mediods algorithm. 6M

**OR**

- 10 Discuss in detail about Data mining Applications. 12M

\*\*\* END \*\*\*