

Reg. No: 

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

SIDDHARTH INSTITUTE OF ENGINEERING &amp; TECHNOLOGY:: PUTTUR

(AUTONOMOUS)

B.Tech III Year I Semester Supplementary Examinations Feb-2021

DATAWARE HOUSING AND DATA MINING

(Computer science &amp; Information Technology)

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? Discuss. **6M**

**OR**

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

**UNIT-II**

- 3 a What are steps in designing the data warehouse? Explain **8M**  
b Compare OLTP and OLAP **4M**

**OR**

- 4 a Write in brief about schemas in multidimensional data model. **7M**  
b Discuss about star schema and Snow Flake schema **5M**

**UNIT-III**

- 5 a Discuss 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 Discuss the applications of association analysis. **7M**

**UNIT-IV**

- 7 a What is prediction? Explain about Linear regression method. **6M**  
b Discuss about Accuracy and Error measures. **6M**

**OR**

- 8 How does the Naïve Bayesian classification works? Explain in detail. **12M**

**UNIT-V**

- 9 a Classify various Clustering methods. **6M**  
b How to access the cluster quality? **6M**

**OR**

- 10 a Write K-means clustering algorithm. **6M**  
b What is the drawback of k-means algorithm? How can we modify the algorithm to diminish? That problem. **6M**

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