Hall Ticket No: ☐☐☐☐☐☐☐☐☐☐  Question Paper Code: 16MCA112

# MADANAPALLE INSTITUTE OF TECHNOLOGY & SCIENCE, MADANAPALLE
(UGC-AUTONOMOUS)
MCA(2Yrs) II Year I Semester (R16) Supplementary End Semester Examinations – October 2020
(Regulations: R16)

## DATA WAREHOUSING AND DATA MINING

Time: 3Hrs                                                                                                          Max Marks: 50

Attempt all the questions. All parts of the question must be answered in one place only.
**In Q.no 1 to 5 answer either Part-A or B only**

Q.1(A)  What is Data Mining? Explain the architecture of data mining.                    10M

**OR**

Q.1(B)  What is Data Pre-processing? Explain in detail about the following          10M
   i)  Data Cleaning
   ii)  Data integration

Q.2(A)  Explain about the architecture of data warehouse with neat diagram.        10M

**OR**

Q.2(B)  Explain about Star schema and Snowflake schema with the help of suitable diagrams.   10M

Q.3(A)  Explain Naïve Bayes classification method.                                                      10M

**OR**

Q.3(B)  Describe the criteria for classifier accuracy.                                                  10M

Q.4(A)  What is outlier? Explain various methods for detecting outliers.              10M

**OR**

Q.4(B)  Discuss the partitioning methods.                                                                   10M

Q.5(A)  Discuss about the similarity search in Multimedia Mining.                         10M

**OR**

Q.5(B)  Discuss about Web mining.                                                                              10M

*** END***

Page 1 of 1