**Programme Name & Branch: M.Tech. (Integrated) Computer Science and Engineering**

**Course Name & Code:** **CSI3010 & Data Warehousing and Data Mining**

**Class Number:** **VL2021220104124** **Slot: B1+TB1**

## Instructions:

1. All the problems must be related with anyone of real time scenario. Avoid involving in malpractice and Reward is zero mark for copycats.

2. Neatly presented handwritten solution Booklet is expected.

3. Each question can carry minimum of 2 pages.

4. Solutions must be uploaded on-time in VTOP.

### Digital Assignment – I

*1. Develop a dataset related to disease information and display all the stages of data pre-processing techniques, also contrast the advantages and limitations. Decide a final dataset which smoothly deal the mining.*

*2. List out the interesting applications (minimum 10) to find frequent item sets and justify it by your commanding results.*