GITAM (Deemed to be University)
[19ECS747]
M.Tech Degree Examination

CSE & DS
I Semester

DATA SCIENCE
(Effective from the admitted batch 2022–23)

Time: 3 Hours **Max.Marks: 60**

Instructions: All parts of the unit must be answered in one place only.
Figures in the right hand margin indicate marks allotted.

SECTION-A

1. **Answer all the questions:** **10×2=20M**

   a) Write the benefits of data science.
   b) Define the characteristics of big data.
   c) Differentiate structured and unstructured data.
   d) Write the importance of data cleaning and filtering.
   e) What is hypothesis testing?
   f) Consider the given data, 30, 45, 46, 62, 64, 72, 75, 85, 90, 97 find mean, median and mode.
   g) What is the process of encoding categorical data into numerical data called?
   h) What is data visualization?
   i) What is perceptron?
   j) What is convolution in CNN?

SECTION-B

Answer the following: **5x8=40M**

UNIT-I

2. Explain data science process in detail .

OR

3. Explain the following terminologies related to data science.
   a) Statistics   b) Machine Learning .

UNIT-II

4. Discuss about data storage and management.

OR

5. List the challenges with using data from multiple sources. Explain briefly.

UNIT-III

6. Explain any three distributions in statistics.

OR

7. Consider the following dataset, apply KNN and predict the class for (Height=170, Weight=58 and K=4).

Height (CM)	Weight (KG)	Class
167	51	UnderWeight
182	62	Normal
176	69	Normal
173	64	Normal
172	65	Normal
174	56	UnderWeight
169	58	Normal
173	57	Normal
170	55	Normal

UNIT-IV

8. Explain about visual coding in data visualization.

OR

9. Discuss about different types of data visualizations.

UNIT-V

10. Explain multi-layer perceptron in detail.

OR

11. Explain about auto encoders briefly.