

**R18**

Code No: 155EQ

**JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD**

**B. Tech III Year I Semester Examinations, January/February - 2023**

**DATA SCIENCE**

**(Computer Science and Engineering – Cyber Security)**

**Time: 3 Hours**

**Max. Marks: 75**

**Note:** i) Question paper consists of Part A, Part B.

ii) Part A is compulsory, which carries 25 marks. In Part A, Answer all questions.

iii) In Part B, Answer any one question from each unit. Each question carries 10 marks and may have a, b as sub questions.

**PART – A**

**(25 Marks)**

- 1.a) Define big data. [2]
- b) What do you mean by Data Analysis? [3]
- c) How to choose k value for kNN algorithm? [2]
- d) What are the basic tools of exploratory data analysis? [3]
- e) State Bayes law. [2]
- f) What is the purpose of Laplace smoothing? [3]
- g) What is the need of web scraping? [2]
- h) List the limitations of decision tree. [3]
- i) What are the trending tools for data visualization? [2]
- j) What is the difference between security and privacy? [3]

**PART – B**

**(50 Marks)**

- 2.a) Discuss the skills required by geeks to become successful data scientists.
- b) What kinds of data need to be handled by a strong data scientist? Give examples for each kind. [5+5]

**OR**

3. Probability distributions are the foundations of statistical models. Explain the following distributions with illustrations:
  - a) Normal distribution
  - b) Poisson distribution
  - c) Gamma distribution. [3+4+3]

- 4.a) How to quantify popularity of a web page? Explain your approach.
- b) Exemplify the stages in the data science process. [5+5]

**OR**

5. Describe various evaluation metrics used for classification and prediction algorithms. [10]

