# SIDDHARTH INSTITUTE OF ENGINEERING & TECHNOLOGY:: PUTTUR
## (AUTONOMOUS)
### B.Tech III Year I Semester Regular & Supplementary Examinations February-2024
### INTRODUCTION TO DATA SCIENCE
#### CSE(Artificial Intelligence & DataScience)

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

(Answer all Five Units **5 x 12 = 60** Marks)

### UNIT-I

| | | | | | |
|---|---|---|---|---|---|
| 1 | a | How will you handling missing data in data science. | CO1 | L2 | 6M |
| | b | Examine K-nearest 1eighbour techniques look at the k-nearest point to make a prediction. | CO1 | L3 | 6M |

**OR**

| | | | | | |
|---|---|---|---|---|---|
| 2 | a | Show the various components of model building. | CO1 | L2 | 6M |
| | b | What are the ways analyzed the data and built a well-performing model. | CO1 | L2 | 6M |

### UNIT-II

| | | | | | |
|---|---|---|---|---|---|
| 3 | a | Define hypothesis testing. | CO2 | L1 | 4M |
| | b | How will you mathematically define confidence. | CO2 | L2 | 8M |

**OR**

| | | | | | |
|---|---|---|---|---|---|
| 4 | a | List and discuss the four measures of significance of association rules. | CO2 | L3 | 6M |
| | b | Give the applications of association rules. | CO2 | L1 | 6M |

### UNIT-III

| | | | | | |
|---|---|---|---|---|---|
| 5 | | Discuss the following with respect to linear regression<br>i) Categorical Variables<br>ii) Confidence Intervals on the Parameters<br>iii) Confidence Interval on the Expected Outcome<br>iv) Prediction Interval on a Particular Outcome | CO3 | L2 | 12M |

**OR**

| | | | | | |
|---|---|---|---|---|---|
| 6 | a | State Bayes' Theorem. | CO4 | L1 | 4M |
| | b | Discuss Naïve Bayes classification method considering an example. | CO4 | L2 | 8M |

### UNIT-IV

| | | | | | |
|---|---|---|---|---|---|
| 7 | a | What is meant by k-means. | CO5 | L1 | 4M |
| | b | Describe k-means algorithm to find k clusters. | CO5 | L3 | 8M |

**OR**

| | | | | | |
|---|---|---|---|---|---|
| 8 | a | List and explain time series components. | CO6 | L1 | 6M |
| | b | Discriminate the steps involved in Box-Jenkins Methodology. | CO6 | L3 | 6M |

### UNIT-V

| | | | | | |
|---|---|---|---|---|---|
| 9 | a | Sketch the flow diagram of Text analysis process. | CO6 | L2 | 6M |
| | b | Illustrate in detail the steps involved in the process of text analysis done by organizations. | CO6 | L3 | 6M |

**OR**

| | | | | | |
|---|---|---|---|---|---|
| 10 | | Illustrate the main challenges of text analysis. | CO6 | L3 | 12M |

*** END ***

64