**U.S.N** [ ][ ][ ][ ][ ][ ][ ][ ][ ][ ]

# P.E.S. College of Engineering, Mandya - 571 401
*(An Autonomous Institution affiliated to VTU, Belagavi)*
**Second Semester, M. Tech - Computer Science and Engineering (MCSE)**
**Semester End Examination; October - 2021**
**Data Science**

Time: 3 hrs                                                                       Max. Marks: 100

*Note: I) Answer any FIVE full questions, selecting ONE full question from each unit.*
  *II) Any THREE units will have internal choice and remaining TWO unit questions are compulsory.*
  *III) Each unit carries 20 marks.*

| Q. No. | Questions | Marks | BLs | COs | POs |
|---|---|---|---|---|---|
| | **UNIT - I** | **20** | | | |
| 1 a. | With a neat diagram, explain Drew Conway's Venn diagram of data science. | 10 | L2 | CO1 | PO1,2 |
| b. | Discuss statistical and fitting a model in detail. | 10 | L2 | CO1 | PO1,2 |
| | **UNIT - II** | **20** | | | |
| 2 a. | With a neat diagram, explain data science process with data scientist's role in detail. | 10 | L2 | CO2 | PO1,2 |
| b. | Discuss K-mean algorithm in detail. | 10 | L3 | CO2 | PO2,3 |
| | **OR** | | | | |
| 2 d. | Discuss K-Nearest Neighbors algorithm in detail. | 10 | L3 | CO2 | PO2,3 |
| e. | Explain the significance of Exploratory Data Analysis (EDA) in detail. | 10 | L3 | CO2 | |
| | **UNIT - III** | **20** | | | |
| 3 a. | Illustrate Bayes Law with spam filter for individual words in detail. | 10 | L4 | CO3 | PO3 |
| b. | Describe, why linear regression doesn't work for filtering spam in detail? | 10 | L3 | CO3 | PO2,3 |
| | **UNIT - IV** | **20** | | | |
| 4 a. | Discuss singular value decomposition and principal component analysis in detail. | 10 | L3 | CO4 | PO2,3 |
| b. | Explain two aspects of wrappers in detail. | 10 | L2 | CO4 | PO2 |
| | **OR** | | | | |
| 4 d. | Discuss some problems with nearest neighbor's poses. | 10 | L2 | CO4 | PO2 |
| e. | Discuss decision tree algorithm in detail. | 10 | L3 | CO4 | PO2,3 |
| | **UNIT - V** | **20** | | | |
| 5 a. | Describe any two technique used for direct discovery of communities in a social network. | 10 | L3 | CO5 | PO2 |
| b. | Illustrate Girvan-Newman algorithm in detail. | 10 | L4 | CO5 | PO2,3 |
| | **OR** | | | | |
| 5 d. | Give essential characteristics of a social network. Also explain social network as graph with an example. | 10 | L3 | CO5 | PO2 |
| e. | Discuss two tools used for data visualization. | 10 | L3 | CO5 | PO2,3 |

∗ ∗ ∗ ∗