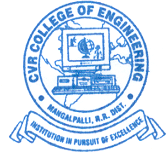
**CVR COLLEGE OF ENGINEERING**



***UGC Autonomous Institution -*** Affiliated to JNTUH

Vastunagar, Mangalpalli(V), Ibrahimpatan (M), R.R.Dist–501 510.

**B.Tech (CSIT) III Year I Semester Assignment-I Examinations 2023 Set-1**

**Subject Name: Concepts in data science**

**Time: 1 hour 45 minutes Max. Marks: 08**

**PART A**

1. Define data science? L5 CO1 2M
2. What is datafication. L2 CO1 2M
3. What is Population? L1 CO2 2M
4. Differentiate Feature Generation and Feature Selection. L1 CO2 2M
5. What is the adjected R2 coefficient and explain its use? L2 CO3 2M

**PART B**

1. Explain different data science applications? L3 CO1 10M
2. Explain the process of data science with a neat diagram? L2 CO1 10M
3. What is feature generation and feature selection? Explain with an example. L2 CO2 5M
4. What is feature selection? explain the role of Information Gain in feature selection?

L2 CO2 5M

1. Find the regression coefficients for the following data L4 CO3 10M

|  |  |  |
| --- | --- | --- |
| Hours worked | Pre exam marks | score |
| 1 | 15 | 80 |
| 2 | 28 | 95 |
| 1 | 14 | 76 |
| 1 | 13 | 70 |
| 2 | 22 | 81 |
| 2 | 28 | 97 |
| 3 | 30 | 98 |
| 3 | 29 | 95 |
| 1 | 14 | 74 |
| 2 | 25 | 85 |

1. What is logistic regression? Explain in detail. L2 CO3 (5+5)M

What is multiple Linear regression? Explain in detail.