



CVR COLLEGE OF ENGINEERING
An UGC Autonomous Institution - Affiliated to JNTUH
B.Tech. IV Year I Sem. II Mid Examinations, December - 2021
Subject: Fundamentals of Data Science (OE-1)

Date: 17-12-2021(FN)

Time: 1 hr 45 min

Max. Marks: **40 M**

PART – A

Answer **ALL** questions

5x2 = 10 M

1. What is Web Scraping? L1 (CO 3)
2. Define a data product. L2 (CO 4)
3. What is a recommendation engine? L3 (CO 4)
4. Write about feature selection related to predictive models. L4 (CO 5)
5. Explain customer retention using data science L2 (CO 5)

PART – B

Answer **ALL** questions

3x10= 30 M

6. There are four medicines and each medicine has two attributes X (Weight Index) and Y (PH value). The data on these medicines are given below. Classify the medicines into two groups using K-Means method. L3 (CO3)

Medicine	X	Y
A	1	1
B	2	1
C	4	3
D	5	4

(OR)

7. Playing cricket(Y) depends on outlook(X_1), temperature(X_2), humidity(X_3) and wind(X_4). The data on these variables are given in the following table: L3 (CO3)

Day	X_1	X_2	X_3	X_4	Y
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	Yes
4	Rain	Cool	Normal	Weak	Yes
5	Rain	Cool	Normal	Strong	No
6	Overcast	Cool	Normal	Strong	Yes
7	Sunny	Mild	High	Weak	No
8	Sunny	Cool	Normal	Weak	Yes
9	Rain	Mild	Normal	Weak	Yes
10	Overcast	Mild	High	Strong	Yes
11	Overcast	Hot	Normal	Weak	Yes
12	Rain	Mild	High	Strong	No

Find Y when X_1 =Sunny, X_2 =Cool, X_3 = High, X_4 = Strong using Naive Bayes method.

P.T.O

8. Perform the principal component analysis to the following data:

L6 (CO4)

X_1	3	10	4	8	5
X_2	10	8	9	8	5

(OR)

9. Find the Singular Value Decomposition of the following matrix:

L5 (CO4)

2	-1
2	1

10. Draw the decision tree to the data given in Question 7.

L4 (CO5)

(OR)

11. Discuss the wrapper methods for feature selection.

L2 (CO5)
