

Total No. of Printed Pages: 02

SUBJECT CODE NO:- H-4024
FACULTY OF SCIENCE AND TECHNOLOGY
Final B.Tech. (ETC) (Sem-VII)
Open Elective-II
Data Science
[Revised]

[Time: Three Hours]

[Max.Marks:80]

N.B

Please check whether you have got the right question paper.

1. Question No. 1 and 6 is compulsory.
2. Attempt any **two** others questions from the remaining in each section.
3. Figures to the right indicate full marks.
4. Assume suitable data if necessary.

Section A

- | | | |
|-----|---|----------|
| Q.1 | Solve any Five | 10 |
| | <ol style="list-style-type: none"> a) Enlist types of data b) Define big data c) List types of sampling d) Define population e) What is datafication f) What is statistical modeling g) What is SPAM | |
| Q.2 | <ol style="list-style-type: none"> a) List EDA tools and explain the data sciences process? b) To draw the histogram and boxplot graph.
(4, 8, 12, 18, 24, 30, 36, 42, 56, 50, 44, 38, 32, 28, 24, 20, 16, 14, 8, 4) | 08
07 |
| Q.3 | <ol style="list-style-type: none"> a) What is API? Explain its type in detail. b) Explain current landscape perspective of data sciences. | 07
08 |
| Q.4 | <ol style="list-style-type: none"> a) Define statistics? Explain probability distribution in detail. b) In Moses lake 7.9% of workers carpool to work. Conduct a sample of 41 workers and find. <ol style="list-style-type: none"> a) Distribution $X \approx B$ b) How many would you expect to carpool to work? c) What is standard deviation d) What is the probability for 3 of 41 carpool to work. | 07
08 |
| Q.5 | <ol style="list-style-type: none"> a) Explain KNN in detail? b) What is regression analysis? Write its applications? Define logistic regression? | 08
07 |

Section B

- Q.6 Solve any FIVE 10
- What is graph?
 - What is decision tree?
 - What is cluster analysis?
 - What is a wrapper?
 - What is dimensionality reeducation?
 - List neighborhood properties of graphs?
 - Define random forest

- Q.7 a) Explain the steps for K-Means and Apply K-means clustering for given data set 08

Customer	Age	Income	Number of credit card
Suresh	35	35K	3
Ramesh	22	50K	2
Ram	63	200K	1
Sham	59	170K	1
Anand	25	40K	4
John	37	50K	2

- b) Explain in brief feature generation algorithm. 07

- Q.8 a) What is PCA? Apply PCA to following data set and find the value of Eigen vectors. 15

X	Y
2.5	2.4
0.5	0.7
2.2	2.9
1.9	2.2
3.1	3.0
2.3	2.7
2	1.6
1	1.1
1.5	1.6
1.1	0.9

- Q.9 a) Describe different ethical issues of data sciences 07
- b) List supervised learning algorithm and explain any one in detail. 08

- Q.10 a) Explain the varieties of social networks 07
- b) What is spam? Explain different type of Spam. 08