

Roll No.:

SGT UNIVERSITY

END TERM THEORY EXAMINATION JULY-2022

| | | | |
|---------------------------|-----------------------------------|----------------|--------------------------|
| Faculty/College of Study: | Engineering & Technology | Year/Semester: | 6 th Semester |
| Program: | B.Tech. (CSE) MLAI | Duration: | 03:00 Hrs |
| Course/Subject: | Data Science tools and techniques | Maximum Marks: | 60 |
| Course/Subject Code: | 13020602 | Batch: | 2019 |

Instructions:-

1. Write Your Roll No. on the Question Paper.
2. Candidate should ensure that they have been provided correct question paper. Complaint(s) in this regard, if any should be made within 15 minutes of the commencement of the exam. No complaint(s) will be entertained thereafter.
3. All Questions are compulsory. Marks are indicated against each question.
4. Illustrate your answer with diagram wherever required.

SECTION-A

(Very Short Answer Type Questions)

Note: All Questions are compulsory: -

[12X1=12 Marks]

| S. No. | Question | Marks Allotted |
|--------|---|----------------|
| 1 | Which methodologies, tools, and techniques is used to explore the data in order to find patterns in the data. | 1 |
| 2 | What is Distributed Computing? | 1 |
| 3 | What is the objective of distributed computing? | 1 |
| 4 | How much data HDFS can store? | 1 |
| 5 | What happen when a data node fail? | 1 |
| 6 | Write a code to find all the number from given string :- abc123xyz' | 1 |
| 7 | What do you mean by finditer()? | 1 |
| 8 | While we predict what the next word in a sentence will be based on the ? | 1 |
| 9 | The smallest unit of a language is known as? | 1 |
| 10 | How we can check Hadoop version? | 1 |
| 11 | What do you mean One-Hot Encoding? | 1 |
| 12 | What is morphological segmentation? | 1 |

SECTION-B
(Short Answer Type Questions)

Note: All Questions are compulsory: -

[4X2=8 Marks]

| S. No. | Question | Marks Allotted |
|--------|--|----------------|
| 13 | Differentiate Name Node and data Node? | 2 |
| 14 | Explain advantages of Distributed computing? | 2 |
| 15 | What do you mean by Heuristic based NLP? | 2 |
| 16 | What do you mean Text Analysis? | 2 |

SECTION-C
(Descriptive Answer Type Questions)

Note: All Questions are compulsory: -

[4X4=16 Marks]

| S. No. | Question | Marks Allotted |
|--------|--|----------------|
| 17 | Explain the use of Machine Learning and Deep Learning for NLP. | 4 |
| 18 | What do you mean by Vector Space Models? | 4 |
| 19 | Explain the use of Tokenizer? | 4 |
| 20 | How Naïve Bayes works in NLP? | 4 |

SECTION-D
(Long Answer Type Questions)

Note: All Questions are compulsory: -

[4X6=24 Marks]

| S. No. | Question | Marks Allotted |
|---------------|--|-----------------------|
| 21 | Explain Architecture of HDFS? | 6 |
| 22 | What do you mean by NLP Project Pipeline? Explain Briefly. | 6 |
| 23 | Explain Recurrent Neural Network for NLP? | 6 |
| 24 | What do you mean by Email harvesting? Explain briefly. | 6 |