

Total No. of Questions : 5]

SEAT No. :

[Total No. of Pages : 2

**P5160**

[5823] - 604

T.Y.B.Sc.

**COMPUTER SCIENCE**

**CS-364 : Data Analytics**

**(CBCS 2019 Pattern) (Semester -VI)**

*Time : 2 Hours]*

*[Max. Marks : 35*

*Instructions to the candidates:*

- 1) *Figures to the right indicate full marks.*
- 2) *All questions are necessary.*
- 3) *Neat diagrams must be drawn wherever necessary.*

**Q1)** Attempt any EIGHT of the following :

**[8×1=8]**

- a) Define Data Analytics.
- b) Define Tokenization.
- c) Define Machine Learning.
- d) What is clustering?
- e) What is Frequent Itemset?
- f) What is data characterization?
- g) What is outlier?
- h) What is Bag of words?
- i) What is Text Analytics?
- j) Define Trend Analytics?

**Q2)** Attempt any FOUR of the following :

**[4×2=8]**

- a) What is confusion matrix?
- b) Define support and confidence in association rule mining.
- c) Explain any two Machine Learning (ML) Applications.
- d) Write a short note on stop words.
- e) Define supervise Learning and unsupervise Learning.

*P.T.O.*

**Q3) Attempt any Two of the following :** **[2×4=8]**

- a) What is prediction? Explain any one regression model in detail.
- b) Differentiate between Stemming and Lemmatization.
- c) Describe types of Data Analytics.

**Q4) Attempt any TWO of the following :** **[2×4=8]**

- a) Consider the following transactional database and find out Frequent Itemsets using Apriori algorithm with minimum support count=2

TID	List _ of _ Item_IDs
T <sub>1</sub>	I <sub>1</sub> , I <sub>2</sub> , I <sub>5</sub>
T <sub>2</sub>	I <sub>2</sub> , I <sub>4</sub>
T <sub>3</sub>	I <sub>2</sub> , I <sub>3</sub>
T <sub>4</sub>	I <sub>1</sub> , I <sub>2</sub> , I <sub>4</sub>
T <sub>5</sub>	I <sub>1</sub> , I <sub>3</sub>
T <sub>6</sub>	I <sub>2</sub> , I <sub>3</sub>
T <sub>7</sub>	I <sub>1</sub> , I <sub>3</sub>
T <sub>8</sub>	I <sub>1</sub> , I <sub>2</sub> , I <sub>3</sub> , I <sub>5</sub>
T <sub>9</sub>	I <sub>1</sub> , I <sub>2</sub> , I <sub>3</sub>

- b) Which are the challenges in social media analytics?
- c) Explain Reinforcement learning.

**Q5) Attempt any ONE of the following :** **[1×3=3]**

- a) Write a short note on support vector machine.
- b) Explain life cycle of Data Analytics.

