Reg. No. : ☐☐☐☐☐☐☐☐☐☐☐☐

Question Paper Code : 70060

B.E./B.Tech. DEGREE EXAMINATIONS, NOVEMBER/DECEMBER 2023.

Seventh Semester

Artificial Intelligence and Data Science

AD 8702 – TEXT ANALYTICS

(Regulations 2017)

Time : Three hours   Maximum : 100 marks

Answer ALL questions.

PART A — (10 × 2 = 20 marks)

1. Mention the use of Rapid Automatic Keyword Extraction.

2. Differentiate precision and recall.

3. What is Latent morpho-semantic analysis?

4. Define PARAFAC2 method.

5. Write short note on nonnegative matrix factorization.

6. Define information gain.

7. What is visualization?

8. Write a note on Future Lens.

9. Define feature extraction.

10. What is semantics?

PART B — (5 × 13 = 65 marks)

11. (a) Outline the steps in RAKE algorithm.

Or

(b) What is stoplist? Outline the steps in stoplist generation with an example.

12. (a) Illustrate K-means clustering with an example.

Or

(b) Discuss Tucker 1 method for document clustering with an example.

13. (a) Explain any one machine-learning algorithms, which have been successfully used in text classification with suitable example.

Or

(b) What is spam email classification? Outline the steps with an example.

14. (a) Elaborate Data Exploration with an example.

Or

(b) Discuss the adaptive threshold method for anomaly detection.

15. (a) Explain vector space modeling with an example.

Or

(b) Explain latent semantic analysis with an example.

PART C — (1 × 15 = 15 marks)

16. (a) Discuss the Adaptive threshold setting in novelty mining.

Or

(b) Outline the steps in embedding semantics in LDA topic modells with an example.

_____