SRM Institute of Science and Technology, Kattankulathur
Faculty of Engineering and Technology, Department of Information Technology
15IT330E - TEXT MINING
Cycle Test – III

Specialisation: IT
Date: 27.10.2018
Max. Marks: 50

Degree : B.Tech
Year/Sem: IV/VII
Duration : 2 period

Part-A [Answer any five]

(5x4=20 Marks)

1. Explain the phases and method of supervised model for document classification
2. Give any two real time application where Expectation maximization algorithm is applicable.
3. Give definition of polysemy and synonyms with example.
4. What are the two parameters of Naïve Bayes assumption?.
5. Give any five examples of features in Text Oriented Data Set.
6. Give a diagrammatic explanation of basic unit of computation in neural network
7. Explain any three activation function with its mathematical formula and diagram

Part-B

(2x15=30 Marks)

8. a) Derive and explain probabilistic latent semantic analysis in detail

(OR)

8. b) Explain in detail about Latent Dirichnet Allocation with example. How it is used in Information Retrieval?

9. a) Explain in detail about the support vector machine classification algorithm in detail.

(OR)

9. b) Consider the following data points. Apply k-mean clustering algorithm to find the final clusters. Where k=2.

Individual	Variable 1	Variable 2
1	1.0	1.0
2	1.5	2.0
3	3.0	4.0
4	5.0	7.0
5	3.5	5.0
6	4.5	5.0
7	3.5	4.5