



## **Sub Committee for Curriculum Development**

### **Data Science Specialization Post Graduate**

**Course Name :** Text Analytics

**(UG/PG):** PG

**Number of Credits:** 3

**Level:** 4

**Course Code:** T3448

**Learning Objective(s):** To understand the concepts of Natural Language Processing.  
To be able to analyze comments and questions in forums in order to find common pain points your customers face when using your products.  
To be able to perform sentiment analysis using text analytics in order to measure positive or negative perceptions of a company, brand, or product

#### **Pedagogy:**

Extensive use of case studies.  
Lecture and class discussions  
White paper oriented learning.  
Analyzing Data and implementing Text Analytics using R/ Python

#### **Pre-learning:**

Basic knowledge of Statistical Methods  
Basic Knowledge of Data Mining

## Course Outline:

Sr.No.	Topic	Hours
1	Introduction Text Analytics Text Analytics Methods Business Value of Text Analytics Application areas of Text Analytics Tools	3
2	Natural Language Processing I Parts of Speech and Morphology Phase Structure Corpus Based Work	6
3	Natural Language Processing II Collocations n-gram Models Parts of Speech Tagging chunking, syntax parsing and named entity recognition.	9
4	Document Representation Representing unstructured text documents with appropriate format and structure Statistical Text and Work Alignment	3
3	Text categorization and Text clustering Naive Bayes, k Nearest Neighbor (kNN) and Logistic Regression Support Vector Machines and Decision Trees, connectivity-based clustering and centroid-based clustering	6
4	Topic modeling Probabilistic Latent Semantic Indexing (pLSI) and Latent Dirichlet Allocation (LDA), and their variants for different application scenarios, including classification, image annotation, collaborative filtering, and hierarchical topical structure modeling.	6
5	Applied Topics Entity Extraction Document summarization Text Analytics in Social Media Sentiment analysis Text visualization	12
	<b>Total</b>	<b>45</b>

### Books Recommended:

1. Practical Text Analytics: Interpreting Text and Unstructured Data for Business Intelligence (Marketing Science), by Steven Struhl, 3 Jul 2015
2. Practical Text Mining And Statistical Analysis For Non-Structured Text Data Applications by Gary Miner By Gary Miner, Published by Elsevier Scienc
3. Mining Text Data. by Charu C. Aggarwal and ChengXiang Zhai, Springer, 2012.
4. Foundations of Statistical Natural Language Processing by Christopher D. Manning, Hinrich Schütze, 1999.
5. Speech & Language Processing. Dan Jurafsky and James H Martin, Pearson Education, India, 2000.

### Suggested Evaluation Methods:

Quizzes  
Case analysis  
Project report  
Presentation  
Midterm test  
End term test

### Parallel/Similar courses the existing curriculum:

S.No.	Name of the course	Institute where it was offered

Name of Member	Dr. Ajey Kumar				
Designation					
Org. / Inst.	SCIT				
Signature					

Name of the Expert:

Signature:

Date:

