

Module Code	DA3111	Semester 5	Module Title			Text Analytics					
Credits	2	Hours/Week			C	E	O	Evaluation %		Prerequisites	
GPA/NGPA	GPA	Lectures	Lab / Tutorials	Self-study				CA	WE	DA2111	
		-	4	2.66				100	-		
Module Aim:	This module aims to enable participants to get a hands-on experience on text mining.										
Learning Outcomes											
After completing this module, the students should be able to:											
LO-1	Identify the linguistic features of natural languages										
LO-2	Demonstrate the understanding of principles of various natural language processing and text retrieval techniques										
LO-3	Apply text analytics algorithms for real world problem solving										
Syllabus Outline										Learning Outcomes	
1	Overview [5 hrs] <i>What is NLP, basics linguistics, ambiguity and uncertainty in language</i>									LO-1	
2	Regular Expressions [5 hrs] <i>Regular languages, finite-state automata, morphology</i>									LO-2	
3	Basic Text Processing [5 hrs] <i>Tokenization and segmentation, normalization, morphological analysis</i>									LO-2, LO-3	
4	Semantics [5 hrs] <i>Representing meaning, semantic analysis, lexical semantics</i>									LO-2, LO-3	
5	Topic Modelling [5 hrs] <i>Probabilistic Semantic Indexing, Latent Dirichlet Allocation</i>									LO-2, LO-3	
6	Advanced Topics [5 hrs] <i>POS tagging, named entity recognition, machine translation</i>									LO-2, LO-3	
Assessments											
Assessment							Weight		Learning outcomes		
Continuous Assessments (CA)		Individual Assessment (Take home)					50% [4 hrs]		LO-1, LO-2		
		Group Assessment (Report & Presentation)					50% [3 hrs]		LO-2, LO-3		
References											