

Course Code	UDS21D06T	Course Name	SOCIAL MEDIA AND TEXT ANALYTICS	Course Category	D	Discipline Specific Elective	L	T	P	C
							4	0	0	4

Pre-requisite Courses	Nil	Co-requisite Courses	Nil	Progressive Courses	Nil
Course Offering Department	Computer Applications	Data Book / Codes/Standards	Nil		

Course Learning Rationale (CLR):		The purpose of learning this course is to:			Learning			Program Learning Outcomes (PLO)														
CLR-1 :	To leverage the power of the R eco-system to extract, process, analyze, visualize and model social media data				1	2	3	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
CLR-2 :	Visualize and analyze data from social media platforms to understand and model complex relationships using various concepts and techniques				Learning (Bloom)	Efficiency (%)	Retention (%)	Knowledge	Concepts	Related Disciplines	Knowledge	Specialization	Core Knowledge	Learning	Interpret Data	Skills	Learning Skills	Communication Skills	Soft Skills	Behavior	Learning	
CLR-3 :	Understand the fundamentals of text mining																					
CLR-4 :	Utilize text for prediction techniques																					
CLR-5 :	Understand the relevance between information retrieval and text mining																					
CLR-6 :	Analyze different case studies related to text mining																					

Course Learning Outcomes (CLO):	At the end of this course, learners will be able to:	Level of Thinking (Bloom)	Expected Proficiency (%)	Expected Attainment (%)	Fundamental Knowledge	Application of Concepts	Link with Related Disciplines	Procedural Knowledge	Skills in Specialization	Ability to Utilize Knowledge	Skills in Modeling	Analyze, interpret Data	Investigative Skills	Problem Solving Skills	Communication Skills	Analytical Skills	ICT Skills	Professional Behavior	Life Long Learning
CLO-1 :	Understand the basics of social media analytics and R language	3	80	70	L	H	-	H	L	-	-	-	L	L	-	H	-	-	-
CLO-2 :	Analyze data from major social media channels such as Twitter & Flickr	3	85	75	M	H	L	M	L	-	-	-	M	L	-	H	-	-	-
CLO-3 :	Acquire knowledge on fundamentals of text mining	3	75	70	M	H	M	H	L	-	-	-	M	L	-	H	-	-	-
CLO-4 :	Perform prediction from text and evaluate it	3	85	80	M	H	M	H	L	-	-	-	M	L	-	H	-	-	-
CLO-5 :	Perform document matching	3	85	75	H	H	M	H	L	-	-	-	M	L	-	H	-	-	-
CLO-6 :	Understand how text mining is implemented	3	80	70	L	H	-	H	L	-	-	-	L	L	-	H	-	-	-

Duration (hour)	12	12	12	12	12
S-1	SLO-1 Getting Started with R and Social Media Analytics	Visualizing data	Overview of Text Mining	Using Text for Prediction	Finding Structure in a Document Collection
	SLO-2 Understanding Social Media	Managing packages	What's Special About Text Mining?	Recognizing that Documents Fit a Pattern	
S-2	SLO-1 Advantages and Significance of Social Media	Data analytics - Analytics workflow	Structured or Unstructured Data	How Many Documents Are Enough?	Clustering Documents by Similarity
	SLO-2 Disadvantages and Pitfalls of Social Media	Machine learning techniques	Is Text Different from Numbers?	Document Classification	Similarity of Composite Documents
S-3	SLO-1 Social media analytics	Supervised learning, Unsupervised learning	What Types of Problems Can Be Solved?	Learning to Predict from Text	k-Means Clustering
	SLO-2 A typical social media analytics workflow	Text analytics	Document Classification	Similarity and Nearest-Neighbor Methods	

S-4	SLO-1	Data access, Data processing and normalization, Data analysis and Insights	Understanding Twitter, APIs	Information Retrieval	Document Similarity	Hierarchical Clustering
	SLO-2	Opportunities and Challenges	Registering an application	Clustering and Organizing Documents	Decision Rules	
S-5	SLO-1	Getting started with R	Connecting to Twitter using R	Information Extraction	Decision Trees	The EM Algorithm
	SLO-2	Environment setup	Extracting sample Tweets	Prediction and Evaluation	Scoring by Probabilities	
S-6	SLO-1	Data types	Trend analysis	From Textual Information to Numerical Vectors	Linear Scoring Methods	What Do a Cluster's Labels Mean?
	SLO-2	Data structures-Vectors		Collecting Documents	Evaluation of Performance - Estimating Current and Future Performance	Applications, Evaluation of Performance
S-7	SLO-1	Arrays	Sentiment analysis	Document Standardization	Getting the Most from a Learning Method	Case Study: Market Intelligence from the Web
	SLO-2	Matrices	Key concepts of sentiment analysis - Subjectivity, Sentiment polarity	Tokenization	Errors and Pitfalls in Big Data Evaluation	
S-8	SLO-1	Lists	Opinion summarization	Lemmatization-Inflectional Stemming	Information Retrieval and Text Mining	Case Study: Lightweight Document Matching for Digital Libraries
	SLO-2	Data Frames	Features	Stemming to a Root	Is Information Retrieval a Form of Text Mining?	
S-9	SLO-1	Functions - Built-in functions	Sentiment analysis in R	Vector Generation for Prediction	Key Word Search	Mining Social Media
	SLO-2	User-defined functions		Multiword Features	Nearest-Neighbor Methods	
S-10	SLO-1	Controlling code flow - Looping constructs	Follower graph analysis	Labels for the Right Answers, Feature Selection by Attribute Ranking	Measuring Similarity -Shared Word Count	E-mail Filtering
	SLO-2	Conditional constructs	Flickr Data Analysis	Sentence Boundary Determination	Word Count and Bonus, Cosine Similarity	
S-11	SLO-1	Advanced operations	Accessing Flickr's data	Part-of-Speech Tagging	Web-Based Document Search - Link Analysis	Emerging Directions
	SLO-2	apply, lapply	Understanding Flickr data	Word Sense Disambiguation	Document Matching	Summarization
S-12	SLO-1	sapply, tapply	Understanding interestingness - similarities	Phrase Recognition, Named Entity Recognition, Parsing	Inverted Lists	Active Learning
	SLO-2	mapply	Are your photos interesting? - Preparing the data -Building the classifier	Feature Generation	Evaluation of Performance	Learning with Unlabeled Data

Learning Resources	1. Raghav Bali, Dipanjan Sarkar, Tushar Sharma, (2017), "Learning Social Media Analytics with R", Packt Publishing.	2. Sholom M. Weiss, Nitin Indurkha, Tong Zhang, (2015), "Fundamentals of Predictive Text Mining", Second Edition, Springer London.
---------------------------	---	--

Learning Assessment											
	Bloom's Level of Thinking	Continuous Learning Assessment (50% weightage)								Final Examination (50% weightage)	
		CLA – 1 (10%)		CLA – 2 (10%)		CLA – 3 (20%)		CLA – 4 (10%)#			
		Theory	Practice	Theory	Practice	Theory	Practice	Theory	Practice	Theory	Practice
Level 1	Remember	40%	-	40%	-	40%	-	40%	-	40%	-
	Understand										
Level 2	Apply	40%	-	40%	-	40%	-	40%	-	40%	-
	Analyze										
Level 3	Evaluate	20%	-	20%	-	20%	-	20%	-	20%	-
	Create										
	Total	100 %		100 %		100 %		100 %		100 %	

Course Designers		
Experts from Industry	Experts from Higher Technical Institutions	Internal Experts
Mr.Jothi, Periyasamy , Chief AI Architect DeepSphere AI, CA, USA	Dr.S.Gopinathan, Associate Professor, University of Madras, Chennai	Mrs.S Chandrakala, SRMIST
		Mrs. Kanmani, SRM IST