

Branch Name:	Information and Communication Technology
Program Code:	ET208
Course Name:	Natural Language Processing (Elective-IV)
Course Code:	2ET2080305T
Pre-requisite Course:	Probability and statistics, Programming and data structures

Course Objective:

1. To familiarize the concepts and techniques of Natural language Processing for analyzing Words based on Morphology and CORPUS.
2. Tolerate mathematical foundations, Probability theory with Linguistic essentials such as Syntactic and semantic analysis of text.
3. To apply the Statistical learning methods and cutting-edge research models from deep Learning

Teaching and Examination Scheme:

Credit	Teaching Scheme Hours per week				Evaluation Scheme (Marks)								Total (Marks)
					Theory (Marks)				Practical (Marks)				
	L	T	P	Total	SEE		CCA		SEE		CCA		
					Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.	
3	3	-	-	3	24	60	16	40	-	-	-	-	100

Subject Contents:

Sr. No	Topic	Total Hours	Weight (%)
1	Introduction to NLP What is NLP? Why NLP is Difficult? History of NLP, Advantages of NLP, Disadvantages of NLP, Components of NLP, Applications of NLP, How to build an NLP pipeline? Phases of NLP, NLP APIs, NLP Libraries	06	15
2	Language Modeling and Part of Speech Tagging Unigram Language Model, Bigram, Trigram, N-gram, Advanced smoothing for language modeling, Empirical Comparison of Smoothing Techniques, Applications of Language Modeling, Natural Language Generation, Parts of Speech Tagging, Morphology, Named Entity Recognition	12	25
3	Words and Word Forms Bag of words, skip-gram, Continuous Bag-Of-Words, Embedding representations for words Lexical Semantics, Word Sense Disambiguation, Knowledge Based and Supervised Word Sense Disambiguation	07	15
4	Text Analysis, Summarization and Extraction Sentiment Mining, Text Classification, Text Summarization, Information Extraction, Named Entity Recognition, Relation Extraction, Question Answering in Multilingual Setting; NLP in Information Retrieval, Cross-Lingual IR	10	25

5	Machine Translation and Applications of NLP Need of MT, Problems of Machine Translation, MT Approaches, Direct Machine Translations, Rule-Based Machine Translation, Knowledge Based MT System, Statistical Machine Translation (SMT), Parameter learning in SMT (IBM models) using EM), Encoder-decoder architecture, NeuralMachine Translation, Applications of NLP	10	20
---	---	----	----

Text Books:

1. Speech and Language Processing: AnIntroduction to Natural Language Processing, Computational
2. Linguistics and Speech Recognition Jurafsky, David, and James H. Martin, PEARSON

Reference Books:

1. Foundations of Statistical Natural Language Processing, Manning, Christopher D., and Hinrich Schütze, Cambridge, MA: MIT Press
2. Natural Language Understanding, James Allen. The Benjamin/Cummings Publishing Company Inc.
3. Natural Language Processing with Python – Analyzing Text with the Natural Language ToolkitStevenBird, Ewan Klein, and Edward Loper.

Course Learning Outcomes (CLO): On completion of this course, the students will be able to:

CLO	Description	Bloom's Taxonomy Level
CLO1	Understand comprehend the key concepts of NLP and identify the NLPchallenges and issues	2 Understanding
CLO2	Develop Language Modeling for various text corpora across the different languages	6 Creating
CLO3	Illustrate computational methods to understand language phenomena of word sense disambiguation	2 Understanding 3 Applying
CLO4	Design and develop applications for text or information extraction/summarization/classification	3 Applying, 2 Understanding
CLO5	Apply different Machine translation techniques for translating a source to target language(s)	2 Understanding

Mapping of CLOs with POs & PSOs

Course Learning Outcomes	Program Outcomes (POs)												Program Specific Outcomes(PSOs)			
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	PSO3	PSO4
CLO1	H	M	M	L							L		H	M		
CLO2	H	M	M	L	M					L			H	M	H	M
CLO3	H	M	M	L							L		H	M	M	M
CLO4	M	L	L		H					L			M	M	M	M
CLO5	H	M	M										M		H	M

H: High, M: Medium, L: Low