

Natural Language Processing

Course Code	20CS4703D	Year	IV	Semester	1
Course Category	PEC	Branch	CSE	Course Type	Theory
Credits	3	L-T-P	3-0-0	Prerequisites	Machine Learning
Continuous Internal Evaluation :	30	Semester End Evaluation:	70	Total Marks:	100

Course Outcomes

Upon successful completion of the course, the student will be able to

CO1	Understand the fundamental concepts of Natural language processing.	L2
CO2	Apply basic evaluating language models for the probability of the test set.	L3
CO3	Apply techniques for extracting limited forms of semantic content from texts.	L3
CO4	Analyze parsing algorithms through the use of context-free grammars.	L4

Contribution of Course Outcomes towards achievement of Program Outcomes & Strength of correlations

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO10	PO11	PO12	PSO1	PSO2
CO1	√													
CO2										√			√	
CO3	√									√				
CO4		√							√	√		√		

Syllabus		
Unit No	Contents	Mapped CO
I	Regular Expressions, Text Normalization, Edit Distance- Regular Expression, Words, Corpora, Text Normalization, Minimum Edit Distance. N-Gram Language Models- NGrams, Evaluating Language Models, Generalization and Zeros, Smoothing, Kneser-Ney Smoothing, The web and stupid Backoff.	CO1,CO2
II	Labeling for Parts of Speech- English Word Classes, Part-of-Speech Tagging, Named Entities and Named Entity Tagging, HMM Part-of-Speech Tagging, Conditional Random Fields (CRFs), Evaluation of Named Entity Recognition.	CO1,CO2
III	Formal Grammars of English- Constituency, Context-Free Grammars, Some Grammar Rules for English, Treebanks, Grammar Equivalence and Normal Form, Lexicalized Grammars. Syntactic Parsing- Ambiguity,CKY Parsing: A Dynamic Programming Approach, Span-Based Neural Consistency Parsing.	CO1,CO4
IV	Dependency Parsing- Dependency Relations, Dependency Formalisms, Dependency Tree banks,Transition-Based Dependency Parsing, Graph- Based Dependency Parsing, Evaluation. Representation of Sentence Meaning- Computational desiderata for Representations, Model-Theoretic Semantics, First-Order Logic, Event and State Representations, Description Logics.	CO1,CO3
V	Semantic Parsing, Information Extraction- Relation Extraction, Relation Extraction Algorithms, Extracting Times, Extracting Events and their Times, Template Filling. Lexicons for Sentiment, Affect and Connotation- Defining Emotion, Available Sentiment and Affect Lexicons, Creating Affect Lexicons by Human Labeling, Semi-Supervised Induction of Affect Lexicons, Supervised Learning of Word Sentiment, Using Lexicons for sentiment Recognition.	CO1,CO3

Learning Resources
Text Books
1. Speech and Language Processing: An introduction to Natural Language Processing,Computational Linguistics and Speech Recognition, Daniel Jurafsky and James H Martin, Third Edition, 2020.
References
1. Natural Language Processing Recipes, Akshay Kulkarni, AdarshaShivananda, 2019, Apress. 2. Applied Text Analysis with Python, Benjamin Bengfort, Tony Ojeda, Rebecca Bilbro, 2018, O'Reilly Media. 3. Natural Language Processing: An information Access Perspective by Kavi Narayana Murthy,2006, EssEss Publications. 4. Statistical Language Learning, Charniack, Eugene, 1993, MIT Press.

e-Resources and other Digital Material

1. <https://web.stanford.edu/~jurafsky/slp3/>
2. https://swayam.gov.in/nd1_noc19_cs56/preview
3. <https://online.stanford.edu/courses/xcs224n-natural-language-processing-deep-learning>
4. <https://www.coursera.org/specializations/natural-language-processing>