

## Natural Language Processing

<b>Course Code</b>	19CS4801A	<b>Year</b>	IV	<b>Semester</b>	II
<b>Course Category</b>	Program Elective-VI	<b>Branch</b>	CSE	<b>Course Type</b>	Theory
<b>Credits</b>	3	<b>L-T-P</b>	3-0-0	<b>Prerequisites</b>	Machine Learning
<b>Continuous Internal Evaluation:</b>	30	<b>Semester End Evaluation:</b>	70	<b>Total Marks:</b>	100

<b>Course Outcomes</b>		
Upon successful completion of the course, the student will be able to		
<b>CO1</b>	Understand the fundamental concepts of natural language processing/generation.	<b>L2</b>
<b>CO2</b>	Apply basic evaluating language models for the probability of the test set.	<b>L3</b>
<b>CO3</b>	Apply techniques for extracting limited forms of semantic content from texts.	<b>L3</b>
<b>CO4</b>	Analyze parsing algorithms through the use of context-free grammars and prepare an effective report.	<b>L4</b>

<b>Syllabus</b>		
<b>Unit No</b>	<b>Contents</b>	<b>Mapped CO</b>
<b>I</b>	<b>Regular Expressions, Text Normalization, Edit Distance-</b> Regular Expression, Words, Corpora, Text Normalization, Minimum Edit Distance. <b>N-Gram Language Models-</b> NGrams, Evaluating Language Models, Generalization and Zeros, Smoothing, Kneser-Ney Smoothing, Add-1 Smoothing, Add k-smoothing, Backoff and Interpolation, Katz backoff, Good-Turing backoff.	<b>CO1, CO2</b>
<b>II</b>	<b>Labeling for Parts of Speech-</b> 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.	<b>CO1, CO2</b>
<b>III</b>	<b>Formal Grammars of English-</b> Constituency, Context-Free Grammars, Some Grammar Rules for English, Treebanks, Grammar Equivalence and	<b>CO1, CO4</b>

	Normal Form, Lexicalized Grammars. <b>Syntactic Parsing-</b> Ambiguity, CKY Parsing: A Dynamic Programming Approach, CCG Parsing.	
<b>IV</b>	<b>Dependency Parsing-</b> Dependency Relations, Dependency Formalisms, Transition-Based Dependency Parsing, Graph- Based Dependency Parsing. <b>Representation of Sentence Meaning-</b> Model- Theoretic Semantics, First-Order Logic, Event and State Representations, Description Logics.	<b>CO1, CO3</b>
<b>V</b>	<b>Semantic Parsing, Information Extraction-</b> Relation Extraction, Relation Extraction Algorithms, Extracting Times, Template Filling. <b>Lexicons for Sentiment, Affect and Connotation-</b> Defining Emotion, Available Sentiment and Affect Lexicons, Semi-supervised Induction of Affect Lexicons, Supervised Learning of Word Sentiment.	<b>CO1, CO3</b>

<b>Learning Resources</b>
<b>Text Books</b>
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.
<b>References</b>
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.
<b>e-Resources and other Digital Material</b>
1. <a href="https://web.stanford.edu/~jurafsky/slp3/">https://web.stanford.edu/~jurafsky/slp3/</a> 2. <a href="https://swayam.gov.in/nd1_noc19_cs56/preview">https://swayam.gov.in/nd1_noc19_cs56/preview</a> 3. <a href="https://online.stanford.edu/courses/xcs224n-natural-language-processing-deep-learning">https://online.stanford.edu/courses/xcs224n-natural-language-processing-deep-learning</a> 4. <a href="https://www.coursera.org/specializations/natural-language-processing">https://www.coursera.org/specializations/natural-language-processing</a>