

III YEAR SEM-II B.Tech CSE	ELECTIVE	L	T	P	C
CS5116	Natural Language Processing	2	2	0	4

OBJECTIVES

1	To provide the student with knowledge of various levels of analysis involved in NLP
2	To understand language modeling
3	To gain knowledge in automated natural language generation and machine translation

Unit I

Origins and challenges of NLP-Language and Grammar-Processing Indian Languages-NLP Applications-Information Retrieval. Introduction-Various Grammar-based Language Models-Statistical Language Model, N-Grams.

Unit II

Introduction- Regular Expressions-Finite-State Automata-Morphological Parsing-Spelling Error Detection and correction-Words and Word classes-Part-of Speech Tagging. Introduction-Context-free Grammar-Constituency Parsing-Probabilistic Parsing

Unit III

Introduction- Meaning Representation-Lexical Semantics Ambiguity-Word Sense Disambiguation. Introduction-cohesion-Reference Resolution Discourse Coherence and Structure.

Unit IV

Introduction-Architecture of NLG Systems, Generation Tasks and Representations-Application of NLG. Introduction-Problems in Machine Translation Characteristics of Indian Languages- Machine Translation Approaches-Translation involving Indian Languages

Unit V

Introduction-Design features of Information Retrieval Systems-Classical, Non-classical, Alternative Models of Information Retrieval - Evaluation

Unit VI

Introduction-WordNet-FrameNet-Stemmers-POS Tagger Research Corpora , HMM in POS Tagging.

Text Books:

1. Tanveer Siddiqui, U.S. Tiwary, "Natural Language Processing and Information Retrieval", Oxford University Press, 2008.

Reference Books:

1. Daniel Jurafsky and James H Martin, "Speech and Language Processing: An introduction to Natural Language Processing, Computational Linguistics and Speech Recognition", Prentice Hall, 2nd Edition, 2008.
2. James Allen, Benjamin Cummings, "Natural Language Understanding", 2nd edition, 1995.