

Nirma University
Institute of Technology, School of Technology
MTech Computer Science and Engineering (Data Science)
Semester – II

L	T	P	C
3	0	2	4

Course Code	3CS42D101
Course Name	Natural Language Computing

Course Learning Outcomes (CLOs):

At the end of the course, students will be able to

1. comprehend the key concepts of NLP which are used to describe and analyse language
2. perform POS tagging and generate context free grammar for English language
3. realize semantics and pragmatics of English language for processing
4. implement natural language processing task

Syllabus:

**Teaching
Hours**

Unit I

Introduction: Knowledge in speech and language processing, Ambiguity, Models and algorithms, Language, Thought, and Understanding, State-of-the-art, History.

2

Unit II

Words: Regular expressions and automata, Morphology and Finite-State Transducers, Computational Phonology and Text-to-Speech, Probabilistic Models of Pronunciation and Spelling, N-grams, HMMs and Speech Recognition

8

Unit III

Syntax: Word classes and Part-of-speech tagging, Context-free grammars for English, Parsing with context-free grammars, Features and unification, Lexicalized and probabilistic parsing

9

Unit IV

Semantics: Representing meaning, Semantic analysis, Lexical semantics, Word Sense Disambiguation and Information Retrieval

9

Unit V

Pragmatics: Discourse, Dialogue and Conversational Agents, Generation, Machine Translation

9

Unit VI

Natural Language Processing: using Deep Learning: Data representation in NLP, Distributed representations, Word2Vec, applying word embeddings, sequence to sequence learning, Application of Deep learning in NLP for author attribution, text classification, word generation

8

Self Study:

The self-study contents will be declared at the commencement of semester. Around 10% of the questions will be asked from self-study contents.

Laboratory Work:

Laboratory work will be based on above syllabus with minimum 5 experiments to be incorporated.

Suggested Readings[^]:

1. Jurafsky, D., & Martin, J. H., Speech and language processing (Vol. 3). London: Pearson.
2. Allen, J., Natural language understanding. Pearson.
3. Charniak, E., Statistical language learning. MIT press.
4. Manning, C. D., Manning, C. D., & Schütze, H., Foundations of statistical natural language processing. MIT press.
5. Dale, R., Moisl, H., & Somers, H., Handbook of natural language processing. CRC Press.
6. Radford, A., Atkinson, M., & Britain, D., Linguistics: an introduction. Cambridge University Press.

L=Lecture, T=Tutorial, P=Practical, C=Credit

[^]this is not an exhaustive list