



MARRI LAXMAN REDDY INSTITUTE OF TECHNOLOGY AND MANAGEMENT

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

COURSE CONTENT

NATURAL LANGUAGE PROCESSING								
I Semester: CSE								
Course Code	Category	Hours / Week			Credits	Maximum Marks		
		L	T	P		C	CIA	SEE
2515807	Foundation	3	0	0	3	40	60	100
Contact Classes: 45	Tutorial Classes: Nil	Practical Classes: Nil			Total Classes: 45			
Prerequisites: Data structures, finite automata and probability theory.								

Course Overview:

This focuses on enabling computers to understand, interpret, and generate human language. This course introduces the fundamental concepts, techniques, and tools required to process and analyze text data using computational methods.

Course Objectives:

1. To understand the fundamental concepts and techniques of Natural Language Processing.
2. To develop skills in morphological, syntactic, and semantic analysis of text data.
3. To apply parsing and language modeling techniques for text processing applications.
4. To design NLP solutions for information retrieval, machine translation, and dialogue systems.
5. To analyze multilingual and cross-lingual language processing methods using statistical and computational approaches.

Course Outcomes: After Completion of the Course, Students should be able to

1. Apply morphological analysis techniques for improving text preprocessing in information retrieval systems.
2. Construct syntactic parsers using treebanks and parsing algorithms to develop multilingual grammar checking tools.
3. Implement semantic parsing and word sense disambiguation to enhance question-answering and knowledge extraction systems.
4. Analyze predicate-argument structures and meaning representation systems to support automated reasoning in intelligent dialogue systems.
5. Evaluate n-gram and cross-lingual models for text prediction and machine translation applications.

UNIT - I: Finding the Structure of Words: Words and Their Components, Issues and Challenges, Morphological Models

Finding the Structure of Documents: Introduction, Methods, Complexity of the Approaches, Performances of the Approaches.

UNIT - II: Syntax Analysis: Parsing Natural Language, Treebanks: A Data-Driven Approach to Syntax, Representation of Syntactic Structure, Parsing Algorithms, Models for Ambiguity Resolution in Parsing, Multilingual Issues.

UNIT-III: Semantic Parsing: Introduction, Semantic Interpretation, System Paradigms, Word Sense Systems, Software.

UNIT - IV: Predicate-Argument Structure, Meaning Representation Systems, Software.

UNIT - V: Discourse Processing: Cohesion, Reference Resolution, Discourse Cohesion and Structure Language Modeling: Introduction, N-Gram Models, Language Model Evaluation, Parameter Estimation, Language Model Adaptation, Types of Language Models, Language-Specific Modeling Problems, Multilingual and Cross Lingual Language Modeling

TEXT BOOKS:

1. Multilingual natural Language Processing Applications: From Theory to Practice – Daniel M. Bikel and Imed Zitouni, Pearson Publication.
2. Natural Language Processing and Information Retrieval: Tanvier Siddiqui, U.S. Tiwary.

REFERENCE BOOKS:

1. Speech and Natural Language Processing - Daniel Jurafsky & James H Martin, Pearson Publications.

ELECTRONIC RESOURCES:

1. <https://www.coursera.org/specializations/natural-language-processing>
2. <https://www.mygreatlearning.com/academy/learn-for-free/courses/introduction-to-natural-language-processing>
3. <https://alison.com/course/natural-language-processing-nlp>
4. <https://www.edx.org/learn/natural-language-processing>

MATERIALS ONLINE:

1. Course template
2. Tutorial question bank
3. Tech talk and Concept Video topics
4. Open-ended experiments
5. Definitions and terminology
6. Assignments
7. Model question paper – I
8. Model question paper – II
9. Lecture notes
10. E-Learning Readiness Videos (ELRV)