



# 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 LABORATORY								
I Semester: CSE								
Course Code	Category	Hours / Week			Credits	Maximum Marks		
2515879	Professional Core courses	L	T	P	C	CIA	SEE	Total
		0	0	4	2	40	60	100
Contact Classes: Nil	Tutorial Classes: 0	Practical Classes: 60			Total Classes: 60			
requisites: Strong foundation in Data Structures								

### Course Overview:

The Natural Language Processing Lab provides practical training in implementing text preprocessing and language modeling techniques using Python

### Course Objectives:

1. To understand the fundamentals of text preprocessing and language modeling techniques in NLP.
2. To develop Python programs for tokenization, stemming, stop-word removal, and POS tagging.
3. To apply morphological analysis, chunking, and word generation techniques in NLP applications.
4. To implement N-gram models and smoothing techniques for probabilistic language processing.
5. To evaluate NLP models and preprocessing methods for real-world text analytics applications

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

1. Implement Python programs for text preprocessing tasks such as tokenization, stop-word removal, and stemming to prepare textual data for NLP applications like chatbots and search engines.
2. Apply part-of-speech tagging and morphological analysis to extract syntactic information for tasks such as grammar checking and text parsing.
3. Develop word analysis and word generation techniques for building applications like auto-completion and predictive text systems.
4. Construct chunking and N-gram models, including smoothing techniques, to perform phrase detection and probabilistic language modelling for speech recognition and text prediction.
5. Evaluate the performance of NLP preprocessing and language modelling methods for real-world applications such as document summarization and sentiment analysis.

## LIST OF EXPERIMENTS

Implement the following using Python

1. Tokenization
2. Stemming
3. Stop word removal (a, the, are,..)
4. Word Analysis
5. Word Generation
6. Pos tagging
7. Morphology
8. chunking
9. N-Grams
10. N-Grams Smoothing

## 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/learn/natural-language-processing>
2. <https://www.geeksforgeeks.org/natural-language-processing/>
3. <https://realpython.com/natural-language-processing-spacy-python/>
4. <https://nptel.ac.in/courses/106/106/106106177/>

## MATERIALS ONLINE:

1. Course template
  2. Open-ended experiments
  3. Definitions and terminology
  4. Lab Manual
-

