

<b>COURSE TITLE</b>	<b>NATURAL LANGUAGE PROCESSING</b>
<b>COURSE CODE</b>	<b>01CE0718</b>
<b>COURSE CREDITS</b>	<b>3</b>

**Objective:**

- 1 The objective of this course is to provide students with a solid foundation in Natural Language Processing by allowing them to analyze and process language, understand semantics, and apply language modeling and speech tagging effectively. Students will also gain the skills necessary to tackle discourse-related challenges and excel in information retrieval tasks in the context of NLP.

**Course Outcomes:** After completion of this course, student will be able to:

- 1 Understand NLP concepts, phases, and applications like spell-checking, summarization, and machine translation
- 2 Apply word-level analysis and syntactic techniques effectively
- 3 Analyse and apply semantic analysis methods and language modelling
- 4 Evaluate language modelling and speech tagging techniques for various tasks
- 5 Create discourse solutions and information retrieval strategies based on real-world needs

**Pre-requisite of course:** Basic Programming Skills, Statistics and Probability, Mathematics Skills

**Teaching and Examination Scheme**

<b>Theory Hours</b>	<b>Tutorial Hours</b>	<b>Practical Hours</b>	<b>ESE</b>	<b>IA</b>	<b>CSE</b>	<b>Viva</b>	<b>Term Work</b>
2	0	2	50	30	20	25	25

<b>Contents : Unit</b>	<b>Topics</b>	<b>Contact Hours</b>
1	<b>Introduction:</b> What is NLP, NLP terminology, Different Phases of NLP, (Morphological, Syntactic, Semantic, Discourse Integration, Pragmatic Analysis), Applications of NLP-Spell-checking, Summarization	4
2	<b>Word Level Analysis and Syntactic Analysis</b> Word Level Analysis: Regular Expressions, Morphological Parsing,, Types of Morphemes , Syntactic Analysis: Concept of Parser, Derivation, Parse Tree Generation, Context Free Grammar	6
3	<b>Semantic Analysis:</b> First order predicate logic (FOPL), Semantic Nets, Frames, Conceptual dependency (CD), Rule-based architecture, Case Grammar, Conceptual Graphs, Relationship Extraction	6

<b>Contents : Unit</b>	<b>Topics</b>	<b>Contact Hours</b>
4	<b>Language Modelling and Speech tagging:</b> Language Modelling: N-gram model, Counting Words - Unsmoothed N gram,, Smoothing- Back off Deleted Interpolation – Entropy, Hidden Markov Models(HMM),Maximum Entropy Markov Models(MEMM), Conditional Random Fields(CRF) , Speech tagging : Rule Based Part of Speech Tagging - Stochastic Part of Speech Tagging, Transformation Based Tagging(TBL)	6
5	<b>Discourse Processing and Information Retrieval</b> Discourse Processing and Information Retrieval, Discourse Cohesion and Structure, Information Retrieval :Basic IR Model, Vector space model, Classical Problem, Ad-hoc retrieval problem- Query Formulation, Inverted Index, Stop Word Elimination, Stemming	6
<b>Total Hours</b>		<b>28</b>

#### Suggested List of Experiments:

<b>Contents : Unit</b>	<b>Topics</b>	<b>Contact Hours</b>
1	<b>Practical 1</b> Word Analysis	2
2	<b>Practical 2</b> Word Generation	2
3	<b>Practical 3</b> Morphologic Analysis	2
4	<b>Practical 4</b> N-Gram Model	2
5	<b>Practical 5</b> N-Grams Smoothing	2
6	<b>Practical 6</b> POS Tagging: Hidden Markov Model	2
7	<b>Practical 7</b> POS Tagging: Viterbi Decoding	2
8	<b>Practical 8</b> Building POS Tagger	2
9	<b>Practical 9</b> Chunking	2
10	<b>Practical 10</b> Building Chunker	2
11	<b>Practical 11</b> Stemming And Lemmatization	2
12	<b>Practical 12</b> Word Sense Disambiguation	2
13	<b>Practical 13</b> Information Retrieval	2

### Suggested List of Experiments:

Contents : Unit	Topics	Contact Hours
14	<b>Practical 14</b> CASE STUDY : Application of NLP- Sentiment Analysis of tweets in Twitter platform	2
<b>Total Hours</b>		<b>28</b>

### Textbook :

- 1 Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech, Daniel Jurafsky , James H. Martin, Pearson Publication, 2014

### References:

- 1 “ Natural Language Processing”, “ Natural Language Processing”, Ela Kumar, IK international Publication, second edition, 2014
- 2 “Introduction to Natural Language Processing”, “Introduction to Natural Language Processing”, Jacob Eisenstein, MIT Press, 2019
- 3 “Natural Language Processing: A Paninian Perspective”, “Natural Language Processing: A Paninian Perspective”, Bharati Akshar, Chaitanya Vineet, Sangal, Rajeev, Bharati Akshar, Chaitanya Vineet, Sangal, Rajeev, 1995

### Suggested Theory Distribution:

The suggested theory distribution as per Bloom’s taxonomy is as follows. This distribution serves as guidelines for teachers and students to achieve effective teaching-learning process

Distribution of Theory for course delivery					
Remember / Knowledge	Understand	Apply	Analyze	Evaluate	Higher order Thinking / Creative
15.00	25.00	25.00	20.00	15.00	0.00

### Instructional Method:

- 1 The course delivery method will depend upon the requirement of content and need of students. The teacher in addition to conventional teaching method by black board, may also use any of tools such as demonstration, role play, Quiz, brainstorming, MOOCs etc
- 2 The internal evaluation will be done on the basis of continuous evaluation of students in the laboratory and class-room.
- 3 Practical examination will be conducted at the end of semester for evaluation of performance of students in laboratory
- 4 Students will use supplementary resources such as online videos, NPTEL videos, e- courses, Virtual Laboratory.

### Supplementary Resources:

- 1 [https://onlinecourses.nptel.ac.in/noc23\\_cs45/preview](https://onlinecourses.nptel.ac.in/noc23_cs45/preview)

**Supplementary Resources:**

- 2 <https://www.coursera.org/specializations/natural-language-processing>
- 3 <https://www.udemy.com/course/natural-language-processing/>
- 4 <https://archive.nptel.ac.in/courses/106/105/106105158/>