SWARNANDHRA
COLLEGE OF ENGINEERING&TECHNOLOGY
(AUTONOMUS)
Seetharamapuram, NARASAPUR – Pin: 534 280

## DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

# LESSON PLAN

| Course Code | Course Title | Semester | Branches | Conduct Periods /Week | A.Y | Date of commencement of Semester |
|---|---|---|---|---|---|---|
| 23AM5T02 | Natural Language processing | V | AI&ML | 6 | 2025-26 | 09/07/2025 |

**COURSE OUTCOMES**

| | |
|---|---|
| 1 | **CO1:** Introduce the fundamentals of NLP, language modeling, and text preprocessing techniques. |
| 2 | **CO2:** Study statistical models for word prediction and part-of-speech tagging techniques. |
| 3 | **CO3:** Understand syntactic parsing techniques using grammar rules and probabilistic models. |
| 4 | **CO4:** Explore meaning representation, semantic roles, and word sense disambiguation techniques. |
| 5 | **CO5:** Study discourse structures, reference resolution, and key lexical resources in NLP. |

| UNITS | Out Comes/ BTKL | Topic No. | Topics/Activity | Text Book / Reference | Conduct Hour | Delivery Method |
|---|---|---|---|---|---|---|
| I | **CO1:** Introduce the fundamentals of NLP, language modeling, and text preprocessing techniques. | | **1. INTRODUCTION** | | | Chalk, Talk, & Tutorials |
| | | 1.1 | Origins and challenges of NLP | T1,T2 | 1 | |
| | | 1.2 | Language Modelling | T1,T3 | 1 | |
| | | 1.3 | Grammar-based LM | T1,T2,R3 | 1 | |
| | | 1.4 | Statistical LM | T1,T2,R3 | 1 | |
| | | 1.5 | Regular Expressions | T1,R3 | 1 | |
| | | 1.6 | Finite-State Automata | R2,R3 | 1 | |
| | | 1.7 | English Morphology | T1,T2 | 1 | |
| | | 1.8 | Transducers for lexicon and rules | T1,R2,R3 | 1 | |
| | | 1.9 | Tokenization | T1, T3 | 1 | |
| | | 1.10 | Detecting and Correcting Spelling Errors | T1, R3 | 1 | |
| | | 1.11 | Minimum Edit Distance | T1,R2,R3 | 1 | |
| | | 1.12 | Assignment 1 | T1,T2,T3 | 1 | |
| | | **C.B.S-1** | Transformer-Based Language Models | T1,T2 | 1 | |
| | | | | **Total** | **13** | |

| | | | | | | |
|---|---|---|---|---|---|---|
| II | CO2: Study statistical models for word prediction and part-of-speech tagging techniques. | **2. WORD LEVEL ANALYSIS** | | | | |
| | | 2.1 | Unsmoothed N-grams | T1, T2 | 1 | Chalk, Talk, & PPT |
| | | 2.2 | Evaluating N-gram | T1,T3, R3 | 1 | |
| | | 2.3 | Smoothing | T1,R3 | 1 | |
| | | 2.4 | Interpolation and Backoff | T1,T2 | 1 | |
| | | 2.5 | Word Classes | T1,T3 | 1 | |
| | | 2.6 | Part- of-Speech Tagging | T3,R1 | 1 | |
| | | 2.7 | Rule-based | T1,T2 | 1 | |
| | | 2.8 | Stochastic and Transformation-based tagging | T2,T3,R1 | 1 | |
| | | 2.9 | Issues in PoS tagging | T1,R1,R2 | 1 | |
| | | 2.10 | Hidden Markov and Maximum Entropy models | T2,T3, R3 | 1 | |
| | | | Assignment 2 | T1,T3,R3 | 1 | |
| | | **C.B.S-2** | Large Language Models (LLMs) and Prompt Engineering | T1,T2 | 1 | |
| | | | | **Total** | **13** | |
| III | CO3: Understand syntactic parsing techniques using grammar rules and probabilistic models. | **3. SYNTACTIC ANALYSIS** | | | | |
| | | 3.1 | Context-Free Grammars | R2,R3 | 1 | Chalk, Talk & PPT |
| | | 3.2 | Grammar rules for English | R1,R2,R3 | 1 | |
| | | 3.3 | Treebanks | T1,R1,R3 | 1 | |
| | | 3.4 | Normal Forms for grammar | T1,T3,R3 | 1 | |
| | | 3.5 | Dependency Grammar | T2,T3,R1 | 1 | |
| | | 3.6 | Syntactic Parsing | T1,T3,R3 | 1 | |
| | | 3.7 | Ambiguity | T1,T2 | 1 | |
| | | 3.8 | Dynamic Programming parsing | T2,T3 | 1 | |
| | | 3.9 | Shallow parsing | T1,R3 | 1 | |
| | | 3.10 | Probabilistic CFG | T1, R3 | 1 | |
| | | 3.11 | Probabilistic CYK | T3,R1,R2 | 1 | |
| | | 3.12 | Probabilistic Lexicalized CFGs | T3,R1 | 1 | |
| | | 3.13 | Feature structures, Unification of feature structures | T1,T2 | 1 | |
| | | | Assignment 3 | T1,T3,R3 | 1 | |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | C.B.S-3 | Word Embeddings & Sentence Embeddings | T1, T3,R1 | 1 | |
| | | | | Total | 15 | |
| IV | CO4: Explore meaning representation, semantic roles, and word sense disambiguation techniques. | | 4. SEMANTICS AND PRAGMATICS | | | Chalk, Talk, & Tutorials, |
| | | 4.1 | Requirements for representation, | T1,T3,R1 | 1 | |
| | | 4.2 | First-Order Logic | T1,T2, | 1 | |
| | | 4.3 | Description Logics | T1, ,T3, | 1 | |
| | | 4.4 | Syntax-Driven Semantic analysis | R2,R3 | 1 | |
| | | 4.5 | Semantic attachments | T1, R3 | 1 | |
| | | 4.6 | Word Senses | T1,R2 | 1 | |
| | | 4.7 | Relations between Senses | T1,T2,R3 | 1 | |
| | | 4.8 | Thematic Roles | T1, T3, 2 | 1 | |
| | | 4.9 | selectional restrictions | T1,T3,R1 | 1 | |
| | | 4.10 | Word Sense Disambiguation | T3,R1 | 1 | |
| | | 4.11 | WSD using Supervised | T1,T3,R1 | 1 | |
| | | 4.12 | Dictionary & Thesaurus | T1,T3,R1 | 1 | |
| | | 4.13 | Bootstrapping methods | T1,T3 | 1 | |
| | | 4.14 | Word Similarity using Thesaurus and Distributional methods | T1,T2 | 1 | |
| | | | Assignment 4 | T1,T2 | 1 | |
| | | C.B.S-4 | Ethics in NLP and Bias in Language Models | T1,R3 | 1 | |
| | | | | Total | 16 | |
| V | CO5: Study discourse structures, reference resolution, and key lexical resources in NLP. | | 5. DISCOURSE ANALYSIS AND LEXICAL RESOURCES | | | Chalk, Talk, & Tutorials |
| | | 5.1 | Discourse segmentation | T1,R3 | 1 | |
| | | 5.2 | Coherence | T1, R1 | 1 | |
| | | 5.3 | Reference Phenomena | T1,T3, R3 | 1 | |
| | | 5.4 | Anaphora Resolution using Hobbs and Centering Algorithm | T1,T2, R3 | 1 | |
| | | 5.5 | Coreference Resolution | T2,T3, R3 | 1 | |
| | | 5.6 | Resources: Porter Stemmer | T2,T3, R3 | 1 | |
| | | 5.7 | Lemmatizer | T1,T2 | 1 | |
| | | 5.8 | Penn Treebank | R1, R3 | 1 | |
| | | 5.9 | Brill's Tagger, WordNet | R2,R3 | 1 | |
| | | 5.10 | PropBank, FrameNet | T1,T2 | 1 | |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | 5.11 | Brown Corpus, British National Corpus (BNC) | T1, R3 | 1 | |
| | | | Assignment 5 | | 1 | |
| | | | | **Total** | **12** | |
| | | | | **Cumulative Proposed Periods** | **69** | |
| Where : **C.B.S** = Content Beyond the Syllabus | | | | | | |

| **Text Books:** | |
|---|---|
| **S.No** | **Authors, Book Title, Edition, Publisher, Year of Publication** |
| T1 | Daniel Jurafsky, James H. Martin,Speech and Language Processing (3rd Edition Draft), Pearson, 2023 |
| T2 | Christopher D. Manning, Hinrich Schütze,Foundations of Statistical Natural Language Processing, MIT Press, 1999 |
| T3 | Steven Bird, Ewan Klein, Edward Loper,Natural Language Processing with Python, O'Reilly Media, 2009 |
| **Reference Books:** | |
| **S.No.** | **Authors, Book Title, Edition, Publisher, Year of Publication** |
| R1 | Frederick Jelinek,Statistical Methods for Speech Recognition,MIT Press, 1997 |
| R2 | Nitin Indurkhya, Fred J. Damerau, Chapman and Hall,Handbook of Natural Language Processing 2nd Edition,CRC, 2010 |
| R3 | Jacob Eisenstein, Introduction to Natural Language Processing, MIT Press, 2019. |
| **Web Details** | |
| **W1** | https://www.nltk.org |
| **W2** | https://huggingface.co/learn/nlp-course |
| **W3** | https://www.geeksforgeeks.org/natural-language-processing-overview |

| S.NO. | Details | Name | Signature |
|---|---|---|---|
| i. | Faculty | Mr. V DURGA RAO | |
| ii. | Course Coordinator | Mr . V DURGA RAO | |
| iii. | Module Coordinator | Dr. G. SUDHAKAR | |
| iv. | Program Coordinator | Dr. B. RAMA KRISHNA | |

**Principal**