Course No.	Type	Subject	L	T	P	Credits	CA	MS	ES	CA	ES	Pre-requisites
COCSE27	**ED**	**Natural Language Processing**	**3**	**1**	**0**	**4**						**CACSC11 AI**

COURSE OUTCOMES

1. Realize the fundamentals of natural language processing.
2. Recognize the basics of Language modelling and Word level analysis.
3. Understand the use of CFG and PCFG in NLP.
4. Recognize the role of semantics of sentences and pragmatic.
5. Identify the recent trends and applications of NLP.

COURSE CONTENTS:

UNIT I

Introduction: Various stages of NLP, Why NLP is hard, Why NLP is useful, Classical problems, Introduction to basic language processing – tokens, sentences, paragraphs, Spelling Correction, Morphological analysis and generation using Finite State Automata and Finite State transducer.

UNIT II

Language Modeling (LM): Grammar-based LM, Statistical LM, Linguistics Fundamentals, Classical approaches to NLP with knowledge bases and linguistic rules; Data Driven and Machine Learning Approaches to NLP, Text classification evaluation, relation extraction
Word Level Analysis: Unsmoothed N-grams, Evaluating N-grams, Smoothing, Interpolation and Backoff – Word Classes, Part-of-Speech Tagging, Rule-based, Stochastic and Transformation-based tagging, Issues in PoS tagging – Hidden Markov and Maximum Entropy models.

UNIT III

Syntactic Analysis: Context Free Grammars, Grammar rules for English, Treebanks, Normal Forms for grammar – Dependency Grammar – Syntactic Parsing, Ambiguity, Dynamic Programming parsing – Shallow parsing – Probabilistic CFG, Probabilistic CYK, Probabilistic Lexicalized CFGs – Feature structures, Unification of feature structures.

UNIT IV

Semantics and Pragmatics: Requirements for representation, First-Order Logic, Description Logics – Syntax-Driven Semantic analysis, Semantic attachments – Word Senses, Relations

between Senses, Thematic Roles, selection restrictions – Word Sense Disambiguation, WSD using Supervised, Dictionary & Thesaurus, Bootstrapping methods – Word Similarity using Thesaurus and Distributional methods

UNIT V

Applications of NLP: Speech Processing, Speech Analysis and Modelling, Machine Translation, Information Retrieval, Text Summarization, Text Classification, Sentiment analysis and opinion mining.

SUGGESTED READINGS

1. Daniel Jurafsky, James H. Martin—Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech, Pearson Publication, 2014.
2. Christopher D. Manning and HinrichSchutze, "Foundations of Natural Language Processing", 6 th Edition, The MIT Press Cambridge, Massachusetts London, England, 2003.
3. Steven Bird, Ewan Klein and Edward Loper, —Natural Language Processing with Python, First Edition, OReilly Media, 2009.
4. Hobson lane, Cole Howard, Hannes Hapke, "Natural language processing in action" MANNING Publications, 2019.