Code No.: **MDS 555**
Paper: **Natural Language Processing** Full Marks: 75
Nature: Theory +Practical (Elective) Credit: 3

Course Description:

The course covers the introductions, methods and approaches used in many real-world NLP applications such as Computational Linguistics, Morphology, Syntax, Semantics, Discourse.

Course Objectives:

After successful completion of this course, the student will be able to

- Provide the students a general overview of the basics as well as the advanced concepts of Natural Language Processing (NLP)
- Apply the different concepts of NLP both theoretically and practically.

Course Contents:

Unit 1: Introduction to NLP **[4 Hrs.]**

Introduction to NLP, Origins and importance of NLP, Challenges in NLP (Difficulties, Ambiguities and Evolution), Language and Knowledge (Syntax, Semantics, Pragmatics and Discourse), A Multi-disciplinary field (Psychology, Information Retrieval), Applications of NLP.

Unit 2 : Words and Morphology **[7Hrs.]**

Finite State Machines (FSM) and Morphology, Introduction to FSM and FST, Morphological Processes, Principles of Word Construction (Suffix, Prefix, Stem, Affixes), Morphological Representation and FSM, Lexicon, Morphotactic and Orthographic rules, Morphological Parsing and FST, Mealy machines, FST operations.

Unit 3: Part of Speech Tagging **[7 Hrs.]**

Parts of Speech (PoS) Tagging and Hidden Markov Models (HMM), PoSTagsets, Rule-based PoS Tagging, Stochastic PoS Tagging, Transformation based tagging.

Unit 4: Syntax **[9Hrs.]**

Syntactic Analysis, Context Free Grammar (CFG) & Probabilistic CFG, Word's Constituency (Phrase level, Sentence level), Parsing (Top-Down and Bottom-Up), CYK Parser, Probabilistic Parsing.

Unit 5: Lexical Semantics **[7 Hrs.]**

Lexical Semantics, Lexeme, Lexicon, Senses, Lexical relations, WordNet (Lexical Database), Word Sense Disambiguation (WSD), Word Similarity.

Unit 6: Discourse **[7Hrs.]**

Pragmatic & Discourse Analysis, Monologue and Dialogue, Reference Resolution, Coherence and Cohesion, Discourse Structure.

Unit 7: Applications of NLP **[7Hrs.]**

Applications of NLP, Question Answering, Machine Translation, Sentiment Analysis, Summary Generation.

Lab and Practical Works:

In the lab and practical works, the students will basically get practical concepts of NLP in the Python Programming Language . A lot of these would be hands-on exercises and writing the codes of NLP problem- solving.

References:

1. Daniel Jurafsky and James H. Martin (2009). *Speech and Language Processing* , Second Edition, Pearson Education.
2. Stephen Bird, Ewan Klein& Edward Loper (2009). *Natural Language Processing with Python*. O'Reilly Media, http://www.nltk.org/book/

***