DS743PE: NATURAL LANGUAGE PROCESSING (Professional Elective – IV)

B.Tech. IV Year I Sem.   L T P C
                                                                                                    3 0 0 3

Prerequisites:
- Data structures and compiler design

Course Objectives:
- Introduction to some of the problems and solutions of NLP and their relation to linguistics and statistics.

Course Outcomes:
- Show sensitivity to linguistic phenomena and an ability to model them with formal grammars.
- Understand and carry out proper experimental methodology for training and evaluating empirical NLP systems
- Able to manipulate probabilities, construct statistical models over strings and trees, and estimate parameters using supervised and unsupervised training methods.
- Able to design, implement, and analyze NLP algorithms; and design different language modeling Techniques.

UNIT - I
Finding the Structure of Words: Words and Their Components, Issues and Challenges, Morphological Models
Finding the Structure of Documents: Introduction, Methods, Complexity of the Approaches, Performances of the Approaches, Features

UNIT - II
Syntax I: Parsing Natural Language, Treebanks: A Data-Driven Approach to Syntax, Representation of Syntactic Structure, Parsing Algorithms

UNIT – III
Syntax II: Models for Ambiguity Resolution in Parsing, Multilingual Issues
Semantic Parsing I: Introduction, Semantic Interpretation, System Paradigms, Word Sense

UNIT - IV
Semantic Parsing II: Predicate-Argument Structure, Meaning Representation Systems

UNIT - V
Language Modeling: Introduction, N-Gram Models, Language Model Evaluation, Bayesian parameter estimation, Language Model Adaptation, Language Models- class based, variable length, Bayesian topic based, Multilingual and Cross Lingual Language Modeling

TEXT BOOKS:
1. Multilingual natural Language Processing Applications: From Theory to Practice – Daniel M. Bikel and Imed Zitouni, Pearson Publication

REFERENCE BOOK:
1. Speech and Natural Language Processing - Daniel Jurafsky& James H Martin, Pearson Publications.
2. Natural Language Processing and Information Retrieval: Tanvier Siddiqui, U.S. Tiwary.