# 19CS42E4 - NATURAL LANGUAGE PROCESSING

|  |  |  |  |
| --- | --- | --- | --- |
| **Course Category:** | Program Elective | **Credits:** | 3 |
| **Course Type:** | Theory | **Lecture – Tutorial – Practical:** | 3-0-0 |
| **Prerequisite:** | Data structures, finite automata and probability theory Course | **Sessional Evaluation:**  **Univ. Exam Evaluation:**  **Total Marks:** | 40  60  100 |
| **Objectives** | Introduce to some of the problems and solutions of NLP and their relation to linguistics and statistics. | | |

|  |  |  |
| --- | --- | --- |
| **Course Outcomes** | Upon the successful completion of the course, the students will be: | |
| CO1 | Able to show sensitivity to linguistic phenomena and an ability to model them with formal grammars. |
| CO2 | Able to understand the mathematical and linguistic concepts of NLP |
| CO3 | Able to understand and carry out proper experimental methodology for training and evaluating empirical NLP systems. |
| CO4 | Able to manipulate probabilities, construct statistical models over strings and trees, and estimate parameters using supervised and unsupervised training methods |
| CO5 | Able to design, implement, and analyze NLP algorithms |
| CO6 | Able to design different language modelling Techniques |
| **Course Content** | 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.  UNIT-II  **Syntax Analysis:** Parsing Natural Language, Treebanks: A Data-Driven Approach to Syntax, Representation of Syntactic Structure, Parsing Algorithms, Models for Ambiguity Resolution in Parsing, Multilingual Issues  UNIT-III  **Semantic Parsing:** Introduction, Semantic Interpretation, System Paradigms, Word Sense Systems, Software, Predicate-Argument Structure, Meaning Representation Systems, Software.  UNIT-IV  **Structure Language Modeling:** Introduction, N-Gram Models, Language Model Evaluation, Parameter Estimation, Language Model Adaptation, Types of Language Models, Language-Specific Modeling Problems, Multilingual and Cross lingual Language Modeling  UNIT-V  **Recognizing Textual Entailment:** Introduction, The recognizing textual entailment task, The framework for recognizing textual entailment, Case Studies.  UNIT-VI  **Multilingual Sentiment and Subjectivity Analysis:** Introduction, definitions, sentiment and subjectivity analysis on English, Word and Phrase-Level Annotations, Sentence-Level Annotations, Document-Level Annotations. | |
| **Text Books and References** | **TEXT BOOK(S):**  1.Multilingual Natural Language Processing Applications : From Theory To Practice-Daniel M.Bikel and ImedZitouni , Pearson Publications.  **REFERENCE BOOKS:**  1. Charniack, Eugene, Statistical Language Learning, MIT Press, 1993.  2. Jurafsky, Dan and Martin, James, Speech and Language Processing, 2nd Edition, Prentice Hall, 2008.  3.Manning, Christopher and Henrich, Schutze, Foundations of Statistical Natural Language Processing, MIT Press, 1999. | |
| **E-Resources** | 1. **https://onlinecourses.nptel.ac.in/noc18\_cs13/preview** 2. **https://books.google.co.in/books?isbn=8126510757** | |