*Subject Code : **IT425** Course Title: **Natural Language Processing***
2. Contact Hours : L: 3 T: 1 P: 0
3. Examination Duration (ETE) (Hrs.) : Theory 3 Hrs Practical 0
4. Relative Weightage : CWS 25 PRS 0 MTE 25 ETE 50 PR - 0
5. Credits : 4
6. Semester : VII
7. Subject Area : DEC
8. Pre-requisite : Theory of Automata
9. Objective : The goal of natural language processing (NLP) is to design and build computer systems that are able to analyze natural languages like German or English, and that generate their outputs in a natural language.

10. Details of Course

1. **Introduction:** The study of Language, Introduction to NLP, Regular Expression, Finite State Automata, Evaluating Language Understanding Systems, Different levels of Language Analysis, Representations and Understanding, Linguistic Background. **6 hours**

2. **Grammars and Parsing:** Top-Down and Bottom-Up Parsers, Transition Network Grammars, Top-Down Chart Parsing, Feature Systems and Augmented Grammars, Morphological Analysis and the Lexicon, Parsing with Features, Augmented Transition Networks. **7 hours**

3. **Grammars for Natural Language:** Auxiliary Verbs and Verb Phrases, Movement Phenomenon in Language, Handling questions in Context-Free Grammars, Hold mechanisms in ATNs, Human preferences in Parsing, Encoding uncertainty, Deterministic Parser. **6 hours**

4. **Ambiguity Resolution:** Statistical Methods, Probabilistic Language Processing, Estimating Probabilities, Part-of-Speech tagging, Obtaining Lexical Probabilities, Probabilistic Context-Free Grammars, Dependency Parsing, Best First Parsing, Semantics and Logical Form, Word senses and Ambiguity, Encoding Ambiguity in Logical Form**.** **7 hours**

5. **Advanced Features and Syntax, Features and Unification:** Feature structures – Unification of feature structures – Features structures in the grammar – Implementing unification – Parsing with unification constraints – Types and Inheritance. Lexicalized and Probabilistic Parsing: Probabilistic context-free grammar – problems with PCFGs– Probabilistic lexicalized CFGs – Dependency Grammars – Human parsing. **8 hours**

6. **Application of NLP:** Intelligent Work Processors, Machine Translation, User Interfaces, Man-Machine Interfaces, Natural language Querying Tutoring and Authoring Systems, Speech Recognition Commercial use of NLP, Semantic Interpretation, Information Retrieval. **8 hours**

TOTAL 42 hours

Text Books

1. James Allen, Natural Language Understanding, 2/e, Pearson Education (ISBN 13: 9788131708958), 2003
2. Foundation of Statistical Natural Language Processing, Manning and Schutze, (**ISBN-13:** 978-0262133609), 1998
3. D. Jurafsky, J. H. Martin, Speech and Language Processing, Pearson Education,(**ISBN-13:** 978-8131716724), 2008

Reference books

1. Bharati, Chaitanya and Sangal: Natural Language Processing- a Paninian perspective(**ISBN-13:** 978-8120309210), 1995
2. Leonard Bolc. (Ed.): Natural Language Parsing Systems, Springer Verlag, (**ISBN-13:** 978-0387175379) 1986