NATURAL LANGUAGE PROCESSING

B Tech III Year II Sem

Course Code	Category	Hours/ Week			Credits	Maximum Marks		
		L	T	P		CIE	SEE	TOTAL
23AM603	Professional Course	3	0	0	3	40	60	100
Contact Classes: 45	Tutorial Classes: 15	Practical Classes: Nil				Total Classes:60		

Pre-requisites:

1. Data structures and compiler design.

Course Objectives:

1. Introduction to some of the problems and solutions of NLP and their relation to linguistics and statistics.

Course Outcomes:

1. Show sensitivity to linguistic phenomena and an ability to model them with formal grammars.
2. Understand and carry out proper experimental methodology for training and evaluating empirical NLP systems
3. Manipulate probabilities, construct statistical models over strings and trees, and estimate parameters using supervised and unsupervised training methods.
4. 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 BOOK:

1. Multilingual natural Language Processing Applications: From Theory to Practice - Daniel M. Sikel and lmed Zitouni, Pearson Publication.

REFERENCE BOOKS:

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.

NATURAL LANGUAGE PROCESSING LAB

B Tech III Year II Sem

Course Code	Category	Hours/ Week			Credits	Maximum Marks		
		L	T	P		CIE	SEE	TOTAL
23AM604	Professional Course	0	0	3	1.5	40	60	100
Contact Classes: Nil	Tutorial Classes: Nil	Practical Classes: Nil				Total Classes:32		

Prerequisites:

1. Data structures, finite automata and probability theory.

Course Objectives:

1. To Develop and explore the problems and solutions of NLP.

Course Outcomes: Students will be able to:

1. Show sensitivity to linguistic phenomena and an ability to model them with formal grammars.
2. Knowledge on NLTK Library implementation
3. Work on strings and trees, and estimate parameters using supervised and unsupervised training methods.

Exercises:

1. Write a Python Program to perform following tasks on text.
    - Tokenization
    - Stop word Removal
2. Write a Python program to implement Porter stemmer algorithm for stemming
3. Write Python Program for a) Word Analysis b) Word Generation
4. Create a Sample list for at least 5 words with ambiguous sense and Write a Python program to implement WSD
5. Install NLTK tool kit and perform stemming
6. Create Sample list of at least 10 words POS tagging and find the POS for any given word
7. Write a Python program to
    - Perform Morphological Analysis using NLTK library
    - Generate n-grams using NLTK N-Grams library
    - Implement N-Grams Smoothing
8. Using NLTK package to convert audio file to text and text file to audio files.

TEXT BOOKS:

2. Multilingual natural Language Processing Applications: From Theory to Practice - Daniel M. Sikel and lmed Zitouni, Pearson Publication.
3. Oreilly Practical natural Language Processing, A Comprehensive Guide to Building Real World NLP Systems.
4. Daniel Jurafsky, James H. Martin-Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech, Pearson Publication, 2014.

REFERENCES:

1. Steven Bird, Ewan Klein and Edward Loper, -Natural Language Processing with Python, First Edition, O'Reilly Media, 2009.