NATURAL LANGUAGE PROCESSING

(Professional Elective-III - Online)

Course Code : 15IT11M2

L	T	P	C
3	0	0	3

Pre-requisites:

Data Structures, Artificial Intelligence, Introduction to Machine Learning.

Course Outcomes:

At the end of the Course, the Student will be able to:

CO 1 Understand the basic concepts of NLP

CO 2 Understand Indian Language Processing techniques

CO 3 Understand Hidden Markov Models and forward backward algorithms

CO 4 Understand the basic IR models

CO 5 Understand the techniques of Word Sense Disambiguation

UNIT-I: **(10 Lectures)**

Introduction,Stages of NLP, Stages of NLP, Two approaches to NLP, Sequence Labeling and Noisy Channel, Noisy Channel: Argmax Based Computation, Noisy Channel Applications to NLP.

UNIT-II: **(11 Lectures)**

Brief on Probabilistic Parsing & Start of Part of Speech Tagging, Part of Speech Tagging, Part of Speech Tagging counted & Indian Language in Focus; Morphology Analysis,PoS Tagging Indian Language Consideration; Accuracy Measure, PoS Tagging; Fundamental Principle; Why Challenging; accuracy, PoS Tagging; Accuracy Measurement; Word categories,AI and Probability: HMM

UNIT-III: (10 Lectures)

HMM, HMM:

Viterbi, Forward Backward Algorithm, Forward Backward Algorithm,HMM: Forward Backward Algorithms, Baum Welch Algorithm, Natural Language Processing and Informational Retrieval,CLIA; IR Basics, IR Models: Boolean Vector

UNIT-IV: (9 Lectures)

IR MODELS:

NLP and IR Relationship, NLP and IR: How NLP has used IR, Toward Latent Semantic,Least Square Method; Recap of PCA; Towards Latent Semantic Indexing(LSI), PCA; SVD; Towards Latent Semantic Indexing(LSI),Wordnet and Word Sense Disambiguation,Wordnet; Metonymy and Word Sense Disambiguation,Word Sense Disambiguation

UNIT-V: (10 Lectures)

Word Sense Disambiguation; Overlap Based Method; Supervised Method,Word Sense Disambiguation: Supervised and Unsupervised methods, Word Sense Disambiguation: Semi - Supervised and Unsupervised method, Resource Constrained WSD; Parsing, Parsing, Parsing Algorithm, Parsing Ambiguous Sentences; Probabilistic Parsing

REFERENCES:

1. Allen, James, “Natural Language Understanding”, 2nd Edition, Benjamin/Cumming, 1995.
2. Charniack, Eugene, “Statistical Language Learning”, MIT Press, 1993.
3. Jurafsky, Dan and Martin, James, “Speech and Language Processing”, 2nd Edition, Prentice Hall, 2008.
4. Manning, Christopher and Heinrich, Schutze, “Foundations of Statistical Natural Language Processing”, MIT Press, 1999.

WEB REFERENCES:

1. http://www.nptelvideos.in/2012/11/natural-language-processing.html