Course III

GROUP (F) : NATURAL LANGAUGE PROCESSING

Total Marks: 100 (End-Semester: 60 & In-Semester: 40)

INTRODUCTION:

Natural Language Processing (NLP) is a branch of artificial intelligence that uses natural language to handle interactions between humans and computers. The ultimate goal of NLP is to read, understand, and decode human language in a valuable way. Most NLP technologies rely on machine learning to extract meaning from human language.

COURSE OBJECTIVE:

1. To learn the basics of Natural Language Processing.
2. To apply the techniques of Natural Language Processing.
3. To understand the role of sentence semantics and pragmatics.

LEARNING OUTCOME:

1. Students will be able to understand Natural Language Processing.
2. Analyze semantics and discourse differentiation from an NLP perspective.
3. Application of a probabilistic model of defining language and techniques.

DETAILED SYLLABUS:

Unit I:
Marks: 20

Words – Structure (spellcheck, morphology using FSTs) – Words - Semantics (Basic ideas in Lexical Semantics, WordNet and WordNet based similarity measures, Distributional measures of similarity, Concept Mining using Latent Semantic Analysis) – Words - Semantics (Word Sense Disambiguation; supervised, unsupervised and semi-supervised approaches) – Words - Parts of Speech Tagging

Unit II:
Marks: 20

Sentences: Basic ideas in compositional semantics, Classical Parsing (Bottom up, top down, Dynamic Programming: CYK parser) - Sentences: Parsing using Probabilistic Context Free Grammars and EM based approaches for learning PCFG parameters - Language Modelling (basic ideas, smoothing techniques) - Machine Translation (rule based techniques, Statistical Machine Translation (SMT), parameter learning in SMT (IBM models) using EM) - Information Extraction: Introduction to Named Entity Recognition and Relation Extraction

Unit III:
Marks: 20

Natural Language Generation: the potential of using ML for NLG - Additional topics: Advanced Language Modelling (including LDA), other applications like summarization, question answering

REFERENCE BOOKS:

1.   Jurafsky Dan and Martin James H. "Speech and Language Processing",3rd Edition, 2018.

2.   Jurafsky D. and Martin J. H., "Speech and language processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition", 2nd Edition, Upper Saddle River, NJ: Prentice-Hall, 2008.