

UNIT I INTRODUCTION

Natural Language Processing – Linguistic Background- Spoken language input and output Technologies – Written language Input - Mathematical Methods - Statistical Modeling and Classification Finite State methods Grammar for Natural Language Processing – Parsing – Semantic and Logic Form – Ambiguity Resolution – Semantic Interpretation.

UNIT II INFORMATION RETRIEVAL 9

Information Retrieval architecture - Indexing- Storage – Compression Techniques – Retrieval Approaches – Evaluation - Search engines- commercial search engine features- comparison- performance measures – Document Processing - NLP based Information Retrieval – Information Extraction.

UNIT III TEXT MINING 9

Categorization – Extraction based Categorization- Clustering- Hierarchical Clustering- Document Classification and routing- finding and organizing answers from Text search – use of categories and clusters for organising retrieval results – Text Categorization and efficient Summarization using Lexical Chains – Pattern Extraction.

UNIT IV GENERIC ISSUES 9

Multilinguality – Multilingual Information Retrieval and Speech processing - Multimodality – Text and Images – Modality Integration - Transmission and Storage – Speech coding- Evaluation of systems – Human Factors and user Acceptability.

UNIT V APPLICATIONS 9

Machine Translation – Transfer Metaphor - Interlingua and Statistical Approaches - Discourse Processing – Dialog and Conversational Agents – Natural Language Generation – Surface Realization and Discourse Planning.

TOTAL = 45

TEXT BOOKS:

1. Daniel Jurafsky and James H. martin, “ Speech and Language Processing” , 2000.
2. Ron Cole, J.Mariani, et.al “Survey of the State of the Art in Human Language Technology”, Cambridge University Press, 1997.
3. Michael W. Berry “ Survey of Text Mining: Culstering, Classification and Retrieval”, Springer Verlag, 2003.
4. Christopher D.Manning and Hinrich Schutze, “ Foundations of Statistical Natural Language Processing “, MIT Press, 1999.

REFERENCES:

1. James Allen “ Natural Language Understanding “, Benjamin/ Cummings Publishing Co. 1995.
2. Gerald J. Kowalski and Mark.T. Maybury, “Information Storage and Retrieval systems”, Kluwer academic Publishers, 2000.
3. Tomek Strzalkowski “ Natural Language Information Retrieval “, Kluwer academic Publishers, 1999.
4. Christopher D.Manning and Hinrich Schutze, “ Foundations of Statistical Natural Language Processing “, MIT Press, 1999.