MANGALORE UNIVERSITY

DEPARTMENT OF COMPUTER SCIENCE

CSS 457: NATURAL LANGUAGE PROCESSING		
Hours/Week: 4 **Credits : 4**		**I.A. Marks: 30** **Exams. Marks: 70**

Course Outcomes:

CO1: After successful completion of this course, student will be able to

CO2: Understand approaches to syntax and semantics in NLP.

CO3: Understand approaches to discourse, generation, dialogue and summarization within NLP.

CO4: Understand current methods for statistical approaches to machine translation.

CO5: Understand machine learning techniques used in NLP, including hidden Markov models and probabilistic context-free grammars, clustering and unsupervised methods, log-linear and discriminative models, and the EM algorithm as applied within NLP

	UNIT-I	**12 Hrs.**

OVERVIEW AND LANGUAGE MODELING : *Overview*: Origins and challenges of NLP-Language and Grammar-Processing Indian Languages- NLP Applications-Information Retrieval. Language Modeling: Various Grammar- based Language Models-Statistical Language Model.

	UNIT-II	**12 Hrs.**

WORD LEVEL AND SYNTACTIC ANALYSIS: Word Level Analysis: Regular Expressions-Finite-State Automata-Morphological Parsing-Spelling Error Detection and correction-Words and Word classes-Part-of Speech Tagging. Syntactic Analysis: Context-free Grammar-Constituency-Parsing-Probabilistic Parsing.

	UNIT-III	**12 Hrs.**

SEMANTIC ANALYSIS AND DISCOURSE PROCESSING : Semantic Analysis: Meaning Representation-Lexical Semantics- Ambiguity-Word Sense Disambiguation. Discourse Processing: cohesion-Reference Resolution- Discourse Coherence and Structure.

	UNIT-IV	12 Hrs.

NATURAL LANGUAGE GENERATION AND MACHINE TRANSLATION :
Natural Language Generation: Architecture of NLG Systems- Generation Tasks and Representations- Application of NLG. Machine Translation: Problems in Machine Translation- Characteristics of Indian Languages- Machine Translation Approaches-Translation involving Indian Languages.

REFERENCE BOOKS:

1. Edward Loper, Ewan Klein, and Steven Bird, Natural Language Processing with Python, O'Reilly Publication 2009.;
2. Christopher D. Manning, Hinrich Schütze, Foundations of Statistical Natural Language Processing, MIT press,1999.
3. Dan Jurafsky, James H. Martin, Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition, Prentice Hall, 2009.