



BONAM VENKATA CHALAMAYYA INSTITUTE OF TECHNOLOGY & SCIENCE (A)
DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
BR25 M.TECH CSE COURSE STRUCTURE AND SYLLABUS

II Semester	NATURAL LANGUAGE PROCESSING (D252BA02)	L	T	P	C
		3	1	0	4

Course Objectives: This course introduces the fundamental concepts and techniques of natural language processing (NLP).

- Students will gain an in-depth understanding of the computation alproperties of natural languages and the commonly used algorithms for processing linguistic information.
- The course examines NLP models andalgorithms usingboth the traditional symbolic and the more recent statistical approaches.
- Enable students to becapable to describe the application based on natural language processing and to show the points of syntactic, semantic and pragmatic processing.

Course Outcomes: At the end of the course, student will be able to (Four to Six)

		Knowledge Level (K)#
CO1	Demonstrate a given text with basic Language features	K5
CO2	Design an innovative application using NLP components	K3
CO3	Explaina rule based system to tack lemorphology/syntax of alanguage	K3
CO4	Design at agset to be used forstatistical processing forreal-time applications	K3
CO5	Compareand contrast the use of different statistical approaches for Different types of NLP applications	K5

Based on suggested Revised BTL

Mapping of course outcomes with program outcomes

	PO1	PO2	PO3	PO4	PO5	PO6
CO1	2		2	3		
CO2	3	2	3	3	3	3
CO3	2		3	3		
CO4	2		3	3	2	
CO5	2	2	3	3	3	2

Dr.N.Rama Krishnaiah, Professor of CSE,UCEK & Control of Examination JNTUK, kakinada.	Dr.C.Krishna Mohan, Professor of CSE,IIT, Kandi, Hyderabad.	Dr.P.Radha Krishna, Professor of CWQSE,NI T, Warangal	Mr.Rajesh Bobburi Chief Operating Officer, HighQ Labs Private Limited, Rajahmundry	Dr.Lakshmi Haritha Medida, Associate Professor, R.M.K.Engin eering College,Kava raipettai,Tam ilnadu	Dr.K.Srinivas, Professor & HoD Department of CSE, B.V.C.I.T.S, Batlapalem
---	---	---	--	--	---



BONAM VENKATA CHALAMAYYA INSTITUTE OF TECHNOLOGY & SCIENCE (A)
DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
BR25 M.TECH CSE COURSE STRUCTURE AND SYLLABUS

(Please fill the above with Levels of Correlation, viz., L, M, H)

UNIT	CONTENTS	Contact Hours
UNIT-1	INTRODUCTION: Origins and challenges of NLP–Language Modeling:Grammar-based LM, Statistical LM–Regular Expressions, Finite-StateAutomata –English Morphology,Transducers for lexiconand rules, Tokenization, Detecting and Correcting Spelling Errors, Minimum Edit Distance	10Hrs
UNIT-2	WORD LEVEL ANALYSIS: Unsmoothed N-grams, Evaluating N-grams, Smoothing, Interpolation and Backoff – Word Classes, Part- of-Speech Tagging, Rule-based, Stochastic and Transformation-based tagging, Issues in PoStagging–Hidden Mark ovand Maximum Entropy models	12Hrs
UNIT-3	SYNTACTIC ANALYSIS: Context-Free Grammars, Grammarrulesfor English, Treebanks, Normal Forms for grammar– Dependency Grammar –Syntactic Parsing, Ambiguity, Dynamic Programming parsing–Shallow parsing Probabilistic CFG, Probabilistic CYK, Probabilistic Lexicalized CFGs – Feature structures, Unification of feature structures	12Hrs

Dr.N.Rama Krishnaiah, Professor of CSE,UCEK & Control of Examination JNTUK, Kakinada.	Dr.C.Krishna Mohan, Professor of CSE,IIT, Kandi, Hyderabad.	Dr.P.Radha Krishna, Professor of CWQSE,NI T, Warangal	Mr.Rajesh Bobburi Chief Operating Officer, HighQ Labs Private Limited, Rajahmundry	Dr.Lakshmi Haritha Medida, Associate Professor, R.M.K.Engin eering College,Kava raipettai,Tam ilnadu	Dr.K.Srinivas, Professor & HoD Department of CSE, B.V.C.I.T.S, Batlapalem
---	---	---	--	--	---



BONAM VENKATA CHALAMAYYA INSTITUTE OF TECHNOLOGY & SCIENCE (A)
DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
BR25 M.TECH CSE COURSE STRUCTURE AND SYLLABUS

UNIT-4	SEMANTICS AND PRAGMATICS: Requirements for representation, First-Order Logic, Description Logics–Syntax-Driven Semantic analysis, Semantic attachments – Word Senses, Relations between Senses, Thematic Roles, selectional restrictions – Word Sense Disambiguation, WSD using Supervised, Dictionary & Thesaurus, Bootstrapping methods – Word Similarity using Thesaurus and Distributional methods.	12Hrs
UNIT-5	DISCOURSE ANALYSIS AND LEXICAL RESOURCES: Discourse segmentation, Coherence – Reference Phenomena, Anaphora Resolution using Hobbs and Centering Algorithm – Coreference Resolution – Resources: Porter Stemmer, Lemmatizer, Penn Treebank, Brill's Tagger, WordNet, PropBank, FrameNet, Brown Corpus, British National Corpus (BNC)	12Hrs
Total		58Hrs

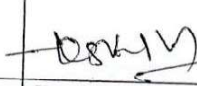
*Note:

Text Books:

1. Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech, 2nd Edition, Daniel Jurafsky, James H. Martin - Pearson Publication, 2014.
2. Natural Language Processing with Python, First Edition, Steven Bird, Ewan Klein and Edward Loper, O'Reilly Media, 2009.

Reference Books:

1. Language Processing with Java and Ling Pipe Cookbook, 1st Edition, Breck Baldwin, Atlantic Publisher, 2015.
2. Natural Language Processing with Java, 2nd Edition, Richard M Reese, O'Reilly Media, 2015.
3. Hand book of Natural Language Processing, Second, Nitin Indurkha and Fred J. Damerau, Chapman and Hall/CRC Press, 2010. Edition Natural Language Processing and Information Retrieval, 3rd Edition, Tanveer Siddiqui, U.S. Tiwary, Oxford University Press, 2008.

Dr.N.Rama Krishnaiah, Professor of CSE, UCEK & Control of Examination JNTUK, Kakinada.	Dr.C.Krishna Mohan, Professor of CSE, IIT, Kandi, Hyderabad.	Dr.P.Radha Krishna, Professor of CWQSE, NIT, Warangal	Mr.Rajesh Bobburi Chief Operating Officer, HighQ Labs Private Limited, Rajahmundry	Dr.Lakshmi Haritha Medida, Associate Professor, R.M.K.Engineering College, Kavaraitpettai, Tamilnadu	 Dr.K.Srinivas, Professor & HoD Department of CSE, B.V.C.I.T.S, Batlapalem
--	--	---	--	--	--