	National Institute of Technology Meghalaya An Institute of National Importance	**CURRICULUM**

Programme	**Bachelor of Technology in Computer Science and Engineering**	Academic Year of Regulation	**2018-19**
Department	**Computer Science and Engineering**	Semester	**VIII**

Course Code	Course Name	Credit Structure				Marks Distribution			
		L	T	P	C	INT	MID	END	Total
CS418	**Natural Language Processing**	**3**	**0**	**0**	**3**	**50**	**50**	**100**	**200**

Course Objectives	Course Outcomes	
This course introduces foundational linguistic and mathematical concepts and algorithms for analysis of natural languages.	CO1	**Able to choose techniques for basic linguistic processing for phonetic analysis, phonological analysis and morphological analysis.**
This course introduces the advantages and disadvantages of different NLP technologies in different real-life applications.	CO2	**Able to construct computational models of natural language text data in order to gain broader understanding of text data.**
This course familiarizes some statistical approaches and machine learning techniques used in Natural Language Processing (NLP) tasks.	CO3	**Able to solve common NLP tasks using models, methods, and algorithms for statistical NLP.**
	CO4	**Able to create software implementations of relevant pre-processing steps for different NLP problems.**
	CO5	**Able to solve common NLP tasks using machine learning algorithms.**

No.	COs	Mapping with Program Outcomes (POs)												Mapping with PSOs		
		PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	PSO3
1	CO1	3	1	1	0	0	0	1	0	0	0	0	1	0	0	0
2	CO2	3	3	3	2	2	0	0	0	0	0	0	0	1	1	1
3	CO3	3	3	3	2	2	1	0	0	1	0	1	0	0	1	1
4	CO4	3	3	3	2	1	1	0	0	0	0	0	0	0	2	3
5	CO5	3	3	3	2	2	1	0	0	1	1	1	0	0	1	1

SYLLABUS

No.	Content	Hours	COs
I	**Introduction; Motivation and challenges of Natural Language Processing (NLP); Tokenisation and Sentence Segmentation**	**02**	**CO1**
II	**Lexical Analysis: Morphology, Finite State Morphology**	**03**	**CO1, CO2**
III	**Syntactic Analysis: Linguistic Background - An outline of English Syntax, Grammars for Natural Language, Parsing techniques, Linking Syntax and Semantics; Semantic Analysis: Lexical Semantics, Word Sense Disambiguation; Pragmatics and Discourse Analysis: Dialogue and Conversational agents, Co-reference resolution; Natural Language Generation**	**12**	**CO2, CO3**
IV	**Overview of NLP applications: POS tagging, Information Retrieval, Question Answering, Information Extraction, Dialogue Systems, Text and Intent Mining, Machine Translation; Data pre-processing for NLP tasks**	**11**	**CO3**
V	**Empirical techniques for NLP tasks; machine learning techniques for NLP tasks; NLP application examples in real-life; Performance evaluation metrics for NLP systems**	**08**	**CO4, CO5**
	Total Hours	**36**	

Essential Readings

1. D. Jurafsky and J. H. Martin, “Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition,” Pearson Education India, 2nd edition, 2013.
2. Akshar Bharati, Vineet Chaitanya, Rajeev Sangal, “Natural Language Processing: A Paninian Perspective”, PHI Learning Pvt. Ltd., 1st edition, 1995.
3. Daniel M. Bikel, “Multilingual Natural Language Processing Applications: From Theory to Practice”, Pearson Education India, 1st edition, 2012.
4. C. D. Manning, H. Schütze, “Foundations of Statistical Natural Language Processing”, MIT Press, 1st edition, 1999.

Supplementary Readings

1. Jacob Perkins, “Python 3 Text Processing with NLTK 3 Cookbook”, Packt Publishing Limited, 1st edition, 2014.
2. Breck Baldwin, Krishna Dayanidhi, “Natural Language Processing with Java and LingPipe Cookbook”, Packt Publishing Limited, 1st edition, 2014.
3. Nitin Indurkhya and Fred J. Damerau, “Handbook of Natural Language Processing”, Taylor and Francis, 2nd edition, 2010.