NIRMA UNIVERSITY
Institute of Technology
B Tech Computer Science and Engineering
Semester VI
Department Elective-III

L	T	P	C
2	0	2	3

Course Code	2CSDE70
Course Title	Natural Language Processing

Course Outcomes:

At the end of the course, students will be able to –

1. discuss about major NLP issues and solutions
2. illustrate computational methods to understand language phenomena of word sense
3. design and develop applications with natural language capabilities.

Syllabus:

Teaching Hours:30

Unit I

Introduction and Text Classification: NLP overview, Text pre-processing, feature extraction from text, Neural networks for words and characters, Text Mining case study — **05**

Unit II — **12**

Language Modelling: N gram models, Smoothing, Part of speech tagging, Hidden Markov models, Viterbi algorithm, Forward - backward algorithm, EM training, Models for Named Entity Recognition, Neural Language Models - Recurrent Neural Networks and Long Short term Memory networks

Unit III

Vector Space Models: Matrix factorization, Word2Vec and Doc2Vec, Word - character and sentence embedding, Topic modelling — **04**

Unit IV — **04**

Maximum Entropy Classifiers: The maximum entropy principle, and its relation to maximum likelihood. Maximum entropy classifiers and their application to document classification and sentence segmentation

Unit V — **05**

Sequence to Sequence Modelling: Introducing machine translation, Encoder-decoder architecture, Attention mechanism, implementing a conversational chat-bot

Self-Study:

The self-study contents will be declared at the commencement of semester. Around 10% of the questions will be asked from self-study contents.

Laboratory Work:

Laboratory work will be based on the above syllabus with minimum 10 experiments to be incorporated.

Suggested Readings^:

1. Jurafsky, David, and James H. Martin. Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition, PEARSON
2. Manning, Christopher D., and HinrichSchütze. Foundations of Statistical Natural Language Processing. Cambridge, MA: MIT Press
3. James Allen. Natural Language Understanding. The Benjamin/Cummings Publishing Company Inc..
4. Steven Bird, Ewan Klein, and Edward Loper. Natural Language Processing with Python – Analyzing Text with the Natural Language Toolkit

L=Lecture, T=Tutorial, P=Practical, C=Credit

^this is not an exhaustive list