



<b>Syllabus</b>		
<b>Unit No</b>	<b>Contents</b>	<b>Mapped CO</b>
<b>I</b>	Regular Expressions, Text Normalization, Edit Distance- Regular Expression, Words, Corpora, Text Normalization, Minimum Edit Distance. N-Gram Language Models-NGrams, Evaluating Language Models, Generalization and Zeros, Smoothing, Kneser-Ney Smoothing, The web and stupid Backoff, Advanced Perplexity's Relation to Entropy.	<b>CO1</b> <b>CO4</b>
<b>II</b>	Naive Bayes and Sentiment Classification: Naive Bayes Classifiers , Training the Naive Bayes Classifier , Worked example, Optimizing for Sentiment Analysis ,Naive Bayes for other text classification tasks, Naive Bayes as a Language Model, Evaluation: Precision, Recall, F-measure, Test sets and Cross-validation, Statistical Significance Testing, Avoiding Harms in Classification	<b>CO1</b> <b>CO2</b> <b>CO4</b>
<b>III</b>	Vector Semantics and Embeddings- Lexical Semantics, Vector Semantics, Words and Vectors, Cosine for measuring similarity, TF-IDF: Weighing terms in the vector, Applications of the TF-IDF vector model, Word2vec, Visualizing Embeddings, Semantic properties of embeddings, Bias and Embeddings, Evaluating Vector Models.	<b>CO1</b> <b>CO3</b> <b>CO4</b>
<b>IV</b>	Sequence Labeling for Parts of Speech and Named Entities- English Word Classes, Part-of-Speech Tagging, Named Entities and Named Entity Tagging , HMM Part-of-Speech Tagging, Conditional Random Fields (CRFs), Evaluation of Named Entity Recognition	<b>CO1</b> <b>CO3</b> <b>CO4</b>
<b>V</b>	Applications of NLP- Question Answering Information Retrieval IR-based Factoid Question Answering, Entity Linking, Knowledge-based Question Answering, Using Language Models to do QA , Classic QA Models, Evaluation of Factoid Answers, Chatbots & Dialogue Systems, Properties of Human Conversation, Chatbots, GUS: Simple Frame-based Dialogue Systems, The Dialogue-State Architecture, Evaluating Dialogue Systems, Dialogue System Design	<b>CO1</b> <b>CO2</b> <b>CO3</b> <b>CO4</b>

<b>Learning Resources</b>
<b>TextBooks</b>
<ol style="list-style-type: none"> <li>1. Speech and Language Processing: An introduction to Natural Language Processing, Computational Linguistics and Speech Recognition by Daniel Jurafsky and James HMartin,3<sup>rd</sup> Edition,PrenticeHall,2019.</li> <li>2. Natural Language Processing: An information Access Perspective by Kavi Narayana Murthy, Ess Publications, 2006.</li> </ol>
<b>References</b>
<ol style="list-style-type: none"> <li>1. Applied Text Analysis with Python by Benjamin Bengfort, Tony Ojeda, Rebecca Bilbro, O'Reilly Media, June 2018.</li> <li>2. Natural Language Processing Recipes by Akshay Kulkarni, Adarsha Shivananda, Apress,2019</li> </ol>
<b>E-Resources and other Digital Material</b>
<ol style="list-style-type: none"> <li>1. Natural Language Processing by Pawan Goyal, IIT Kharagpur, <a href="https://swayam.gov.in/nd1_noc19_cs56/preview">https://swayam.gov.in/nd1_noc19_cs56/preview</a></li> <li>2. Natural Language Processing offered by deeplearning.ai on Coursera <a href="https://www.coursera.org/specializations/natural-language-processing">https://www.coursera.org/specializations/natural-language-processing</a></li> </ol>