GANPAT UNIVERSITY
FACULTY OF ENGINEERING & TECHNOLOGY

Programme	Bachelor of Technology			Branch/Spec.	Computer Engineering / Information Technology		
Semester	VII			Version	1.0.0.1		
Effective from Academic Year	2025-26			Effective for the Batch admitted in		July 2022	
Course Code	2CEIT7PE7	Course Name		Natural Language Processing			

Teaching Scheme					Examination Scheme (Marks)			
(Per week)	Lecture (DT)		Practical (Lab.)		Total		CE	SEE	Total
	L	TU	P	TW					
Credit	3	-	1	-	4	Theory	40	60	100
Hours	3	-	2	-	5	Practical	30	20	50

Pre-requisites
Basic knowledge of python programming and machine learning statistics

Course Outcomes
On successful completion of the course, the students will be able to:

CO1	Comprehend and apply text processing techniques such as stemming, tokenization, and dependency parsing, to resolve linguistic ambiguities.
CO2	Evaluate language modelling principles and critique NLP models like LSTM, GRU, BERT, and Transformers.
CO3	Design and implement machine translation systems and question-answering bots using advanced NLP techniques.
CO4	Apply and compare state-of-the-art NLP models for advanced NLP tasks.

Theory Syllabus		
Unit	Content	Hrs.
1	**Introduction to NLP:** Introduction to various levels of natural language processing, Ambiguities and computational challenges in processing various natural languages, Introduction to Real life applications of NLP such as spell and grammar checkers, information extraction, question answering, and machine translation.	02
2	**Text Processing:** Ambiguity in language, Segmentation, Stemming, Tokenization, Representation of word, Sentence, Word embedding, Word Senses, Linguistic Structure: Dependency Parsing.	06
3	**Text Classification:** Word Window Classification, Neural Networks for text, N-gram Language Models, Perplexity, Hidden Markov Models, Viterbi algorithm, Recurrent Neural network, Vanishing Gradients and exploding gradients.	09
4	**Language Modelling:** The role of language models, Estimating parameters and smoothing, Evaluating language models, LSTM (Long sort term memory), GRU (Gated recurrent Unit), Part of speech tagging, BERT, XLnet, 1D-CNN for NLP, Sub-word Models, Contextual Representations, Transformers, Self-Attention for Generative Models.	10
5	**Machine Translation:** Statically Machine Translation, Neural Machine Translation, Seq2Seq Modelling, Attention, Question Answering Bot, Natural Language Generation, Neural Machine Translation.	06
6	**Advancement in NLP Models:** Introduction to Generative AI, OpenAI's GPT, GPT3 & Beyond, Google's ALBERT, ULMFiT, Facebook's RoBERTa, Text Summarization, Extractive, Abstractive Text Summarization, Transformer models for Text Summarization.	10
7	**NLP Case Study:**	02

Practical Content
Practicals, assignments and tutorials are based on the above syllabus.

Text Books

1	Speech and Language processing an introduction to Natural Language Processing by Daniel Jurafsky and James H. Martin, Computational Linguis, Prentice Hall, (Latest Edition)
Reference Books	
1	Natural Language Processing with Python by Steven Bird, Ewan Klein and Edward Lopper, O'Reilly, (Latest Edition)
2	Natural Language Processing and Information Retrieval by Siddiqui and Tiwari, Oxford University Press, (Latest Edition)
3	Handbook of Natural Language Processing by Nitin Indurkhya and Fred J. Damerau, Taylor and Francis, (Latest Edition). ISBN 978-1420085921.
4	Natural Language understanding by Allen J., Pearson, (Latest Edition). ISBN 978-0805303346.
ICT/MOOCs Reference	
1	https://nptel.ac.in/courses/106/105/106105158/
2	https://nptel.ac.in/courses/106/106/106106211/

Mapping of CO with PO and PSO:															
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	PSO3
CO1	3	2	2	1	3	0	0	0	1	1	0	3	3	3	3
CO2	3	2	2	1	3	0	0	0	1	1	0	3	2	3	3
CO3	3	3	3	1	3	2	0	0	1	1	1	3	3	3	3
CO4	3	3	3	1	3	2	0	0	1	1	1	3	3	3	3