SEMESTER S8

NATURAL LANGUAGE PROCESSING

(Common to CS/CA/CD)

Course Code	**PECST862**	**CIE Marks**	40
Teaching Hours/Week (L:T:P: R)	3:0:0:0	**ESE Marks**	60
Credits	3	**Exam Hours**	2 Hrs. 30 Min.
Prerequisites (if any)	None	**Course Type**	Theory

Course Objectives:

1. To provide a comprehensive understanding of natural language processing (NLP) and language models, focusing on the principles and techniques of prompt engineering to effectively guide and optimize AI-driven outputs.
2. practical skills necessary to design, implement, and evaluate prompt engineering strategies across various applications, while considering the ethical implications and challenges associated with AI-generated content.

SYLLABUS

Module No.	Syllabus Description	Contact Hours
1	**Introduction to NLP:** Introduction to Natural Language Processing - Various stages of traditional NLP – Challenges - Basic Text Processing techniques - Common NLP Tasks. N-gram Language Models - Naive Bayes for Text Classification, and Sentiment Analysis – Evaluation-Precision, Recall and F-measure-Test sets and cross validation.	7
2	**Traditional NLP Techniques**: Annotating Linguistic Structures - Context-Free Grammars, Constituency Parsing, Ambiguity, CYK Parsing, Dependency Parsing - Transition-Based Dependency Parsing, Graph-Based Dependency Parsing, Evaluation.	7
3	**Neural Networks for NLP:** Word representations - Lexical Semantics, Vector Semantics, TF-IDF, Pointwise Mutual Information (PMI), Neural Word embeddings - Word2vec, GloVe, Contextual Word Embeddings. Evaluating Vector Models - Feedforward Neural Networks for Text Classification	10

4	**Advanced NLP and Applications:** Sequence Modelling - Recurrent Neural Networks, RNNs as Language Models, RNNs for NLP tasks, Stacked and Bidirectional RNN architectures, Recursive Neural Networks, LSTM & GRU, Common RNN NLP Architectures, Encoder-Decoder Model with RNNs, Attention models, Transformers. NLP Applications - Machine Translation, Question Answering and Information Retrieval, Introduction to Large Language Models.	**12**

Course Assessment Method
(CIE: 40 marks, ESE: 60 marks)

Continuous Internal Evaluation Marks (CIE):

Attendance	Assignment/ Microproject	Internal Examination-1 (Written)	Internal Examination- 2 (Written)	Total
5	**15**	**10**	**10**	**40**

End Semester Examination Marks (ESE)

In Part A, all questions need to be answered and in Part B, each student can choose any one full question out of two questions

Part A	Part B	Total
• 2 Questions from each module. • Total of 8 Questions, each carrying 3 marks **(8x3 =24 marks)**	• Each question carries 9 marks. • Two questions will be given from each module, out of which 1 question should be answered. • Each question can have a maximum of 3 subdivisions. **(4x9 = 36 marks)**	**60**

Course Outcomes (COs)

At the end of the course students should be able to:

Course Outcome		Bloom's Knowledge Level (KL)
CO1	Understand the foundational concepts of NLP and apply that to do text processing.	**K3**
CO2	Utilize word representations and evaluate vector models for NLP	**K3**
CO3	Analyse and implement advanced linguistic annotation and parsing techniques	**K4**
CO4	Apply advanced sequence modeling techniques using Neural Networks	**K3**
CO5	Apply NLP techniques in machine translation, question answering, and information retrieval.	**K3**

Note*: K1- Remember, K2- Understand, K3- Apply, K4- Analyse, K5- Evaluate, K6- Create*

CO-PO Mapping Table (Mapping of Course Outcomes to Program Outcomes)

	PO1	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**	**PO8**	**PO9**	**PO10**	**PO11**	**PO12**
CO1	3	3								3		
CO2	3	3			3							
CO3	3	3									3	
CO4	3	3	3		3							
CO5	3	3	3			3						

Note: *1: Slight (Low), 2: Moderate (Medium), 3: Substantial (High), -: No Correlation*

Text Books				
Sl. No	**Title of the Book**	**Name of the Author/s**	**Name of the Publisher**	**Edition and Year**
1	Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition	Dan Jurafsky and James H. Martin.	Pearson	2006
2	Introduction to Natural Language Processing	Jacob Eisenstein	MIT Press	2019
3	Natural Language Processing with Transformers	Lewis Tunstall, Leandro von Werra, and Thomas Wolf	O'Reilly	2022

Reference Books				
Sl. No	**Title of the Book**	**Name of the Author/s**	**Name of the Publisher**	**Edition and Year**
1	Deep learning for Natural Language Processing	Stephan Raaijmakers	Manning	2022
2	Natural Language Processing with PyTorch	Delip Rao and Brian McMahan	O'Reilly	2019
3	Deep Learning	Ian Goodfellow, Yoshua Bengio, Aaron Courville	MIT Press	2016

Video Links (NPTEL, SWAYAM…)	
No.	**Link ID**
1	https://onlinecourses.nptel.ac.in/noc19_cs56