Title of the Course	Natural Language Processing

Credits	T	P	E	C
	3	1	0	4

Course Type (Theory/Practical/Integrated)	Integrated
Course Category	Core Discipline
Pre-Requisite	1. Basic proficiency in Python programming 2. Foundational knowledge of Machine Learning 3. Introductory neural networks/deep learning concepts
Learning Objectives	1. Understand fundamental concepts of Natural Language Processing and core text preprocessing techniques. 2. Apply classical and neural approaches for text classification and sequence labelling tasks. 3. Analyse sequence-to-sequence architectures, including RNNs and Transformer encoder–decoder models, for machine translation. 4. Explain and utilise pretrained language models such as ELMo and BERT for downstream NLP tasks. 5. Design and evaluate NLP applications including summarisation, question answering, and dialogue systems.

Course Outcomes & Bloom's Level

CO Code	Course Outcome Statement	Bloom Level
CO1	Explain core NLP concepts including tokenisation, morphology, distributional semantics, and language modelling.	L2
CO2	Apply probabilistic and neural techniques for text classification and sequence labelling tasks such as POS tagging and NER.	L3
CO3	Analyse sequence-to-sequence models and decoding strategies for machine translation tasks.	L4
CO4	Evaluate contextual word representation models and the pretrain–finetune paradigm for NLP applications.	L5
CO5	Build NLP solutions for summarisation and open-domain question answering problems.	L4
CO6	Design and assess dialogue systems and chatbots using appropriate evaluation metrics.	L6

Course Elements	Course Element	Coverage Level
	Skill Development	Yes
	Entrepreneurship	Yes
	Employability	No
	Professional Ethics	Yes
	Gender	No
	Human Values	Yes
	Environment & Sustainability	No
SDG (Goals)	SDG 4: Quality Education	
Total Hours of Pedagogy	45 hours Theory (15 hours Self-paced content + 30 hours lecture) 30 hours Practical	

Module#	Content	Pedagogy
M-1	**Introduction** What is Natural Language Processing; Tokenization, Stemming and Lemmatization; Distributional Semantics for meaning representation, Word Vectors; N-gram Language Model; Recurrent Neural Network based Language Model	Self-paced content, Lecture, Modular Assignment
M-2	**Text Classification and Sequence Labelling** Text Classification using Naive Bayes, POS tagging using HMM; RNN for Text Classification, POS Tagging, NER	Self-paced content, Lecture, Modular Assignment
M-3	**Sequence to Sequence (Machine Translation)** RNN for Sequence to Sequence, Decoding Algorithms: Beam Search; Transformer Enoder-Decoder	Self-paced content, Lecture, Modular Assignment
M-4	**Pretraining** Pretraining for contextual represetation: ELMo; BERT, RoBERTa -- Pretrain-finetune paradigm with applications: Classification, NER, QA	Self-paced content, Lecture, Modular Assignment
M-5	**Applications-I** Text Summarization; Open-domain Question-Answering	Self-paced content, Lecture, Modular Assignment
M-6	**Applications-II** Dialog Modeling, Chatbots; Dialog Evaluation	Self-paced content, Lecture, Modular Assignment

List of DIY Modular Assignments

1. Text Classification using Classical and Neural NLP Methods
2. Sentiment Analysis Pipeline - From Classical NLP to RNNs
3. Text Classification and Sequence Labeling with Classical and Neural NLP Models
4. Neural Machine Translation with Seq2Seq and Beam Search Decoding
5. Neural Machine Translation Using Transformer Encoder-Decoder
6. Building Summarisation and Open-Domain Question Answering Pipelines
7. Exploring BERT and RoBERTa for Downstream NLP Tasks
8. Building and Evaluating a Task-Oriented Chatbot

Modular Assignment Mapping

S. No.	***DIY Assignment Title***	***Mapped CO(s)***	***Bloom Level***
1	Text Classification using Classical and Neural NLP Methods	CO2	L3
2	Sentiment Analysis Pipeline – From Classical NLP to RNNs	CO2	L4
3	Text Classification and Sequence Labelling with Classical and Neural NLP Models	CO2, CO3	L4
4	Neural Machine Translation with Seq2Seq and Beam Search Decoding	CO3	L4
5	Neural Machine Translation Using Transformer Encoder–Decoder	CO3	L5
6	Building Summarisation and Open-Domain Question Answering Pipelines	CO5	L3
7	Exploring BERT and RoBERTa for Downstream NLP Tasks	CO4	L5
8	Building and Evaluating a Task-Oriented Chatbot	CO6	L6

Reference Books	1. **Jurafsky, D. and Martin, J.H.**, *Speech and Language Processing*, 3rd Edition (Draft), Pearson Education, 2023. 2. **Eisenstein, J.**, *Introduction to Natural Language Processing*, 1st Edition, MIT Press, 2019. 3. **Goldberg, Y.**, *Neural Network Methods for Natural Language Processing*, 1st Edition, Morgan & Claypool Publishers, 2017. 4. **Manning, C.D. and Schütze, H.**, *Foundations of Statistical Natural Language Processing*, 1st Edition, MIT Press, 1999.

Course Articulation Matrix

CO \ PO	**PO1**	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**	**PO8**	**PO9**	**PO10**	**PO11**	**PO12**
CO1	3	2	–	–	1	–	–	–	–	–	–	2
CO2	2	3	2	2	3	–	–	–	1	–	–	2
CO3	2	3	2	3	3	–	–	–	1	–	–	2
CO4	2	2	–	3	3	–	–	–	–	–	–	3
CO5	2	2	3	2	3	–	–	–	1	2	–	2
CO6	2	2	3	2	3	–	–	–	2	2	–	2