

Branch Name:	M.Sc. (Artificial Intelligence & Data Science)
Program Code:	ET214
Course Title:	Natural Language Processing
Course Code:	1ET2140201
Pre-requisite Course:	Machine Learning

Course Objectives:

1. To introduce foundational understanding in natural language processing
2. To understand the principles and methods of statistical NLP.
3. To develop in-depth understanding of algorithms for processing and analysis of natural languages.
4. To perform statistical analysis of textual data and find useful patterns from the data.
5. To gain practical experience using NLP toolkits.

Teaching and Examination Scheme:

Teaching Scheme (Hours per week)				Evaluation Scheme (Marks)				
Lecture (L)	Tutorial (T)	Practical (P)	Credit	Theory (Marks)		Practical (Marks)		Total (Marks)
				University Assessment	Continuous Assessment	University Assessment	Continuous Assessment	
03	00	02	04	50	50	25	25	150

Course Contents:

Unit No	Topic	Total Hours	Weightage (%)
1	Introduction and Basic Text Processing: Speech and Language Processing, ambiguity problem, typical NLP tasks, tokenization, stemming, lemmatization, stop-word removal.	04	10%
2	Formal Language Modeling: Regular Expressions, Text Normalization, Edit Distance, Unigrams, Bigrams, N-grams, Language Models, Smoothing and Entropy	06	15%
3	Sequence Labeling and POS Tagging: Part-of-Speech Tagging, Named Entity Recognition, Hidden Markov Models, Forward and Viterbi algorithms, EM training	12	25 %
4	Vector Semantics and Embeddings: Lexical Semantics, Word Vectors, TF-IDF, Vector Space Models, Word2Vec and Gensim	08	20 %
5	Applications of Text Mining: Text Classification, Sentiment Analysis	04	10 %
6	Deep Learning Models for NLP: Feedforward Neural Networks, Recurrent Neural Networks, LSTM	08	20 %

Text Books:

1. Daniel Jurafsky and James H. Martin, Speech and Language Processing, Pearson, 3rd edition, 2022.
2. Christopher D. Manning and Hinrich Schütze, Foundations of Statistical Natural Language Processing, MIT Press, 1999
3. Steven Bird, Ewan Klein, and Edward Loper, Natural Language Processing with Python, O'Reilly Media, 2009.C

Reference Books:

1. Jacob Eisenstein – Introduction to Natural Language Processing, MIT Press.
2. Palash Goyal, Sumit Pandey & Karan Jain – Deep Learning for Natural Language Processing, Apress.
3. Denis Rothman – Transformers for Natural Language Processing, Packt Publishing.
4. Yoav Goldberg – Neural Network Methods for Natural Language Processing, Morgan & Claypool.
5. Nitin Indurkha & Fred J. Damerau – Handbook of Natural Language Processing, CRC Press.

LIST OF EXPERIMENTS / PRACTICALS

- Text pre-processing using tokenization, stop-word removal, stemming and lemmatization.
- Word frequency analysis and vocabulary extraction.
- POS tagging and dependency parsing.
- TF-IDF matrix creation and n-gram extraction.
- Named Entity Recognition (NER) using NLP libraries.
- Topic modeling using LDA and visualization.
- Sentiment analysis on movie reviews dataset.
- Build and train a text classifier using deep learning.
- Generate text using character-based RNN.

Course Learning Outcomes (CLO):

COs	Statements	Bloom's Level
CO1	Understand fundamental concepts and tasks in NLP.	L2 – Understand
CO2	Apply statistical and machine learning techniques for text processing	L3 – Apply
CO3	Analyze syntax and semantics using NLP models.	L4 – Analyze
CO4	Implement deep learning models for language processing.	L5 – Evaluate
CO5	Develop NLP applications for real-world datasets.	L6 – Create

Mapping of CLO with PO and PSO

Course Learning Outcomes	Program Outcomes (POs)												Program Specific Outcomes (PSOs)			
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	PSO3	PSO4
CLO1	H	M	L	–	M	–	–	–	–	L	–	M	M	L	M	L
CLO2	H	H	M	M	H	–	–	–	–	L	–	M	H	H	H	M
CLO3	M	H	H	M	M	–	–	–	–	M	–	M	H	M	H	M
CLO4	M	M	H	H	H	–	–	–	–	M	M	M	H	H	H	M
CLO5	M	H	H	H	H	M	L	M	M	H	H	H	H	H	H	H

H:High, M:Medium, L:Low