

DSM364 Natural Language Processing

Semester :VI	Course Title: Natural Language Processing	Credit:4
Course Code: DSE364		(3 T + 1 P)

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

CO	COGNITIVE ABILITIES	COURSE-OUTCOMES
CO1	REMEMBERING	Recall fundamental concepts, linguistics basics, and NLP pipeline components.
CO2	UNDERSTANDING	Explain text preprocessing, feature extraction, and statistical & deep learning models for NLP.
CO3	APPLYING	Implement NLP tasks such as tokenization, POS tagging, sentiment analysis, and text classification.
CO4	ANALYSING	Compare traditional machine learning and deep learning approaches for NLP applications.
CO5	EVALUATING	Evaluate NLP models using precision, recall, F1-score, BLEU, and perplexity metrics.
CO6	CREATING	Design and develop end-to-end NLP solutions for real-world applications such as chatbots, translation, and information retrieval.

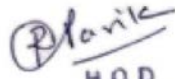
	PSO1	PSO2	PSO3	PSO4	PSO5
CO1	3	2	2	1	1
CO2	2	3	3	2	1
CO3	2	3	3	3	2
CO4	2	3	3	2	2
CO5	3	3	3	2	2
CO6	3	3	3	3	2

UnitNo.	Detailed Syllabus	Teaching Hours
I	Introduction to NLP & Linguistic Foundations - Definition, scope, and applications of NLP (search engines, chatbots, translation, sentiment analysis). - Basics of linguistics: Phonology, morphology, syntax, semantics, pragmatics. - NLP pipeline overview: Text preprocessing, Tokenization, Stop-word removal, Lemmatization, Stemming. - Word representation: Bag-of-Words, TF-IDF, Word embeddings (Word2Vec, GloVe, FastText).	15

II	Traditional NLP Methods • n-grams and Language Models. • POS tagging: Rule-based, statistical, and HMM approaches. • Named Entity Recognition (NER). • Text classification with Naïve Bayes, SVM, Logistic Regression. • Sentiment analysis basics. • Evaluation metrics: Accuracy, Precision, Recall, F1-score, Perplexity.	15
III	Deep Learning for NLP • Neural networks for NLP tasks. • Recurrent Neural Networks (RNN), LSTM, GRU models. • Sequence-to-sequence models and Attention mechanism. • Transformers and BERT: Architecture and applications. • Machine Translation, Summarization, and Question Answering. • Speech recognition basics and integration with NLP.	15
IV	Practical Applications • Text pre-processing with NLTK, SpaCy. • Implementing POS tagging and NER. • Sentiment analysis using supervised ML models. • Word embeddings and similarity measures. • Building a text classification pipeline with Scikit-learn. • Implementing LSTM/Transformer for text generation. • Mini Projects: - o Chatbot development. - o Machine Translation demo. - o Document summarization system.	15

Suggested Reference Books

1. Speech and Language Processing – Daniel Jurafsky & James H. Martin, Pearson, 3rd Edition (Draft), 2023.
2. Foundations of Statistical Natural Language Processing – Christopher D. Manning & Hinrich Schütze, MIT Press, 1999.
3. Natural Language Processing with Python – Steven Bird, Ewan Klein & Edward Loper, O'Reilly, 2009.
4. Neural Network Methods in Natural Language Processing – Yoav Goldberg, Morgan & Claypool, 2017.
5. Transformers for Natural Language Processing – Denis Rothman, Packt Publishing, 2021.


 H.O.D
 Dept of DATA SCIENCE & ANALYTICS
 M.G. Science Institute, Ahmedabad-9.