

Natural Language Processing

- 1. Course number and name:** 020NLPES3 Natural Language Processing
- 2. Credits and contact hours:** 4 ECTS credits, 2x1:15 contact hours
- 3. Name of course coordinator:**
- 4. Instructional materials:** PowerPoint slides; Web tutorials; Group Assignments

References:

- Jurafsky, D. and Martin, J.H., “Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition”.
- DeepLearning.ai
- SimpliLearn platform.
- “All you need to know about Natural Language Processing” by AnalyticsVidhya
- AnalyticsVidhya platform.

5. Specific course information

a. Catalog description:

This Natural Language Processing (NLP) course offers a foundational and practical understanding of key NLP techniques, from text processing and feature extraction to modern machine learning and deep learning approaches. Students will explore core methods like tokenization, sentiment analysis, and topic modeling, using tools such as NLTK and spaCy. The course delves into advanced models, including RNNs, LSTMs, and Transformers like BERT and GPT, highlighting their real-world applications. Through hands-on projects, students will learn to build and evaluate NLP models, understand ethical considerations, and apply NLP techniques across various industries, preparing them for advanced work in AI-driven language processing.

b. Prerequisites: None.

c. Required for CCE Artificial Intelligence Option students; **Selected Elective** for students in the CCE Software Engineering and Telecommunication Networks Options.

6. Educational objectives for the course

a. Specific Course Learning Outcomes:

- Define Natural Language Processing and explain its significance in the field of Artificial Intelligence.
- Implement key NLP techniques including Named Entity Recognition, Sentiment Analysis, and Topic Modeling in real-world scenarios. Also utilize various preprocessing methods such as tokenization, stemming, and lemmatization to prepare textual data for analysis.

- Demonstrate proficiency in using NLP tools and libraries such as NLTK, spaCy, Gensim.
- Follow the complete workflow of building an NLP model, including data collection, preprocessing, feature extraction, model training, and evaluation.
- Apply the latest advancements in NLP, such as Transformers (e.g., BERT, GPT-3), to enhance model performance and achieve state-of-the-art results.

b. PI addressed by the course:

PI	1.2	1.3	2.3	2.4	2.5	4.2	7.1	7.2
Covered	x	x	x	x	x	x	x	x
Assessed	x	x	x	x	x	x		

7. Brief list of topics to be covered

- Introduction to Natural Language Processing (2 lectures)
- Fundamentals of Text Processing: Tokenization and Segmentation, Text Normalization (Stemming, Lemmatization), Part-of-Speech Tagging (2 lectures)
- NLP Techniques and Methods: Named Entity Recognition (NER), Sentiment Analysis, Topic Modeling (LDA) (2 lectures)
- NLP Tools and Libraries: Introduction to NLTK, Working with spaCy, Exploring Gensim (2 lectures)
- Feature Extraction and Representation: Bag-of-Words Model, Term Frequency-Inverse Document Frequency (TF-IDF), Word Embeddings (Word2Vec) (2 lectures)
- Machine Learning for NLP: Classification and Clustering (2 lectures)
- Deep Learning in NLP: Introduction to Neural Networks, Recurrent Neural Networks (RNNs), Long Short-Term Memory Networks (LSTMs), Attention Mechanisms and Transformers (2 lectures)
- Transformers and Modern NLP Models: Overview of Transformers, BERT (Bidirectional Encoder Representations from Transformers), GPT (Generative Pre-trained Transformer), Fine-Tuning Pre-trained Models (4 lectures)
- Building and Evaluating NLP Models: Data Collection and Preprocessing, Feature Extraction Techniques, Model Training and Hyperparameter Tuning, Evaluation Metrics (Accuracy, Precision, Recall, F1 Score) (2 lectures)
- Latest Trends and Future Directions: Generative AI and its Implications, Recent Breakthroughs in NLP Research, Ethical Considerations and Challenges in NLP (2 lectures):
- Hands-On Projects and Case Studies: Implementing an NLP Solution from Start to Finish, Implementing an NLP Solution from Start to Finish, Analyzing Real-World Datasets, Case Studies of NLP Applications in Various Industries (2 lectures)