

NLP for economic decisions

ECTS : 3

Description du contenu de l'enseignement :

General introduction: aims and uses of NLP

Part I: translating texts into quantities

1. Tokenization, preprocessing
2. Word frequency and inverse-document frequency (tf and tf-idf)
3. Word embeddings

Part II: Traditional NLP methods

Methods used before LLMs

1. Text classification
2. Topic modelling and text summarization
3. Sentiment analysis

Part III: NLP with LLMs

1. Short presentation of LLMs: transformers, encoders-only models, decoders-only models
2. Text classification with LLM
3. Topic modelling and text summarization
4. Sentiment analysis
5. Generating prompt for generative LLM
6. Fine-tuning of LLM

Compétence à acquérir :

At the end of this course, the students should:

- have a good knowledge of the main methods used in Natural Language Processing
- be able to implement these methods to data with Python
- be able to interpret the results obtained through these examples

Mode de contrôle des connaissances :

The evaluation is based on a project made by groups of two students

Bibliographie, lectures recommandées :

- Albrecht, Jens, Ramachandran, Sidharth, Winkler, Christian. *Blueprints for text analytics using Python machine learning-based solutions for common real world (NLP) applications*, O'Reilly, 2020
- Alammur Jay and Maarten Grootendorst. *Hands-On Large Language Models, Language Understanding and Generation*, O'Reilly, 2024
- Useful website: Kaggle

Document susceptible de mise à jour - 01/04/2026

Université Paris Dauphine - PSL - Place du Maréchal de Lattre de Tassigny - 75775 PARIS Cedex 16