

# NLP with Python

---

Nom : **PICAULT**

Prénom : **Matthieu**

Année : **M2**

Semestre : **9**

Nature : **CM**

Volume horaire : **12 H**

ECTS / Coef : **2**

Prérequis	Introduction à Python (M1 ESA) Installation de la suite Anaconda
Résumé	<p>Ce cours se concentre sur le traitement des informations textuelles (articles de presse, réseaux sociaux, documents officiels,...) et son utilisation dans les domaines de l'Economie et de la Finance.</p> <p>Afin d'apprendre à extraire les informations importantes d'un texte, les étudiants s'appuient sur les principaux packages de Python en analyse textuelle comme <i>Nltk</i> et <i>Spacy</i> et travaillent à partir de cas pratiques.</p> <p>L'évaluation s'appuie sur un projet de groupe et sur un examen terminal.</p>
Objectifs	<p>L'objectif de ce cours est de permettre aux étudiants de :</p> <ul style="list-style-type: none"><li>- Manipuler des données textuelles (<i>string</i>).</li><li>- Isoler les éléments important d'un texte en étudiant sa structure grammaticale (<i>Part-of-Speech tagging</i>)</li><li>- Calculer la complexité de documents au sein d'un corpus et mesurer la similarité entre les documents.</li><li>- Identifier les occurrences (<i>TF-IDF</i>) et les thèmes principaux d'un corpus de documents (<i>Latent Dirichlet Allocation</i>).</li><li>- Estimer le sentiment associé aux documents du corpus.</li><li>- Introduction à l'utilisation du Machine Learning aux textes.</li></ul>
Bibliographie	<ul style="list-style-type: none"><li>• Bird, S., Klein, E., &amp; Loper, E. (2009). <i>Natural language processing with Python: analyzing text with the natural language toolkit</i>. " O'Reilly Media, Inc."</li><li>• Hardeniya, N., Perkins, J., Chopra, D., Joshi, N., &amp; Mathur, I. (2016). <i>Natural language processing: python and NLTK</i>. Packt Publishing Ltd.</li><li>• Perkins, J. (2014). <i>Python 3 text processing with NLTK 3 cookbook</i>. Packt Publishing Ltd.</li><li>• Beysolow II, T. (2018). <i>Applied Natural Language Processing with Python: Implementing Machine Learning and Deep Learning Algorithms for Natural Language Processing</i>. Apress.</li><li>• Loughran, T., &amp; McDonald, B. (2016). Textual analysis in accounting and finance: A survey. <i>Journal of Accounting Research</i>, 54(4), 1187-1230.</li></ul>

---

## PLAN

---

Introduction: What is NLP and text analysis?

Section 1: Acquiring and preparing a corpus

Section 2: Term-frequency Matrix and word2vec

Section 3: Similarity, complexity,...

Section 4: Identifying the topics across a corpus using the LDA method.

Section 5: Measuring sentiment in a document