Contenus

Code du Cours	Nom du Cours	Semestre du Cours	Cours Théoriques	Travaux Dirigés (TD)	Travaux Pratiques (TP)	Crédit du Cours	ECTS
MAT410	Traitement du Langage Naturel	8	3	0	0	3	5

Cours Pré-Requis	
Conditions d'Admission au Cours	

Langue du Cours	Anglais
Type de Cours	Électif
Niveau du Cours	Licence
Objectif du Cours	Le cours vise à fournir une compréhension approfondie du Traitement Automatique du Langage, explorant à la fois les concepts fondamentaux et les techniques avancées. Il est conçu pour équiper les étudiants avec les connaissances théoriques et les compétences pratiques nécessaires pour appliquer le TAL dans des applications réelles, telles que la classification de texte, l'analyse de sentiments, la traduction automatique et les systèmes de réponse aux questions. Data Structures and Algorithms, Linear Algebra, Introduction to Artificial Intelligence-Machine Learning .
Contenus	Le contenu du cours sera structuré autour des domaines clés suivants : Introduction au TAL : Aperçu du TAL et de ses applications. Traitement du Texte : Prétraitement de base du texte, tokenisation, racinisation, lemmatisation. Modèles de Langage : N-grammes, modèles probabilistes, modèles basés sur les réseaux de neurones. Étiquetage des Parties du Discours et Reconnaissance d'Entités Nommées : Techniques et applications. Syntaxe et Analyse Syntaxique : Analyse de la structure des phrases, analyse de dépendance. Sémantique : Plongements de mots, plongements contextuels (BERT, GPT). Traduction Automatique : Approches de la traduction automatique, modèles de séquence à séquence. Questions-Réponses et Chatbots : Construction de systèmes qui comprennent et génèrent des réponses similaires à celles des humains. Considérations Éthiques dans le TAL : Biais, équité et implications des technologies TAL.
Ressources	Ressource Principale : Les cours en ligne de Chris Manning sur le TAL. Manuels : Lectures recommandées de "Speech and Language Processing" par Jurafsky & Martin et d'autres littératures pertinentes sur le TAL. Logiciels et Outils : Python, Jupyter Notebooks, NLTK, spaCy, TensorFlow et PyTorch. Forums en Ligne : Plateformes comme Stack Overflow et GitHub pour le soutien communautaire et le partage de code.

Intitulés des Sujets Théoriques

Semaine	Intitulés des Sujets
1	Introduction au TAL ; histoire et applications.
2	Bases du traitement de texte ; travail avec des données textuelles.
3	Modèles de langue ; introduction aux n-grammes et aux modèles probabilistes.
4	Modèles de langue avancés ; introduction aux réseaux de neurones dans le TAL.
5	Étiquetage des parties du discours ; compréhension et mise en œuvre des algorithmes d'étiquetage.
6	Reconnaissance d'entités nommées ; techniques et outils.
7	Syntaxe et analyse syntaxique ; analyse de la structure des phrases.
8	Sémantique ; exploration des plongements de mots et des plongements contextuels.
9	Traduction automatique ; compréhension et construction de modèles de traduction.
10	Sujets avancés en traduction automatique ; exploration des modèles à la pointe de la technologie.
11	Questions-réponses et chatbots ; conception de systèmes pour l'interaction.

Semaine	Intitulés des Sujets
12	Considérations éthiques dans le TAL ; discussion sur les biais, l'équité et l'impact social.
13	Présentations de projets de groupe ; application de ce qui a été appris.
14	Clôture du cours ; révision et examens finaux.