

5.00 crédits	30.0 h + 15.0 h	Q1
--------------	-----------------	----

Enseignants	Dupont Pierre ;
Langue d'enseignement	Anglais > Facilités pour suivre le cours en français
Lieu du cours	Louvain-la-Neuve
Préalables	notions de base d'algorithmique telles qu'enseignées par le cours LINFO1103 compétences en probabilité et statistiques tels que visées par le cours LBIR1212 ou LEPL1109 souhaitable : notions approfondies d'algorithmique et structures de données telles que visées par le cours LINFO1121
Thèmes abordés	<ul style="list-style-type: none"> • Différents niveaux d'analyses linguistiques • Modélisation probabiliste du langage • Etiquetage en parties du discours • Sémantique vectorielle et méthodes d'encodage vectoriel de mots • Apprentissage automatique profond appliqué au traitement de la langue • Mécanismes d'attention et transformeurs • Systèmes de questions-réponses et agents conversationnels • Applications en ingénierie linguistique telles que les logiciels de complétion automatique, d'étiquetage automatique, de résumé ou de traduction automatique, de recherche d'informations et de robots conversationnels
Acquis d'apprentissage	<p>A la fin de cette unité d'enseignement, l'étudiant est capable de :</p> <p>Eu égard au référentiel AA du programme « Master ingénieur civil en informatique », ce cours contribue au développement, à l'acquisition et à l'évaluation des acquis d'apprentissage suivants :</p> <p>INFO1.1-3 INFO2.3-4 INFO5.4-5 INFO6.1, INFO6.5</p> <p>Eu égard au référentiel AA du programme « Master [120] en sciences informatiques », ce cours contribue au développement, à l'acquisition et à l'évaluation des acquis d'apprentissage suivants :</p> <p>SINF1.M4 SINF2.3-4 SINF5.3-4 SINF6.1, SINF6.5</p> <p>Les étudiants ayant suivi avec fruit ce cours seront capables de</p> <ul style="list-style-type: none"> • décrire les concepts fondamentaux de la modélisation du langage naturel • maîtriser la méthodologie de l'utilisation de ressources linguistiques, en particulier sous la forme de corpus à grande échelle, éventuellement annotés ou structurés • appliquer de manière pertinente les techniques statistiques de modélisation du langage • mettre en oeuvre des méthodes récentes relevant de l'apprentissage automatique en lien avec le traitement du langage • développer des applications en ingénierie linguistique <p>Les étudiants auront développé des compétences méthodologiques et opérationnelles. En particulier, ils auront développé leur capacité à</p> <ul style="list-style-type: none"> • s'intégrer dans une approche pluridisciplinaire entre l'informatique et la linguistique, en utilisant à bon escient la terminologie, les outils et méthodes existants, • gérer au mieux le temps disponible pour mener à bien des projets de moyenne ampleur, • manipuler et exploiter de grandes quantités de données.

Modes d'évaluation des acquis des étudiants	<p>Calcul de la note globale pour le cours</p> <p>Les projets valent pour 30 % de la note globale du cours, 70 % pour l'examen final (à livre fermé). Les projets ne peuvent pas être refaits en seconde session. La note globale unique des projets est fixée à la fin du quadrimestre et reprise telle quelle dans la note globale du cours en seconde session. L'examen final est, par défaut, un écrit (sur papier ou, le cas échéant, sur un ordinateur).</p> <p>Règles concernant les collaborations entre étudiants et l'usage de ressources externes, y compris les IA génératives</p> <p>Une étude collaborative entre les étudiant.e.s est encouragée lors des séances de suivi de projets et via un forum d'échanges sur Moodle.</p> <p>Chaque étudiant.e doit soumettre une solution personnelle pour chaque projet. Le recours à des ressources publiques (p.ex. stackoverflow.com), ce-inclus des IA génératives (p.ex. chatGPT) est autorisé pour autant que chaque (fragment de) code soumis par l'étudiant.e mentionne spécifiquement toutes les ressources utilisées.</p> <p>La diffusion ou l'échange entre étudiants de (fragments de) code ne sont pas autorisés par quelque moyen que ce soit (GitHub, Facebook, Discord, ...) et ce même après la date limite de remise des projets.</p> <p>Le non-respect de ces règles pour n'importe quel projet entraînera une note de 0 pour l'ensemble des projets.</p> <p>L'examen final sur ordinateur doit se faire sans accès à aucune ressource externe.</p> <p>Ces règles sont exposées et détaillées lors du premier cours (voir site Moodle du cours).</p>
Méthodes d'enseignement	<ul style="list-style-type: none"> • Cours magistraux • Projets pratiques implémentés en Python sur le serveur <i>Inginious</i>
Contenu	<ul style="list-style-type: none"> • Différents niveaux d'analyse linguistique • Modélisation probabiliste du langage • Etiquetage en parties du discours • Sémantique vectorielle et méthodes d'encodage vectoriel de mots • Apprentissage automatique profond appliqué au traitement de la langue • Mécanismes d'attention et transformeurs • Systèmes de questions-réponses et agents conversationnels • Applications en ingénierie linguistique telles que les logiciels de complétion automatique, d'étiquetage automatique, de résumé ou traduction automatique, de recherche d'informations et de robots conversationnels
Ressources en ligne	moodle.uclouvain.be/course/view.php?id=1182
Bibliographie	One recommended textbook - un ouvrage conseillé : <ul style="list-style-type: none"> • Speech and Language Processing, D. Jurafsky and J.H. Martin, Prentice Hall.
Autres infos	<p>Préalables indispensables</p> <ul style="list-style-type: none"> • Connaissance préalable et expérience acquises en programmation en Python • Bases en structures de données et algorithmique <p>ces préalables sont couverts par les cours LINFO1101 et LINFO1103, ou équivalents.</p> <p>Préalables additionnels</p> <ul style="list-style-type: none"> • Connaissances de bases en probabilités (probabilité conditionnelles, indépendance probabiliste, règle de Bayes, ...) <p>Connaissances complémentaires</p> <ul style="list-style-type: none"> • Une connaissance acquise de méthodes d'apprentissage automatique est un plus mais ne constitue pas un préalable
Faculté ou entité en charge:	INFO

Programmes / formations proposant cette unité d'enseignement (UE)				
Intitulé du programme	Sigle	Crédits	Prérequis	Acquis d'apprentissage
Master [120] en science des données, orientation statistique	DATS2M	5		
Master [120] : ingénieur civil en chimie et science des matériaux	KIMA2M	5		
Master [120] : ingénieur civil des constructions	GCE2M	5		
Master [120] : ingénieur civil biomédical	GBIO2M	5		
Master [120] en linguistique	LING2M	5		
Master [120] : ingénieur civil mécanicien	MECA2M	5		
Master [120] : ingénieur civil électricien	ELEC2M	5		
Master [120] : ingénieur civil physicien	FYAP2M	5		
Master [120] : ingénieur civil en informatique	INFO2M	5		
Master [120] en sciences informatiques	SINF2M	5		
Master [120] : ingénieur civil électromécanicien	ELME2M	5		
Master [120] : ingénieur civil en science des données	DATE2M	5		
Master [120] en science des données, orientation technologies de l'information	DATI2M	5		
Master [120] : ingénieur civil en génie de l'énergie	NRGY2M	5		