Master Informatique

Introduction au traitement automatique du langage

Responsables

Frederic BECHET
frederic.BECHET@univ-amu.fr

Carlos RAMISCH
carlos.ramisch@univ-amu.fr

Descriptions

Code : S51IN2A2

Nature : Unité d'enseignement

Domaines : Sciences et Technologies

Informations

Composante : Faculté des Sciences

CONTENU

L'objectif de ce cours est d'apporter aux étudiants des connaissances fondamentales en traitement automatique des langues (TAL). Le cours abordera des modèles de TAL statistiques classiques. Après un bref rappel de notions de probabilité appliquées aux textes, nous aborderons la représentation de documents, c.-à-d. l'encodage de textes, la tokenisation et la représentation de documents sous la forme vectorielle avec des techniques fondées sur les "sac de mots". Nous étudierons également des modèles distributionnels pour la représentation de mots sous la forme de vecteurs, avec des applications en classification de textes et en similarité de mots. Les modèles de langage fondés sur les n-grammes s'en suivent, avec des applications à la génération de textes. Le dernier sujet abordé ce sont les étiqueteurs fondés sur des machines à état (tels que les modèles de Markov cachés) et leurs applications, par exemple en étiquetage morphosyntaxique.

VOLUME HORAIRE

- Volume total: 27 heures
- Cours magistraux: 10 heures
- Travaux dirigés: 7 heures
- Travaux pratiques: 10 heures

CODES APOGÉE

- SINBU02L [ELP]

M3C

Aucune donnée M3C trouvée

POUR PLUS D'INFORMATIONS

Aller sur le site de l'offre de formation...

Dernière modification le 15/01/2024