

Proposition d'action

Techniques d'apprentissage automatique pour le TAL

Jean-Yves Antoine, Anaïs Lefeuvre, Agata Savary

LI, équipe BDTLN, Centre Universitaire de Blois

Denys Duchier

LIFO, équipe CA, Université Orléans

MACHINE LEARNING ET RECHERCHES EN TAL EN REGION CENTRE

Les techniques d'apprentissage automatique sont incontournables en TALN, du fait de la mise en disposition croissante de corpus comme banques de données d'envergure pour l'entraînement de modèles

- “ **LIFO** (CA) – Recherches formelles sur l'apprentissage automatique, en relation avec le TAL
- “ **LI** (BDTLN) – Vision applicative mais également travaux plus approfondis entre chercheurs en TAL et en fouille de données
 - “ Fouille de texte pour la reconnaissance des entités nommées (PhD Damien Nouvel)
 - “ Classification automatique pour la reconnaissance d'auteurs de Blog (stage Eloquenz)

ANCOR : UN CADRE POUR UNE ACTION DE RECHERCHE COMMUNE

- “ **Partenaires** – LI, LLL
- “ **Objectif** – Etude de la coréférence et de l'anaphore en langue orale

loc1 : je vous recommande les Pyrénées

loc2 : c' est magnifique



“ Résultats

ANCOR, plus grande ressource orale annotée en coréférence

CROC [LaTTiCe] : premier système de résolution appris sur ANCOR

ACTIONS ICVL : ML POUR LA RÉOLUTION DES CORÉFÉRENCES

Objectif à court-terme

- “ Réalisation d'un système de résolution des coréférences par adaptation au français oral de BART, plateforme modulaire intégrant plusieurs techniques d'apprentissage / classification
- “ Stage Master (expérience LaTTiCe)

Intérêt pour les partenaires

Meilleure compréhension des techniques d'apprentissage appliquées au TAL

- “ Comparaison des différentes techniques proposées par BART
- “ Etude des relations entre pertinence des traits d'apprentissage et propriétés linguistiques observées en corpus
- “ Etude de l'adaptation de modèles pour l'oral, voire pour les registres de l'oral (interview, standard téléphonique, renseignement touristique..)

Visibilité internationale par participation à des campagnes d'évaluation.