

Introduction au text-mining avec AlvisNLP

02 et 03 juin 2022

mouhamadou.ba@inrae.fr

robert.bossy@inrae.fr

Objectifs Pédagogiques

Formation dédiée à l'analyse de données textuelles (text-mining).

L'objectif est l'acquisition des principales techniques pour la Reconnaissance d'Entités Nommées (REN) à partir de textes.

Format de la formation

- Cours introductifs.
- TPs, manipulation de workflows (plan) de text-mining avec AlvisNLP.
<https://github.com/Bibliome/alvisnlp>

A l'issue de cette formation, vous serez capables de :

- Acquérir des connaissances théoriques sur des approches pour la REN
- Acquérir les compétences pratiques pour effectuer de façon autonome une première approche pour une application de text-mining.
- Être autonome avec AlvisNLP pour créer et exécuter des plans

Contexte de travail

Ressources de la formation

- Le répertoire de la formation est disponible sur ForgeMIA
- <https://forgemia.inra.fr/bibliome/formation-alvisnlp.git>

Supports

- La présentation [presentation/formation-introduction-au-text-mining-avec-alvisnlp.pdf](#)
- Les exercices pratiques (sous-dossiers dans le dossier [formation-alvisnlp/exercices](#))
- La documentation sur les modules <https://bibliome.github.io/alvisnlp/reference/Module-reference>
- La documentation de AlvisNLP <https://bibliome.github.io/alvisnlp/>

Outils dont vous aurez besoin (déjà disponibles sur les postes de formation)

- Singularity
- Image Singularity de AlvisNLP
- Terminal
- Editeur de textes
- Navigateur (Chrome)

Formulaire d'évaluation de la formation à remplir avant la fin

- https://migale.inrae.fr/Evaluation_formation

Contexte de travail

On utilisera un répertoire de travail en local

- sur les postes formation, votre dossier de travail est déjà prêt sous le home
`cd ~/formation-alvisnlp`

Sinon, pour installer le répertoire de travail

- Ouvrir un nouveau terminal
- Taper la commande
`git clone https://forgemia.inra.fr/bibliome/formation-alvisnlp.git`
- Aller dans le répertoire de travail
`cd formation-alvisnlp/`
- Télécharger l'image d'AlvisNLP (singularity doit être déjà installé)
`singularity pull library://migale/default/alvisnlp:0.8.1`

Pour dérouler les exercices via le terminal

- il faut être dans le répertoire de l'exercice en cours de traitement
`cd formation-alvisnlp/exercices/...`
- Ensuite, lancer la commande d'exécution d'AlvisNLP
`alvisnlp -browser [file].plan`

Planning

Jeudi 2 juin	
9h30 – 17h00 Pause : 12h30 – 14h00	Presentation du TM et de la REN
	Projection de lexiques (ex 1-*, 2-* et 3*)
	Évaluation de la REN et application de patrons (ex 4-* et 5-*)
Vendredi 3 juin	
9h30 – 17h00 Pause : 12h30 – 14h00	Notation BIO et apprentissage (ex 6 et 7)
	Suite apprentissage et composition (ex 7 et 8)