

Université Larbi Ben Mhidi

Niveau : 1ère année Master IA

Département MI

Module : TLN

Année : 2024-2025

Responsable Dr.Belhouchette

Série de TD 2 : Prétraitement du texte

Exercice 01

On vous donne le texte suivant réalisez une lemmatisation sur chaque mot du texte

Les chercheurs ont développé des solutions innovantes pour améliorer la performance des systèmes de traitement de données dans des environnements complexes.

Exercice 02

Vous disposez du texte suivant :

Les chercheurs ont étudié différentes stratégies pour améliorer les performances des systèmes informatiques.

Effectuez les opérations suivantes :

- Appliquez la **lemmatisation** sur chaque mot du texte
- Appliquez le **stemming**

Exercice 03

Effectuez la tokenisation de ce texte en :

1. Séparant les mots (tokens) par des espaces
2. Séparant les mots, les signes de ponctuation et les autres symboles en différents tokens

L'analyse des données est cruciale pour la prise de décision dans des projets de grande envergure.

Exercice 04

Identifier et supprimer les stop words du texte suivant :

Le rapport de performance de l'équipe présente des résultats satisfaisants, mais des améliorations sont possibles.