

Université Larbi Ben Mhidi

Niveau : 1ère année Master IA

Département MI

Module : TLN

Année : 2024-2025

Série de TP 3 : Représentation du texte

Exercice 01 :

Créer une représentation de type sac de mots à partir d'un corpus.

Pour ce faire suivez ces étapes :

1. Créez un corpus de phrases
2. Utilisez `CountVectorizer` pour créer le modèle Bag of Words

Exercice 02 :

Appliquer la pondération TF-IDF sur un corpus en utilisant :

1. Le même corpus que précédemment.
2. `TfidfVectorizer`

Exercice 03 :

Utiliser SpaCy pour obtenir des embeddings de mots.