

Université Larbi Ben Mhidi

Niveau : 1ère année Master IA

Département MI

Module : TLN

Année : 2024-2025

Responsable : Dr.Belhouchette

Série de TD 6 : Modèles de langue et classification de texte

Exercice 01 :

Voici un extrait d'un manuel technique sur les bases de données SQL :

"SQL permet de créer des bases de données. Une base de données stocke des informations. Une requête SELECT extrait les données. La clause WHERE filtre les résultats."

1. Construisez un modèle bigram et un modèle trigram à partir du corpus ci-dessus.
2. Déterminez la probabilité des phrases suivantes avec le modèle bigram :

"Une requête SELECT filtre les résultats."

"Les bases de données stockent des requêtes."

Exercice 02 :

Un modèle a été testé sur 1000 articles techniques et a donné les résultats suivants :

- ✓ Vrais positifs (TP) : 450
- ✓ Faux positifs (FP) : 50
- ✓ Vrais négatifs (TN) : 400
- ✓ Faux négatifs (FN) : 100

Calcul des métriques :

- Accuracy
- Precision
- Recall
- F1-score