
Filière : CI 2 : Informatique
Module : Traitement du langage naturel
TP 3: Analyse du discours (POS)

Objectifs:

- Comprendre le concept de POS Tagging et son importance en NLP.
- Expérimenter avec des techniques de POS Tagging basées sur des règles et des méthodes statistiques.
- Analyser les résultats et identifier les forces et faiblesses de chaque approche.

POS Tagging basé sur des règles

1. En se basant sur des règles simples, créez un script Python qui prend une phrase en entrée et renvoie la phrase avec un POS tag pour chaque mot.
2. Testez votre script sur plusieurs phrases pour observer son comportement (Voir exemples).
3. Discutez des limitations des approches basées sur des règles pour le POS Tagging. Donnez des exemples de cas où votre script pourrait échouer.

POS Tagging avec NLTK

1. Importez la bibliothèque NLTK dans votre script Python et téléchargez les ressources nécessaires pour le français (nltk.download('punkt')).
2. Utilisez la fonction nltk.pos_tag pour étiqueter automatiquement les mots d'une phrase en POS tags.
3. Comparez les résultats obtenus avec votre script basé sur des règles et ceux de la fonction NLTK.
4. Expliquez les avantages et les inconvénients de l'utilisation de NLTK pour le POS Tagging.

Exploration

1. Essayez d'améliorer votre script basé sur des règles en ajoutant des règles plus complexes.
2. Explorez d'autres bibliothèques Python pour le POS Tagging (spaCy).

Exemples de texte :

1. "The quick brown fox jumps over the lazy dog."
2. Football federations of Morocco, Portugal and Spain officially signed, on Wednesday, the bid agreement to host the 2030 FIFA World Cup.

Compte Rendu : Soumettez votre code Python avec des commentaires explicatifs via la plateforme à la fin de la séance.
