

Introduction au NLP (Natural Language Processing)

Le NLP (traitement du langage naturel) est un domaine de l'intelligence artificielle qui permet aux machines de comprendre, analyser, generer et interagir avec le langage humain.

Comment fonctionne le NLP ?

1. Pretraitement du texte : nettoyage, tokenisation, suppression des mots inutiles.
2. Representation du texte : convertir les mots en vecteurs (word embeddings).
3. Modelisation : entrainer un modele (comme GPT) a comprendre ou generer du texte.
4. Taches de NLP : classification, generation de texte, resume, traduction, etc.

Types de NLP

- Classification de texte
- Analyse de sentiments
- Reconnaissance dentites nommees (NER)
- Resume automatique
- Traduction automatique
- Generation de texte

Projet : Generation de texte

Etapes pour creer un generateur de texte :

1. Choisir un dataset (corpus) de texte.
2. Nettoyer et pretraiter le texte.
3. Utiliser un modele pre-entraine (comme GPT-2).

4. Fine-tuner le modele sur ton propre dataset.
5. Generer du texte avec un prompt (debut de phrase).

Outils recommandes :

- Hugging Face Transformers
- Google Colab
- Python

Idees de projets de generation de texte :

- Generateur dhistoires personnalisees
- Generateur de poemes ou citations
- Completeur de lettres damour ou messages Insta
- Generateur de dialogues de films ou series
- Generateur de resumes

Utilisation dans le domaine du crime :

1. Generateur de dialogues pour entrainer des policiers.
2. Generation de rapports de police a partir de donnees brutes.
3. Creation de faux messages de phishing pour la cybersecurite.
4. Simulation dalibis pour detecter les incoherences dans les recits.
5. Aide a lanalyse de messages suspects ou menaces.

Conclusion :

Le NLP est un domaine puissant avec de nombreuses applications. Le projet de generation de texte

te permet de mieux comprendre son fonctionnement, et tu peux l'appliquer meme dans des contextes serieux comme la securite, l'education ou la cyberscurite.