

Correction de texte après reconnaissance vocale avec Whisper et T5

Introduction

Ce projet propose un pipeline combinant Whisper (ASR) d'OpenAI et T5 (NLP) de Google pour améliorer la reconnaissance vocale grâce à la correction automatique du texte. Le système corrige les erreurs de texte transcrit (par exemple, grammaire, contexte) pour des applications telles que les assistants virtuels et la transcription en temps réel.

Methodology

Architecture du Pipeline:

- Entrée Audio :** Fichier ou flux audio brut.
- Transcription par Whisper :** Conversion de la parole en texte, pouvant comporter des erreurs.
- Correction par T5 :** Amélioration grammaticale et adaptation contextuelle du texte brut.
- Sortie :** Texte corrigé, prêt pour affichage ou traduction.



Technologies utilisées:

Backend : Flask (API intégrant Whisper et T5).
Frontend : Next.js (interface utilisateur interactive)

Group info

Le Nom de group : Team DeepMind
Le numero de group : 1

Membres du group

- Houssame Bouyous
- Brahim Chbab
- Chakir Abderrahmane

Les profs

Le prof de modèle: Pr. El Habib Ben Lahmar

Encadrant: Dr. Zakaria Elfakir

Outils et Environnement

- Python, PyTorch
- Hugging Face Transformers

Contact info



Test model T5

Original	Corrigé
he go to school every day.	he goes to school every day.
I has three apple	I have three apple.
they was playing football yesterday	they were playing football yesterday.

Conclusion

Ce pipeline associe la robustesse de Whisper en reconnaissance vocale aux capacités avancées de correction linguistique de T5, offrant ainsi des transcriptions plus fiables et précises. Les perspectives incluent l’optimisation pour le temps réel et l’extension à plusieurs langues.

Interface

