

Rapport Technique – Projet IA/ML

Assistant Intelligent de Podcast à Base de Modèles Entraînés

April 9, 2025

Contents

1	Introduction	2
2	Concept global du projet	2
3	Fonctionnalités de l'application	2
4	Travail IA/ML à réaliser par nous-mêmes	3
4.1	1. Modèle de transcription vocale	3
4.2	2. Chapitrage automatique	3
4.3	3. Résumé automatique	3
4.4	4. Système Q/R sémantique	3
4.5	5. Recommandation personnalisée	3
5	Conclusion	4

1 Introduction

Les podcasts sont devenus un format incontournable pour s'informer, apprendre ou se divertir. Pourtant, leur consommation peut être chronophage. D'où l'idée de concevoir un **assistant intelligent** capable de :

- Transcrire un podcast automatiquement
- Découper le contenu en chapitres thématiques
- Générer un résumé clair pour chaque section
- Répondre à des questions posées sur le podcast
- Recommander d'autres podcasts similaires

Le but du projet est de réaliser toutes ces tâches **en construisant nos propres modèles IA**, sans simplement utiliser des API pré-entraînés (comme Whisper ou GPT).

2 Concept global du projet

Ce projet est divisé en 5 modules majeurs, reliés entre eux et s'intègrent dans une application web.

1. **Transcription vocale** : Transformer l'audio en texte (reconnaissance vocale)
2. **Chapitrage** : Détecter les parties du podcast et leur donner un titre
3. **Résumé automatique** : Créer un texte court pour chaque chapitre
4. **Système de questions/réponses** : Permettre à l'utilisateur de poser des questions sur le contenu
5. **Recommandation de podcasts** : Proposer des épisodes similaires basés sur le thème, le ton, etc.

3 Fonctionnalités de l'application

L'utilisateur pourra :

- Importer un fichier audio ou lien YouTube
- Lire la transcription complète
- Visualiser les chapitres avec minute + titre
- Lire les résumés par section
- Poser des questions librement ("De quoi parle la 2e partie ?")
- Voir les podcasts recommandés automatiquement

4 Travail IA/ML à réaliser par nous-mêmes

Nous avons choisi de **construire nos propres modèles** à partir de datasets ouverts.

4.1 1. Modèle de transcription vocale

- Dataset: **Common Voice**, **LibriSpeech**
- Objectif: Entraîner un modèle de type CNN + CTC ou petit Wav2Vec2
- Evaluation: Word Error Rate (WER)

4.2 2. Chapitrage automatique

- Dataset: **PodcastSum** (transcription + chapitres)
- Objectif: Entraîner un modèle pour détecter les transitions de sujet (LSTM ou clustering sur embeddings)
- Bonus: Modèle extractif pour générer le titre du chapitre

4.3 3. Résumé automatique

- Dataset: **PodcastSum**, **corpus YouTube manuel**
- Objectif: Entraîner un modèle Seq2Seq (ex: Transformer simple ou T5 réduit)
- Evaluation: ROUGE, BERTScore

4.4 4. Système Q/R sémantique

- Objectif: Entraîner un réseau siamois pour retrouver la bonne partie du texte
- Créer un générateur de réponse simple (GPT-2 ou LSTM)
- Ajouter du feedback utilisateur pour apprentissage

4.5 5. Recommandation personnalisée

- Dataset: **Spotify Dataset**, **ListenNotes scraping**
- Objectif: Entraîner un auto-encoder pour embeddings thématiques
- Comparer par similarité cosinus + clustering

5 Conclusion

Ce projet nous permet d'explorer toutes les compétences de l'IA moderne : traitement audio, NLP, systèmes de recommandation, apprentissage supervisé, et interface utilisateur. En construisant chaque module de A à Z, nous serons capables de présenter un projet ambitieux, utile et **entièrement personnalisé par notre équipe**.