Projet : Analyse de Sentiment et Modélisation de Thèmes des Commentaires YouTube

(Intégrer une lecture des émoji selon le temps disponible)

Comment utiliser l’api de youtube :

1. Avoir un compte google
2. Créer un projet et ses identitifiants
3. Autoriser le projet à utiliser l’api de youtube

API youtube :

Key = API\_KEY

Creation d’un environement virtuel :

python -m venv .venv

activation de l’environement virtuel

source .venv/bin/activate

installation des packages

pip install -r requirements.txt

POC sur l’extraction, collect, stockage des données

Poc sur l’analyse de sentiment avec model pre entrainé et modèle entrainé sur des données labellisé

Poc sur l’entrainemen t d’un modèle sur les données labellisées par le modèle pré-entrainé

Poc sur le topic modeling

Poc sur la restitution

Utiliser MLflow pour la reproductibilité, et le suivi des perfs du modèle et déployer le modèle

# BLOC 3 : ÉLABORATION ET CADRAGE DU PROJET DATA

**Problématique :**

* ~~Comment exploiter la valeur des commentaires youtube pour en faire un outil d’amélioration / en tirer des KPI et améliorer le service produit~~
* **Valorisation des Retours Clients via l'Analyse Automatisée**

Comment mettre en place un système d'analyse automatisé des commentaires YouTube pour extraire des insights permettant d'améliorer en continu le service produit ?

* **Optimisation de l'Expérience Client**  
  Comment transformer les retours qualitatifs des commentaires YouTube en données quantitatives pour améliorer l'expérience utilisateur et le support client ?

**Objectifs :**

Développer un outil permettant de tirer le maximum de valeur des commentaires laissés sur un vidéo YouTube. L’outils doit présenter un certain nombre de KPI pertinents et utiles pour l’amélioration du service. Cela doit être un outil utilisable pour tout utilisateur de YouTube, allant de l’influenceur au simple utilisateur occasionnel. Il doit permettre à l’utilisateur d’avoir une visibilité et une compréhension des retours de ses « spectateurs » (sous condition d’avoir une base de spectateurs importante).

**Livrables :**

* Liste de KPI
* Modèle d’analyse de sentiment
* Modèle de topic modeling
* Traitements sur les données
* Script de collecte
* Base de données pour l’analyse de sentiments
* Plateforme
* Documentation technique
* Cahier des charges

Cadre réglementaire

Les contraintes et les point de vigilance

Attention au phénomène des haters, mettre en place une modération si ce n’est pas déjà fait par YouTube.

Les enjeux RSE

1. Etat de l’art

* **Comprendre ce qui a déjà été fait** dans le domaine (techniques, outils, jeux de données).
* **Identifier les limites ou lacunes** des approches existantes.
* **Justifier tes choix** méthodologiques.
* **Proposer une amélioration, une innovation ou une approche originale.**

 **Introduction au domaine**

 **Revue des méthodes d'analyse de sentiment**

 **Revue des techniques de modélisation de thèmes (Topic Modeling)**

 **Outils et bibliothèques existants**

 **Études similaires / projets passés**

 **Critique / Limites des approches**

 **Positionnement de ton projet**

Les projets existants

Quels modèles

Quels traitements

Cassendra aws, stockage forma colonne

1 modèle pour l’analyse de sentiment propre à chaque vidéo ou à améliorer en ajoutant des données (positif négatif ou neutre / ou score de sentiment)

1 modèle de topic modeling propre à chaque vidéo

Outil d’analyse automatique : KPI, tableau de bord + génération d’un rapport

Faire attention à la réglementation RGPD, IA ACT … LES PROBLEMES D’ÉTHIQUE, DE CONFIDENTIALITÉ