

# Résumé de la Solution pour l'Amélioration des Réponses d'Assistance

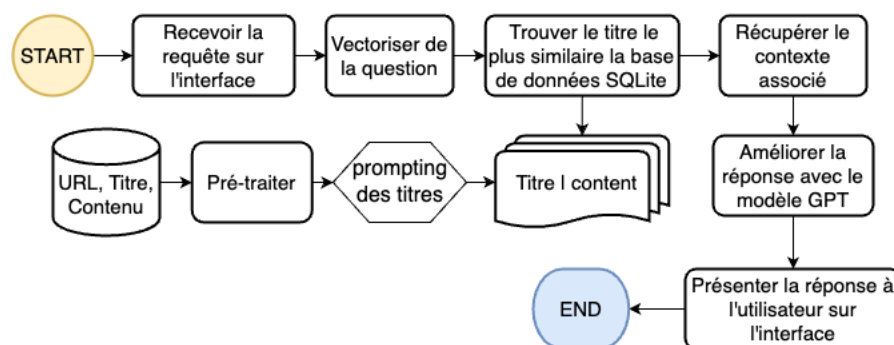
Yuyan QIAN

17/04/2024

**Objectif :** Améliorer la précision et la pertinence des réponses fournies par un chatbot dédié au support technique, en utilisant une base de données structurée d'interactions précédentes.

## Méthodologie :

1. **Compréhension des Données :** Utilisation de SQLite pour examiner et gérer une base de données contenant des interactions utilisateurs sous forme de URL, Titre et Contenu. L'attention se concentre spécialement sur les colonnes **titre** et **content**, où les titres suivent certains formats spécifiques, les contenus comportent des symboles spéciaux et les URLs fonctionnent comme identifiant unique. Le corpus est purement en français.
2. **Prétraitement des Données (*prompting + embedding*) :** Les titres sont d'abord contextualisés par prompting puis vectorisés via des modèles de sentence-transformer multilingues. Bien que des variantes Bert aient été testés, les meilleurs résultats sont obtenus avec des titres plus courts.
3. **Appariement et Récupération de Contenu :** La similarité cosinus entre la requête de l'utilisateur et les titres contextualisés est calculée pour identifier et récupérer le titre le plus proche.
4. **Amélioration de Réponse avec GPT :** Les réponses récupérées sont ensuite passées à travers un modèle de génération de texte basé sur GPT 3.5 turbo pour les reformuler et les enrichir, assurant une réponse plus naturelle et adaptée.
5. **Validation et Tests :** Le système est testé avec un ensemble de questions pour évaluer la précision des réponses générées et l'efficacité globale du système.



## Attendus :

- **Optimisation Continue :** Le système pourrait évoluer en intégrant les requêtes reçues et s'étendre avec des nouvelles données.
- **Amélioration de la Compréhension Contextuelle :** L'incorporation du contenu avec le titre dans l'embedding permet d'ajouter un contexte plus riche, mais cette étape testée introduit les bruits qui légèrement affectant l'appariement.
- **Fine-tuning avec un processus d'entraînement :** Le chatbot pourrait s'affiner via un processus d'entraînement continu en supervision, améliorant ainsi la précision et la capacité de réponse personnalisée. La classification des titres en fonction du mot-clé prédéfini sera utilisée dans cette étape.