**Développement d'un Chatbot Multilingue de Soutien pour le Cancer du Sein**

**Résumé**

Le cancer du sein est l’un des cancers les plus répandus et constitue une priorité mondiale en matière de santé publique. La détection précoce et la sensibilisation sont essentielles pour améliorer les résultats thérapeutiques et sauver des vies. Ce document présente la conception et la mise en œuvre d’un **chatbot interactif multilingue** dédié à fournir des informations fiables, un soutien émotionnel et une personnalisation des réponses pour répondre aux besoins des patients et des proches concernés par le cancer du sein. En utilisant des technologies de traitement du langage naturel (NLP) avancées telles que l'API GPT-3.5 d'OpenAI, combinées avec Streamlit pour une interface intuitive, ce projet établit une plateforme innovante pour l'éducation et le soutien en santé.

**1. Introduction**

Le cancer du sein reste une des principales causes de décès liés au cancer, affectant des millions de personnes chaque année. Face à cette réalité, l'éducation et le soutien psychologique jouent un rôle clé dans la gestion de cette maladie. Cependant, les patients et leurs proches rencontrent souvent des difficultés à trouver des informations fiables et adaptées à leurs besoins.

**Problématique**

Bien que des ressources existent, elles manquent généralement d’interactivité et de personnalisation. Les solutions existantes ne prennent pas en compte les préférences linguistiques des utilisateurs, ni leur besoin d’un soutien empathique.

**Objectif**

Le projet vise à concevoir un **chatbot multilingue** qui :

1. Fournit des informations précises sur le cancer du sein.
2. Répond aux besoins émotionnels et psychologiques des utilisateurs.
3. Permet une interaction dans plusieurs langues pour une meilleure accessibilité.
4. Offre une expérience utilisateur personnalisée grâce à des informations collectées via un formulaire.

**2. Architecture et Conception**

**2.1 Technologie Utilisée**

* **API GPT-3.5 d’OpenAI** : Gère la génération des réponses basées sur les requêtes utilisateur.
* **Streamlit** : Fournit une interface utilisateur simple, dynamique et réactive.
* **LangDetect** : Grâce à la bibliothèque langdetect, le chatbot identifie automatiquement la langue utilisée dans les requêtes utilisateur. Cela permet une réponse immédiate dans la langue détectée (anglais, français ou arabe), rendant le chatbot accessible à un public diversifié.
* **Python** : Langage principal pour l’intégration et le développement.

**2.2 Fonctionnalités Principales**

1. **Formulaire Personnalisé** : Collecte les informations de base telles que le nom, l'âge, le sexe et les symptômes. Ces données permettent de personnaliser l'expérience utilisateur.
2. **Suggestions de Questions** : Des boutons de suggestions permettent aux utilisateurs de poser des questions courantes (par exemple, "Quels sont les symptômes du cancer du sein ?").
3. **Multilingue** : Le chatbot répond dans la langue détectée ou choisie par l'utilisateur.
4. **Historique des Messages** : Affiche les conversations précédentes dans une interface claire.

**2.3 Flux de Travail**

Le fonctionnement du chatbot suit les étapes suivantes :

1. **Collecte des Informations** : L'utilisateur remplit un formulaire initial.
2. **Détection de Langue** : Le système identifie la langue utilisée grâce à LangDetect.
3. **Personnalisation de la Réponse** : Le chatbot génère des réponses adaptées au contexte, en tenant compte du nom et des données utilisateur.
4. **Affichage de l'Historique** : Toutes les conversations sont affichées dans une interface interactive.
5. **Entrée Persistante** : Un champ de saisie permanent permet d’envoyer facilement des requêtes.

**3. Implémentation Technique**

**3.1 Interface Utilisateur**

L'interface est conçue avec **Streamlit**, offrant :

* Suggestions cliquables : Offrent des questions prédéfinies en haut de la page.
* Historique des Messages : Liste les messages échangés dans une disposition verticale claire.
* Barre de Saisie Fixe : Permet à l'utilisateur de poser des questions supplémentaires.

**3.2 Détection de la Langue**

L'intégration de **LangDetect** permet d’identifier automatiquement la langue utilisée dans les requêtes utilisateur. Les réponses générées s’adaptent à cette langue. Cette fonctionnalité améliore l’accessibilité pour les non-anglophones.

**3.3 Génération de Réponses**

Les réponses sont générées par **GPT-3.5** à l'aide d'un modèle de traitement du langage naturel sophistiqué. L'intégration utilise l'API OpenAI et des instructions spécifiques pour inclure des informations personnalisées dans les réponses initiales.

**3.4 Personnalisation**

Le nom de l'utilisateur est utilisé uniquement dans le premier message pour éviter des répétitions inutiles. Une variable d'état vérifie si le nom a été mentionné.

**3.5 Calcul de la Similarité Cosinus**

Le chatbot utilise la **similarité cosinus** pour améliorer la pertinence des réponses. Une fois que le modèle GPT-3.5 génère plusieurs réponses possibles à une requête, elles sont converties en vecteurs numériques grâce au modèle text-embedding-ada-002.  
La similarité cosinus est ensuite calculée pour comparer ces vecteurs à celui de la requête utilisateur. La réponse ayant le score le plus élevé est sélectionnée comme étant la plus pertinente.  
Cette méthode garantit des réponses contextuellement alignées avec les attentes de l'utilisateur, réduisant les réponses génériques ou inadaptées.

**3.6 Gestion des Erreurs**

Si une requête utilisateur est ambiguë ou difficile à comprendre, le chatbot est conçu pour :

* S'excuser de manière empathique.
* Inviter l'utilisateur à reformuler ou à fournir plus de détails.  
  Cette fonctionnalité assure une interaction fluide et conviviale, même en cas de problèmes de compréhension.

**3.7 Journalisation (Logging)**

Un système de journalisation (logging) est intégré pour suivre :

* Les requêtes utilisateur.
* Les réponses générées par le modèle.
* Les scores de similarité cosinus pour chaque réponse.  
  Ces logs sont enregistrés dans un fichier, permettant une analyse approfondie des performances et des ajustements pour améliorer la précision des réponses.

**4. Résultats**

**4.1 Fonctionnalités Clés Testées**

* **Personnalisation** : L'utilisateur se sent directement impliqué grâce à la mention de son nom dans le premier message.
* **Précision Linguistique** : Le chatbot peut répondre avec précision en anglais, français et arabe.
* **Interaction Fluide** : L'utilisateur peut poser des questions en continu via une barre de saisie fixe.

**4.2 Exemple Pratique**

**Requête utilisateur :** "Quels sont les symptômes du cancer du sein ?"  
**Réponses générées :**

* R1 : "Les symptômes incluent des douleurs et des gonflements."
* R2 : "Consultez un médecin pour tout problème mammaire."
* R3 : "Les symptômes courants incluent des nodules, des douleurs, et des rougeurs."  
  **Score de similarité cosinus :**
* R1 : 0.89
* R2 : 0.65
* R3 : 0.92  
  **Résultat :** La réponse R3 est sélectionnée comme la plus pertinente.

**5. Améliorations Futures**

**5.1 Évolutions Techniques**

1. **Intégration Vocale** : Ajouter une fonctionnalité de reconnaissance vocale pour faciliter l’accès aux utilisateurs ayant des difficultés à taper.
2. **Base de Connaissances Médicales** : Intégrer une base de données pour garantir que les réponses soient toujours en conformité avec les dernières recherches médicales.
3. **Analyse des Émotions** : Détecter et adapter les réponses aux émotions de l’utilisateur (exemple : anxiété ou tristesse).

**5.2 Optimisation de l'Expérience Utilisateur**

* Réduire les temps de latence pour les réponses générées.
* Ajouter des animations visuelles pour enrichir l’interactivité.

**6. Conclusion**

Ce projet démontre comment l’intelligence artificielle conversationnelle peut être utilisée pour sensibiliser et soutenir les patients concernés par le cancer du sein. Grâce à une interface simple, un support multilingue et des capacités de personnalisation avancées, ce chatbot représente une solution innovante et accessible. Avec les améliorations proposées, il a le potentiel d’évoluer en un outil de référence pour l’éducation et le soutien en santé.

Avec l'intégration de la similarité cosinus, de la gestion des erreurs, et des capacités multilingues, ce chatbot offre une expérience utilisateur enrichie, garantissant des réponses adaptées aux besoins spécifiques des utilisateurs.

**Références**

1. OpenAI. “Documentation GPT-3.5,” 2024.
2. Streamlit. “Framework Documentation,” <https://streamlit.io>.