

Projet NLP

Projet 1 : Création d'un Chatbot FAQ Utilisant des Données Médicales ou d'autres données de votre choix

Description : Ce projet vise à développer un chatbot FAQ (Foire Aux Questions) qui peut répondre aux questions médicales courantes des utilisateurs. Vous utiliserez des techniques de traitement automatique du langage naturel (NLP) pour entraîner le chatbot sur un ensemble de données médicales.

Étapes du projet :

1. Collecte de données :

- Vous collecterez des ensembles de données de questions-réponses médicales à partir de sources telles que des forums médicaux en ligne, des FAQ de sites Web médicaux, ou des chatbots médicaux existants.
- Exemples de datasets :
 - Le dataset "COVID-19 Open Research Dataset (CORD-19)" contient des articles de recherche sur le COVID-19, qui peuvent être utilisés pour répondre aux questions sur le virus.
 - Le dataset "Medical Information Mart for Intensive Care (MIMIC)" contient des données médicales provenant de patients admis en soins intensifs, utiles pour répondre aux questions sur les conditions médicales graves.

2. Nettoyage et prétraitement des données :

- Les données brutes seront nettoyées pour éliminer les balises HTML, les liens, les informations personnelles, etc.

- Vous effectuerez des tâches de tokenization, de lemmatisation, de suppression des mots vides etc.

3. Construction du modèle de chatbot :

- Vous choisirez une architecture de modèle de chatbot, telle qu'un modèle séquentiel de réseaux de neurones récurrents (RNN) ou LSTM ou encore un modèle à transformer comme BERT, etc.
- Vous entraînerez le modèle sur les données de questions-réponses médicales pour qu'il puisse répondre aux questions des utilisateurs.

4. Évaluation du chatbot :

- Les performances du chatbot seront évaluées en utilisant des métriques telles que la précision des réponses, la couverture des questions, la pertinence des réponses, etc.

5. Optimisation du modèle :

- Vous pourrez expérimenter avec différentes architectures de modèle, différentes techniques de prétraitement, ou différentes stratégies d'apprentissage pour optimiser les performances du chatbot.

6. Déploiement du chatbot :

- Une fois le chatbot entraîné et évalué, déployez le sur une application streamlit pour la demo.

Livraison du projet : Vous présenterez le projet sous forme d'une présentation powerpoint, décrivant chaque étape du processus, ainsi que les résultats obtenus. Vous pourrez également présenter votre chatbot lors d'une démonstration en direct, en interagissant avec lui pour montrer ses capacités et ses performances.

Projet 2 : Développement d'un Chatbot de Support Technique

Description : Ce projet consiste à développer un chatbot de support technique capable de répondre aux questions des utilisateurs concernant les problèmes techniques courants. Vous utiliserez des techniques de traitement automatique du langage naturel (NLP) pour entraîner le chatbot sur un ensemble de données de support technique.

Étapes du projet :

1. Collecte de données :

- Vous collecterez des ensembles de données de support technique à partir de forums en ligne, de bases de connaissances techniques, de tickets de support, ou d'autres sources similaires.

Exemple de dataset :

- Stack Overflow : un ensemble de données de questions et réponses sur des problèmes de programmation et informatiques.
- Dataset d'Amazon Alexa Reviews : contient des commentaires sur les fonctionnalités et performances des appareils Amazon Alexa.

2. Nettoyage et prétraitement des données :

- Les données brutes seront nettoyées pour éliminer les balises HTML, les liens, la ponctuation, etc.
- Vous effectuerez des tâches de tokenization, de lemmatisation et de suppression des mots vides, etc.

3. Construction du modèle de chatbot :

- Une architecture de modèle de chatbot sera choisie, telle qu'un modèle de réseaux de neurones récurrents (RNN), un modèle de transformers, ou une combinaison de ces deux.
- Le modèle sera entraîné sur les données de support technique pour qu'il puisse répondre aux questions des utilisateurs.

4. Évaluation du chatbot :

- Les performances du chatbot seront évaluées en utilisant des métriques telles que la précision des réponses, la pertinence des réponses, etc.

5. **Optimisation du modèle :**

- Vous pourrez expérimenter avec différentes architectures de modèle, techniques de prétraitement, ou techniques d'apprentissage pour améliorer les performances du chatbot.

6. **Déploiement du chatbot :**

- Une fois le chatbot entraîné et évalué, déployez le sur une application streamlit pour la demo.

-

Livraison du projet : Vous présenterez le projet sous forme d'une présentation powerpoint, décrivant chaque étape du processus, ainsi que les résultats obtenus. Vous pourrez également présenter votre chatbot lors d'une démonstration en direct, en interagissant avec lui pour montrer ses capacités et ses performances.

NB : Un repo github avec description sur le projet, l'entrainement du modèle et comment lancer le chatbot.