

Détection d'intention dans un chatbot

Data Science

Contexte

ILLUIN développe, à travers son produit ILLUIN Dialogue, une suite conversationnelle permettant de créer et gérer des *chatbots* via différents canaux: site web, messagerie instantanée, serveur vocal, application mobile...

Dans la conception d'un chatbot, l'une des premières étapes de traitement d'un message utilisateur (un *verbatim*) consiste à classifier l'intention exprimée avec un label prédéfini qui permet dans un second temps de déclencher un parcours adapté.

Cette étape de classification est appelée *détection d'intention*.

Problème

Proposez un algorithme permettant, étant donné un *verbatim* en entrée, de classifier l'intention exprimée. Nous nous placerons dans le cas d'un chatbot d'assistance sur le site d'une grande agence de tourisme. Les différentes classes sont les suivantes :

translate: l'utilisateur souhaite traduire une phrase dans une autre langue

travel_alert: l'utilisateur demande si sa destination est concernée par une alerte de voyage

flight_status: l'utilisateur demande des informations sur le statut de son vol

lost_luggage: l'utilisateur signale la perte de ses bagages

travel_suggestion: l'utilisateur souhaite une recommandation de voyage

carry_on: l'utilisateur souhaite des informations sur les bagages à main

book_hotel: l'utilisateur souhaite réserver un hôtel

book_flight: l'utilisateur souhaite réserver un vol

L'agence vous signale que l'intention ***lost_luggage*** redirige vers un service client avec des conseillers téléphoniques ayant un coût élevé.

De plus, il est possible que l'utilisateur formule une demande dite "hors-scope" c'est-à-dire qui n'appartient à aucune des 8 classes ci-dessus. Dans ce cas, on attend que l'algorithme renvoie une réponse spéciale ***out_of_scope***.

Vous disposez de quelques exemples de verbatims annotés, y compris des demandes hors-scope, disponibles dans [ce fichier](#).

Objectif

L'objectif est de proposer une approche pertinente que nous pourrions challenger ensemble lors d'un entretien physique. Il n'y a pas une bonne réponse mais de nombreuses possibilités d'arriver au résultat. En plus de la pertinence des approches utilisées, la qualité du code sera également évaluée.

Livrable

- Choisir une ou plusieurs métriques permettant d'évaluer la qualité de l'algorithme.
- Module Python exécutable à l'aide d'une CLI permettant d'effectuer une prédiction et une évaluation sur un dataset de test au format CSV. L'utilisation de conventions Python (typing, classes) sera appréciée.
- La prédiction pour un verbatim doit pouvoir être effectuée en un temps raisonnable (ordre de grandeur de la seconde maximum)
- La comparaison de plusieurs algorithmes est encouragée