Titre du projet : Annotation de textes médicaux

Type de projet : (  ) recherche      (X) développement

Résumé du projet (300 mots max) :

Ce projet se déroule dans le cadre d’un projet de recherche avec le Centre de Pharmacovigilance du CHU de Nice qui a pour objectif d’extraire des rapports médicaux les effets indésirables médicamenteux qu’ils décrivent.

Une part du prototype est réalisé sur la base de rapport médicaux en anglais et déjà annoté. C’est à dire que l’on sait déjà quels sont les mots qui représentent un symptôme et ceux qui représente un médicament.

L’objectif du travail est de réaliser un prototype permettant d’annoter des rapports médicaux issu du CHU et donc en français afin de savoir, comme pour les rapports médicaux en anglais quels sont les mots qui décrivent un médicament et ceux qui décrivent un symptôme. Pour faire ce travail, les pistes de départ sont les suivantes :

* Utiliser des dictionnaires qui contiennent les noms des médicaments et des symptômes
* Utiliser un système d’annotation existant via une API (http://bioportal.lirmm.fr) （选择ontology：medic，symp、结果分类）
* Utiliser un réseau de neurones pré-entrainé (<https://spacy.io>)

Et très certainement, combiner ces différentes approches pour obtenir un résultat proche de la perfection.

En complément, il s’agit aussi de développer un petit prototype permettant de visualiser les annotations produites.

Mots-clés (5 mots qui caractérisent la thématique et le scope du projet) :

Outil d'annotation, reconnaissance d'entités nommées, dossiers médicaux électroniques, traitement du langage naturel.

Références bibliographiques :

|  |  |
| --- | --- |
| [1] | B. Dura, P. Wajsburt, T. Petit-Jean, A. Cohen, C. Jean and R. Bey, *EDS-NLP: efficient information extraction from French clinical notes,* 2022. |
| [2] | H.-J. Song, B.-C. Jo, C.-Y. Park, J.-D. Kim and Y.-S. Kim, "Comparison of named entity recognition methodologies in biomedical documents," *BioMedical Engineering OnLine,* vol. 17, p. 158, November 2018. |
| [3] | P. D. Soomro and S. Kumar, "Bio-NER: Biomedical Named Entity Recognition using Rule-Based and Statistical Learners," *International Journal of Advanced Computer Science and Applications,* vol. 8, p. 8, 2017. |
| [4] | A. Tchechmedjiev, A. Abdaoui, V. Emonet, S. Zevio and C. Jonquet, "SIFR annotator: ontology-based semantic annotation of French biomedical text and clinical notes," *BMC Bioinformatics,* vol. 19, p. 405, November 2018. |

Nom et affiliation de l’encadrant :

ZHANG Xuchun, doctorant Univ Côte d’Azur, [xuchun.zhang@inria.fr](mailto:xuchun.zhang@inria.fr)

Co-encadrant(s) :

RIVEILL Michel, PR Univ Côte d’Azur, michel.riveill@univ-cotedazur.fr

commentaires！

打分：precision、recall、f1-score

precision：resultat bien?

Recall: combien trouve