

Travaux dirigés

Introduction NLP

Objectif :

À l'aide d'un Notebook, produire les analyses suivantes portant sur un jeu de données d'environ 16 000 tweets dont le sujet tourne autour des vaccins de la Covid19 :

- Réaliser une brève analyse descriptive du jeu de données ;
- Identifier les tweets les plus influents ;
- Comparer les différents moteurs d'analyse de sentiments :
 - TextBlob
 - nltk
 - vader
 - sklearn
 - allennlp
- Réaliser une ou plusieurs analyses temporelle par type de vaccin avec un axes d'analyse de type : appréciation (résultat de l'analyse sentimentale), nombre de tweet, nombre de retweet, etc) ;
- Réaliser une analyse NER sur le champ text et par vaccin. Visualiser le résultat en créant un nuage de mots par type d'entité ;
- Analyse géographique : réaliser une cartographie représentant par vaccin l'appréciation de ces derniers en fonction de leur opinion par pays/région.

Références :

[Kaggle](#)

Exemple d'analyse : [Sentiment Analysis of COVID-19 Vaccine Tweets](#)

[Kaggle dataset](#)