

# ANALYSE DES AVIS DE FOURNISSEURS D'ENERGIE

Hela Dellagi, Ayoub Ben Chaliah, Joslin Martinez, Nicolas Bourgain



Tuteur: Emmanuel Haback



## Contexte du projet



Les sites d'avis contiennent des informations précieuses pour les entreprises. Ces dernières peuvent ainsi connaître leurs forces & faiblesses auprès des consommateurs et se comparer avec la concurrence.

## Récupération des données



Afin de pouvoir analyser les données du site Trustpilot.com nous avons réalisé un scraping des données. Nous avons retenu quatre paramètres : les noms des fournisseurs d'énergie, le corps de l'avis, la date de l'avis et la notation sous forme d'étoiles sur cinq.

## Nettoyage des données



Afin de faciliter et rendre possible l'analyse nous avons nettoyé les données en plusieurs étapes. Tout d'abord nous avons supprimé la ponctuation, les accents ainsi que les smileys puis nous avons effacé les mots jugés non essentiels.



## Analyse des sentiments avec Vader et LDA



Pour exploiter les avis nettoyés nous avons recours à Vader qui classe les avis selon un score qui reflète son caractère positif, négatif ou neutre. Mais pour savoir de quoi parle exactement les utilisateurs de trustpilot.com nous utilisons également le modèle LDA pour dégager les thèmes les plus souvent cités dans les avis.

## Deep Learning avec BERT



Pour aller plus loin que Vader dans l'analyse, nous utilisons également du Deep Learning avec le modèle BERT. Ce dernier offre la possibilité de traiter les mots non pas un par un mais par groupes entiers qui sont donc plus faciles à interpréter.

## Contenu de notre dashboard



Notre dashboard contient des graphes interactifs de favorabilité, de proportion d'avis, de sentiment par topic et de topics principaux selon le fournisseur d'énergie. Nous avons ajouté à la fin du dashboard un tableau contenant une multitude de filtres pour naviguer dans les données.