

# Analyse de sentiments

## Présentation du projet

Pour une entreprise, il est important de bien comprendre ses clients. Il n'est pas toujours évident de savoir ce que vos clients pensent de votre entreprise. Aujourd'hui, grâce aux réseaux sociaux, les utilisateurs donnent leur ressenti de façon libre.

Une bonne application est d'utiliser ces données pour prendre la température des réseaux sociaux à propos de votre entreprise. Le but est de savoir si vos clients parlent en bien ou en mal de vos entreprises et de vos produits. Cette température peut ensuite être vue comme un indicateur à améliorer via des campagnes de communication ou une amélioration des offres ou des produits.

L'application est assez simple, il vous suffit d'entraîner un modèle d'analyse de sentiments. Pour chaque tweet, votre modèle doit être capable de déterminer le sentiment de la personne l'ayant écrit.

## Consignes

- Créer un répertoire [Github](#) et sauver votre travail régulièrement.
- Utiliser ce jeu de [données](#) pour l'entraînement.
- Utiliser une architecture de deep learning utilisant les RNN.
- Utiliser une technique de word embedding pour la transformation des mots en vecteurs.
- Créer une pipeline qui prend en entrée une phrase brute, qui effectue tous les traitements et prédictions.
- Utiliser l'[API yelp](#) pour tester votre pipeline.

## Conseils

- Dans un premier temps, prenez seulement une petite partie des données pour réduire le temps de calcul.
- Réutilisez au maximum les exemples vus dans les TPs précédents.
- N'hésitez pas à chercher de l'aide sur internet.

# Rendus

Un notebook :

- disponible sur votre Github,
- un sommaire doit être mis en place,
- chaque étape doit être justifiée en utilisant le markdown,
- Le code doit être clair et commenté

Un rapport sera rédigé reprenant :

- le contexte, le sujet du projet,
- les données et l'approche utilisées,
- les détails de votre solution,
- les améliorations possibles de cette version.

Le rapport ne devra pas faire plus de 10 pages.

Le rapport et le notebook seront à rendre par email à l'adresse suivante au plus tard le 28 février à minuit.