

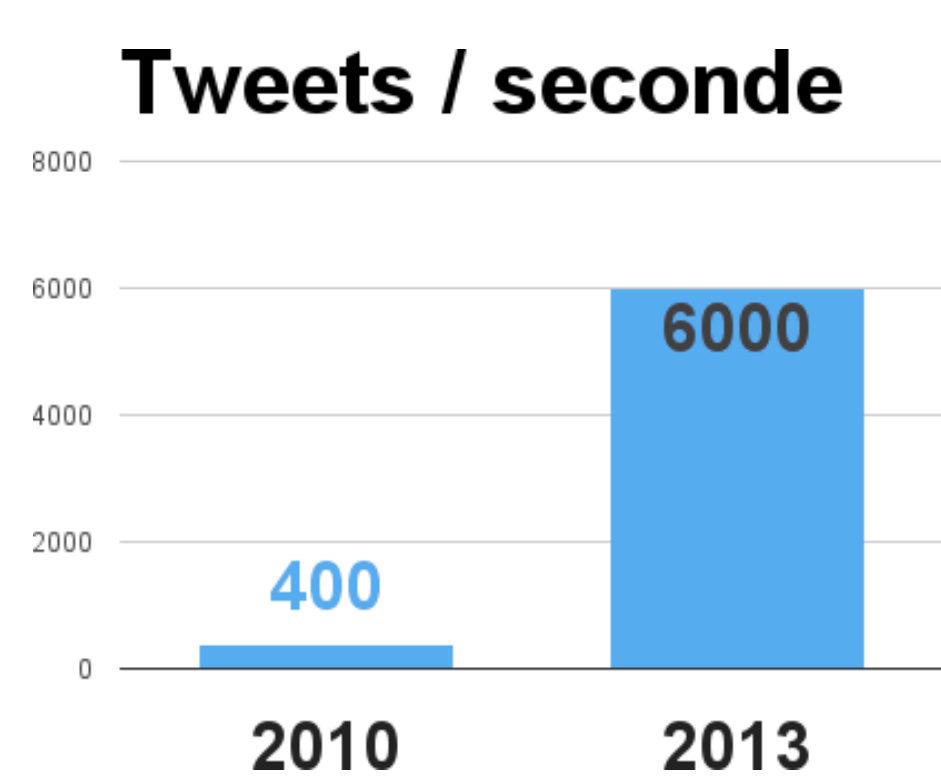
Intégration du contexte spatio-temporel et social pour l'analyse de sentiments sur Twitter

Ophélie Fraisier

Institut de Recherche en Informatique de Toulouse
CEA Tech Midi-Pyrénées

✉ ophelie.fraisier@irit.fr

Contexte



Depuis quelques années, les médias sociaux permettent à tout le monde de s'exprimer publiquement sur le web.

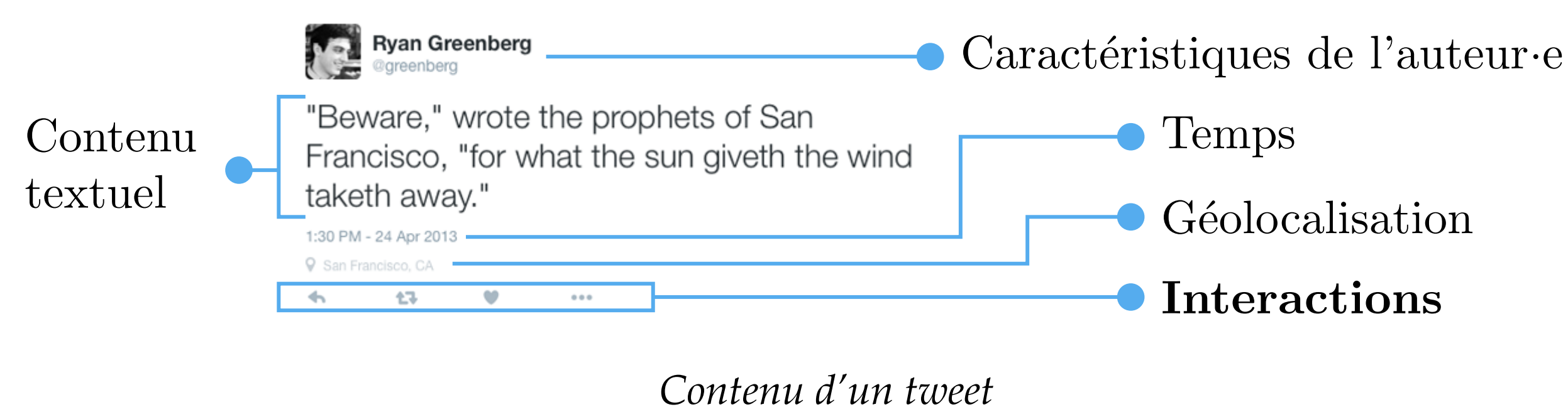
Ces données présentent un intérêt certain pour de nombreux domaines : marketing, politique, cybersécurité, ...

Plusieurs outils et méthodes permettent d'extraire les sentiments du texte :

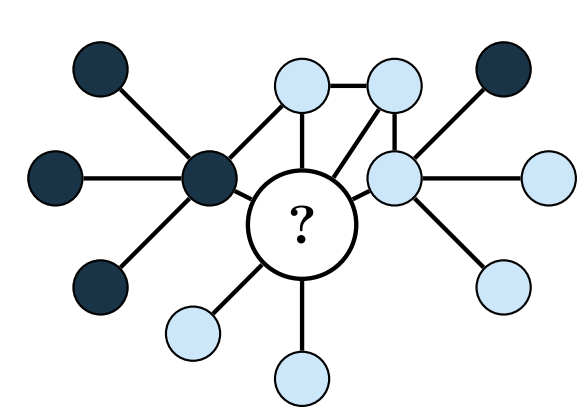
- Traitement automatique des langues
- Lexiques spécialisés
- Méthodes d'apprentissage automatique

Caractéristiques de Twitter

- Plateforme de microblogs la plus connue et la plus utilisée
- Tweet : message de 140 caractères maximum
- Abonnements possibles entre utilisateurs·trices
- Réactions possibles aux tweets :
 - Réponse
 - Retweet : diffusion du tweet à ses propres abonné·e·s
 - Placement du tweet en favori



Problématique



Les modèles classiques de détection des sentiments se concentrent sur les caractéristiques textuelles. Les interactions entre utilisateurs·trices – composante essentielle des réseaux sociaux – ne sont pas exploitées. Or l'hypothèse d'homophilie stipule que les personnes appartenant aux mêmes communautés ont des points de vue similaires.

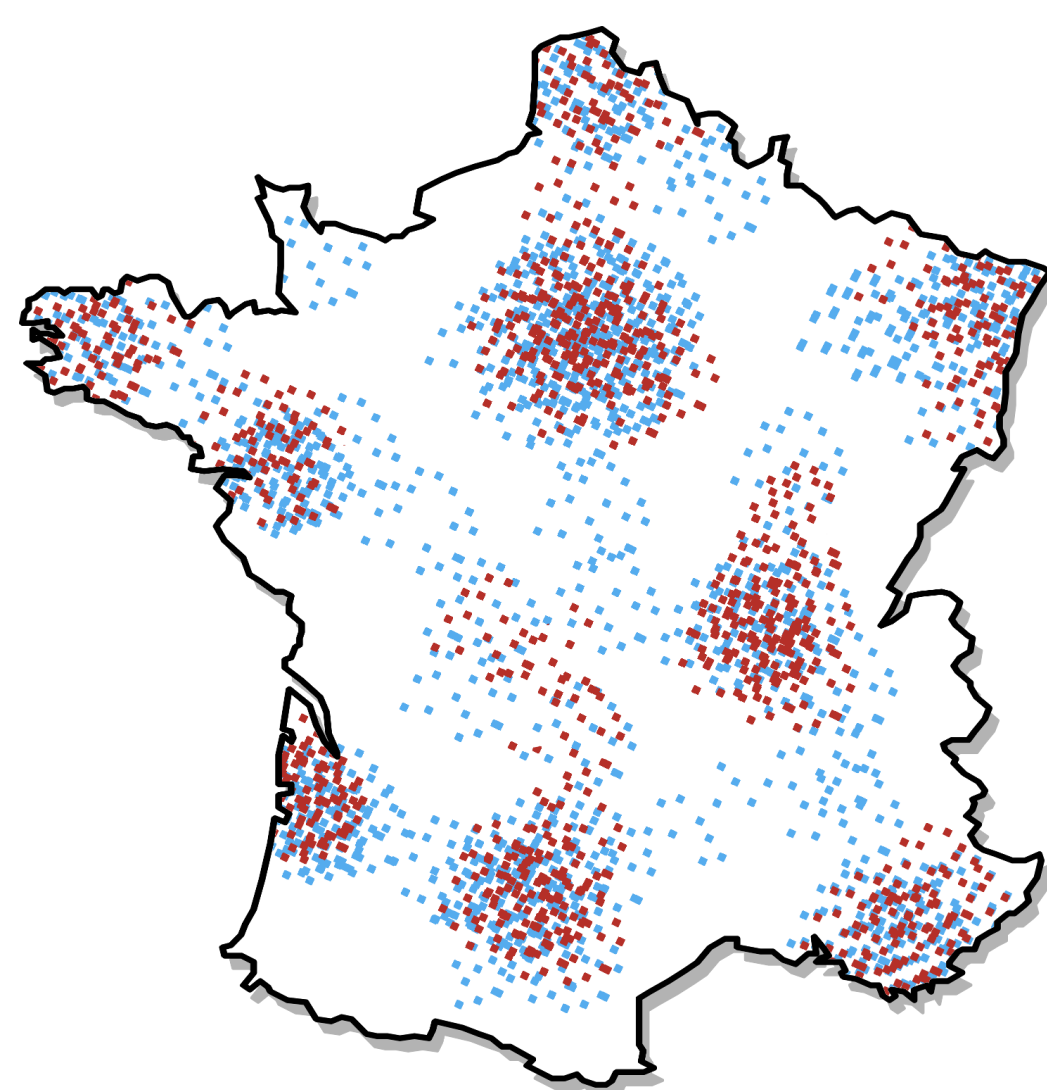
Notre modèle propose donc :

1. d'utiliser les interactions sociales pour détecter les communautés
2. de déterminer ainsi le point de vue de l'auteur·e afin d'améliorer la détection des sentiments exprimés dans ses tweets.

Agrégation des sentiments

Les éléments contextuels seront également utiles pour agréger les sentiments :

- géolocalisation : "carte des sentiments"
- temps : le modèle doit prendre en compte le fait qu'une personne puisse changer d'avis
- caractéristiques de l'auteur·e : sexe, âge, catégorie sociale, personnalité, ...



Contribution principale

On peut distinguer 2 types de sentiments :

Explicites : portés par des mots exprimant une émotion (aimer, merveilleux, terrible, ...)

Implicites : beaucoup plus difficiles à détecter car dépendant du contexte plus que des termes employés

"La loi Travail est considérée comme adoptée par l'Assemblée Nationale."

- Sans contexte : pas de sentiment particulier
- Si l'auteur·e est opposé·e à la loi Travail : sentiment négatif

L'utilisation du point de vue de l'auteur·e permettra de mieux détecter les sentiments implicites, contrairement à un modèle basé uniquement sur le texte.

Références

Liu, B. (2012). Sentiment Analysis and Opinion Mining. *Synthesis Lectures on HLT*, 5(1), 1–167. doi: 10.2200/S00416ED1V01Y201204HLT016

Tang, J., Chang, Y., & Liu, H. (2014). Mining social media with social theories: a survey. *ACM SIGKDD Explorations Newsletter*, 15(2), 20 – 29. doi: 10.1145/2641190.2641195

Remerciements : Ce travail a été réalisé grâce à l'obtention d'un financement Contrat Laboratoire-Entreprise numéro 14050975 soutenu par la Région Languedoc-Roussillon Midi-Pyrénées.