

Analyse des sentiments à partir des commentaires Facebook publiés en Arabe standard ou dialectal marocain par une approche d'apprentissage automatique

Abdeljalil Elouardighi*, Mohcine Maghfour*, Hafdalla Hammia*, Fatima-Zahra Aazi**

* Laboratoire de Modélisation Mathématiques et de Calculs Economiques
Faculté des Sciences Juridiques Economiques et sociales, Université Hassan 1^{er},
Km 3, route de Casablanca, B.P. : 784, Settat, Maroc.
abdeljalil.elouardighi@uhp.ac.ma, m.maghfour.@gmail.com, hhammia@gmail.com

** ESCA Ecole de Management,
7, Rue Abou Youssef El Kindy, 20 070 Casablanca, Maroc.
faazi@esca.ma

Résumé. L'analyse des sentiments est un processus pendant lequel la polarité (positive, négative ou neutre) d'un texte donné est déterminée. Nous nous intéressons dans ce travail à l'analyse des sentiments à partir des commentaires Facebook, réels, partagés en arabe standard ou dialectal marocain par une approche basée sur l'apprentissage automatique. Ce processus commence par la collecte des commentaires et leur annotation à l'aide du crowdsourcing suivi d'une phase de prétraitement du texte afin d'extraire des mots arabes réduits à leur racine. Ces mots vont être utilisés pour la construction des variables d'entrée en utilisant plusieurs combinaisons de schémas d'extraction et de pondération. Pour réduire la dimensionnalité, une méthode de sélection de variables est appliquée. Les résultats obtenus des expérimentations sont très prometteurs.

1 Introduction

L'analyse des sentiments (AS) devient un domaine d'étude très ouvert à la recherche. L'objectif étant d'analyser, à partir des textes partagés sur les réseaux sociaux, les opinions, les sentiments, les attitudes et les émotions des communautés sur différents sujets. En général, on distingue deux catégories d'approches pour l'AS des textes publiés sur les réseaux sociaux : la première est basée sur le lexique, consiste à utiliser une collection prédéfinie de mots et d'annoter chacun avec une valeur traduisant sa polarité (sentiment positif, négatif ou neutre). La deuxième est basée sur des techniques d'apprentissage automatique. L'AS dans ce cas peut être vu comme étant un problème de classification supervisée de texte. Sur les réseaux sociaux, comme Facebook, les commentaires partagés en arabe prend généralement la forme de l'Arabe Standard Moderne (ASM) ou l'Arabe dialectal (Duwairi et Qarqaz, 2014). Nous nous intéressons dans ce travail à l'AS à partir des commentaires Facebook écrits en ASM ou en Arabe Dialectal Marocain (ADM) en utilisant une approche basée sur l'apprentissage automatique.