

Modèles prédictifs pour la détection précoce des canulars diffusés sur Twitter

Didier Henry*, Erick Stattner*

* LAMIA, Université des Antilles, Guadeloupe, France
didier.henry ; erick.stattner@univ-antilles.fr

1 Introduction

De nos jours, les médias sociaux sont largement utilisés quotidiennement pour accéder aux nouvelles. Malheureusement, ces plateformes utilisées par des millions de personnes ne sont pas à l'abri de la désinformation. Les rumeurs de la mort des célébrités sur les médias sociaux se propagent très largement dans un court laps de temps et sont difficilement vérifiables. Ces types de rumeurs peuvent conduire à des situations inquiétantes ou stressantes. Dans ce travail, nous avons abordé le problème de détection des canulars sur la mort de célébrités dans les médias sociaux.

2 Collecte et extraction de données

Notre objectif est d'observer et de comprendre les différences entre les vraies et les fausses rumeurs de mort de célébrités au niveau de la diffusion, des tweets et des utilisateurs. Le principal défi que nous avons rencontré était d'accéder aux tweets sur ces rumeurs. Grâce au moteur de recherche de Twitter¹ nous avons collecté un grand nombre de tweets liés à 25 rumeurs relatives à la mort de célébrités, dont 13 réelles (c'est-à-dire que la célébrité est vraiment morte) et 12 fausses (c'est un canular sur la mort de la célébrité).

Dans un premier temps, nous avons pris en compte plusieurs caractéristiques de diffusion au moment t_n (le nombre de tweets postés, le nombre de retweets publiés, le pourcentage de tweets publiés, le pourcentage de retweets publiés, le pourcentage de nouveaux messages, un booléen égal à vrai si la rumeur se propage aussi en français et en espagnol en moins d'une heure)

Dans un deuxième temps, nous nous sommes intéressés aux caractéristiques des messages. Tout d'abord, nous effectuons une analyse du contenu des messages au moment t_n (le nombre moyen de caractères des messages, le pourcentage de messages contenant une URL et une URL distincte (pointant vers des sites autres que Twitter), le pourcentage de messages contenant l'acronyme RIP, le pourcentage de messages avec au moins un smiley triste, etc.). Deuxièmement, nous avons utilisé TextBlob² une API en langage Python afin d'extraire les sentiments

1. <https://twitter.com/search-advanced>

2. <http://textblob.readthedocs.io/en/dev/>