



Master Bioinformatique

HAI817 – Machine learning 1 (méthodes classiques)

Professeurs : P. Poncelet, K. Todorov, E. Raoufi

Classification d'assertions venant d'X (Twitter) selon leur rapport à la science

Par : Tiziri Tamani (22415178), Ciriàn Mahony (22400729), Raphaël Ribes (22401925), Dalia Belmadi (22208849)

Lien vers le projet : <https://github.com/RaphaelRibes/machine-learning/>

Table des matières

1- Introduction	1
2- Matériel et Méthodes	2
3- Résultats	3
4- Discussion	4
5- Conclusion	5
6- Références	6
7- Annexes	7

1- Introduction

2- Matériel et Méthodes

On regarde le nombre de #, @ et lien par tweet selon le type de tweet scientifique (figure 1).

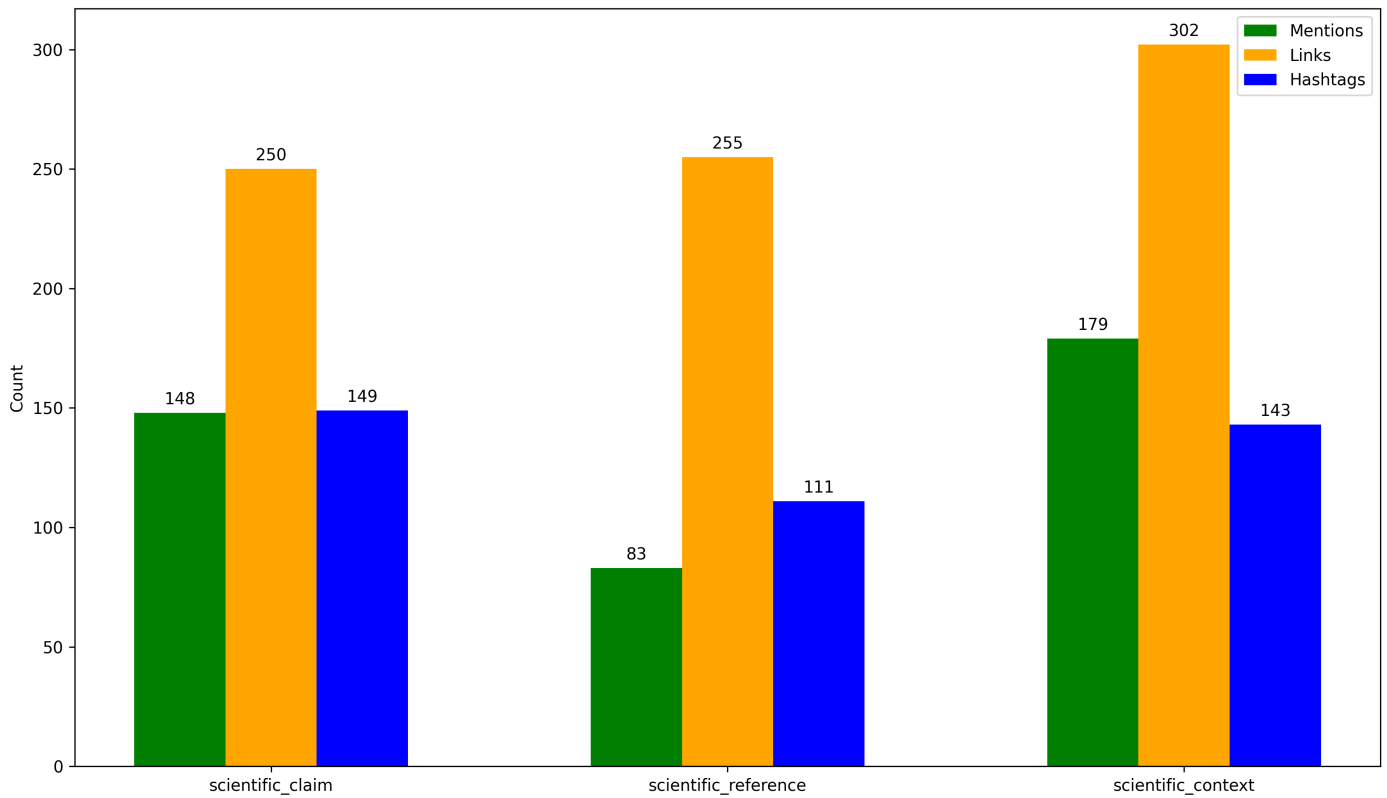


FIGURE 1 – Nombre de hashtags par tweet selon le type de tweet scientifique.

Après avoir réalisé des tests de student sur chaque variable, les *p-values* ne descendent pas en dessous de 0.18 donc on ne remarque pas de différence significative entre les tweets scientifiques et non scientifiques. On peut ainsi conclure que ces variables ne sont pas pertinentes pour la classification des tweets scientifiques, on les retirera de notre dataset.

3- Résultats

4- Discussion

5- Conclusion

6- Références

7- Annexes