Adverse event detection in Twitter: results from the WEB-RADR consortium

Lucie M. Gattepaille, Sara Hedfors Vidlin, Tomas Bergvall, Johan Ellenius
Uppsala Monitoring Centre, Uppsala, Sweden
Lucie.gattepaille@who-umc.org

1 Introduction

With hundreds of millions of active users sharing openly their thoughts and experiences, Twitter has the potential to be a useful resource for post-marketing surveillance of medicines, complementing by its unsolicited nature, timeliness and breadth of patient coverage the traditional pharmacovigilance tools such as spontaneous case reports databases [Sarker *et al*., 2015]. However, whether this type of data can be harnessed and reliably utilized for pharmacovigilance remains an open question.

Supported by the Innovative Medicines Initiative, the WEB-RADR consortium aimed, among other things, at investigating the potential of social media data for identifying medicine safety issues [WEB-RADR]. We present the results of this investigation on Twitter data.

2 Methods

A computational processing pipeline was developed for detecting and normalizing medicinal products, medical events and characterizing their relationship as Adverse Events (AE) or not. First, a relevance filter of tweets was applied for removing posts with low probability of containing an AE, using a previously published method [Freifeld *et al*., 2014]. Medicinal product detection used a dictionary lookup method based on WHODrug, a global dictionary of medicinal products. Medical event detection used a collection of dictionary lookups, natural language processing methods and logistic regression classifiers to normalize the identified text to the Medical Dictionary for Regulatory Activities (MedDRA) preferred terms. Consequently, a logistic regression classifier based on a variety of features (document or context based, relational, syntactic and semantic from Word2Vec [Mikolov *et al*., 2013] embeddings) was trained to identify AE relationships.

The pipeline was trained on a proprietary dataset of 196,533 tweets manually curated and annotated by trained MedDRA coders at Epidemico, a WEB-RADR consortium member, and evaluated on an independent benchmark Reference dataset of 57,481 manually annotated tweets produced by WEB-RADR. We present comparative performance results of the pipeline on the WEB-RADR Reference dataset with a previously published method [Freifeld *et al*., 2014] that has reported good performance on other datasets.

3 Results

The Reference dataset contains 1,058 posts with at least one medicinal product - AE pair (MP/AE pair) and a total of 1,398 MP/AE pairs. Our pipeline could recognize 316 of the pairs while identifying 597 false positive pairs, yielding thus a recall of 0.23 and precision of 0.35 (F-score 0.27). 498 MP/AE pairs were lost in the relevance filter module, 410 in the medicinal product and medical event detection modules combined and 174 in the AE relation classification module, leading to marginal recalls of 0.64, 0.54 and 0.64 respectively. In total, 2,323 MP/medical event pairs were submitted for AE relation classification, of which 1,833 were non-AE pairs. Comparatively, the previously published method, which reports 0.86 recall and 0.72 precision (F-score 0.78) in the publication, achieved a recall of 0.32 and a precision of 0.18 (F-score 0.23) on the Reference dataset.

4 Discussion

This study highlights two major difficulties with developing methods of automatic detection of adverse events in Twitter posts: 1) detecting and normalizing medical events is a challenge in Twitter posts, probably due to the noisiness of the data (e.g. misspellings, abbreviations, diversity of laymen expressions), 2) the transferability of models outside the universe of the training data to external datasets is poor, despite the use of a training/validation/test setup. The latter difficulty is poorly understood and should be the object of more research, to investigate the true ability of AE detection algorithms to harness social media data.

References

[Sarker *et al*., 2015] Abeed Sarker, Rachel Ginn, Azadeh Nikfarjam, Karen O'Connor, Karen Smith, Swetha Jayaraman, Tejaswi Upadhaya, Graciela Gonzales. Utilizing social media data for pharmacovigilance: a

review. *Journal of Biomedical Informatics,* 54:202-212, April 2015.

[WEB-RADR] WEB-RADR consortium. https://web-radr.eu/about-us/ , accessed May 7th, 2018

[Freifeld *et al*., 2014] Clark C. Freifeld, John S. Brownstein, Christopher M. Menone, Wenjie Bao, Ross Filice, Taha Kass-Hout, Nabarun Dasgupta. Digital drug safety surveillance: monitoring pharmaceutical products in Twitter. *Drug Safety*, 37(5):343–350, April 2014.

[Mikolov *et al*., 2013] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S. Corrado, Jeff Dean. Distributed representations of words and phrases and their compositionality. In *Proceedings of Advances in Neural Information Processing Systems*, 26, 2013.