Abstract

The aim of this paper is present and regroup 70 research works that have been done for detecting Adverse Drug Events from social media. For each work, we focus on its description, its aim, its approach, the used models, the used datasets, its novelty and its limitation.

1 Introduction

Detection of ADEs (drug side effects) is one of the main tasks in the pharmaceutical industry. ADEs can have profound effects on patients’ quality of life and is one of the leading causes of increased mortality internationally. The massive use of social media, and the abundance of discussions relating to healthcare in general (where drugs and ADEs are the most widely discussed categories), let them represent an excellent source for extracting ADEs. More recently, different works have been proposed for extracting ADEs from social media. The role of this paper is to briefly present, classify and analyse these works. To the best of our knowledge, no prior survey was proposed on ADEs detection in social media.

2 Adverse Drug Events in social media: related works

The works on ADEs detection in social media can be grouped into different categories: classification, extraction, normalisation, corpus creation and other analysis related to ADEs such as the correlation between Drugs-ADEs or sentiment analysis regarding ADEs.

3.1 ADEs classification

The classification corresponds to assigning the right class to tweets, posts, texts, etc. In the majority of cases, it is a binary classification where we only have two classes ADE (for texts including ADEs) and NoADE (for texts without ADEs). The classification represents the first step for detecting if a given text includes reference to ADEs or not. Hence different works focus on this tasks (Liu et al., 2019; Ribeiro et al., 2021; Dai and Wang, 2019; Booth et al., 2018; Rakhsha et al., 2021; Aji et al., 2021; Kayastha et al., 2021; Mane et al., 2018; Hsu et al., 2021; Pimpalkhute et al., 2021; Wang et al., 2018; Habibabadi et al., 2022).

3.2 ADEs detection

Two common approaches have been used for medical entity extraction in general (including ADEs extraction): lexicon-based and machine learning methods (Xie et al., 2018). The majority of the most recent studies rely on the machine learning approach (Wahbeh et al., 2021; Xie et al., 2018; Gattepaille et al., 2020; Wang et al., 2021; Lavertu et al., 2021; Shen et al., 2020, 2021; Zhang et al., 2020, 2021; Rakhsha et al., 2021; Bollegala et al., 2018b). However, we also observe that some approaches can not be classified into those categories.
where the authors are extracting ADEs from a corpus that was manually annotated without using any lexicon or machine learning techniques (Alex et al., 2020). Some other approaches exploit various lexical, semantic, and syntactic features, and integrated ensemble learning and semi-supervised learning in order to detect ADEs (Liu et al., 2018). Some authors start by training their own embedding model that they use after for the detection (Hoang et al., 2018) (where AC-SPASM, a Bayesian model for the authenticity and credibility aware detection of potential ADEs from social media is trained and used). Finally, in addition to detecting ADEs, some approaches also highlight the correlation between drugs and ADEs (De Rosa et al., 2021).

### 3.3 Normalisation

The normalisation consists in assigning (mapping) ADEs to their corresponding codes in medical ontologies such as Unified Medical Language System (UMLS), SNOMED CT, Medical Dictionary for Regulatory Activities (MedDRA), etc. This task is in most cases associated with the detection (extraction) where the ADEs are first extracted automatically and then mapped to an existing ontology. To the best of our knowledge, it has no works dedicated to normalisation only without involving the detection (Ji et al., 2021). To map the extracted ADEs to MedDRA, these authors first apply Neural Transition-based Model for named entity recognition (NER) and then link each extracted mention to its MedDRA code.

### 3.4 Resources creation

Some authors start dedicating their efforts to constructing such resources (Dietrich et al., 2020; Laksito et al., 2018; Alvaro et al., 2017; Karimi et al., 2015). Other studies focus in validating the constructed corpus either by classifying ADEs (Smith et al., 2018; Shen et al., 2019; Habibabadi et al.; Duval and Silva, 2019; Jiang et al., 2018; Li et al., 2020) or by extracting them (Li et al., 2020; Arnoux-Guenegou et al.). For this category of works, Twitter was also the predominant source for collecting data.

### 3.5 Classification, detection and normalisation

Some works focus on a pipeline including both tasks, such as (Yaseen and Langer, 2021), which did not only perform binary classification of the ADE text, but also extracted them. Many other studies (Fuentes-Carbajal et al., 2022; Wang et al., 2022; Guo et al., 2021; Zhang et al., 2019; Bolle-gala et al., 2018a; Tang et al., 2018; Saha et al., 2021; Kim et al., 2020) followed the same pattern, while others (Sakhovskiy et al., 2021; Zhou et al., 2021; El-karef and Hassan, 2021; Magge et al., 2021; Jagannatha et al., 2019; Dima et al., 2021; Ramesh et al., 2021; Barry and Uzuner, 2019) have added normalisation to the pipeline as a technique for transforming features to be on a similar scale like associate or map extracted ADEs to code.

### 3.6 ADEs analysis

The last category of work is dedicated to carrying on some analysis related to ADEs (Golder et al., 2019; Lentzen et al., 2022; Lyu et al., 2020; Golder et al., 2021; Clemens et al., 2022; Chalasani et al., 2018; Saha et al., 2021; Nawar et al., 2022; Zhou et al., 2020; Suragh et al., 2018). These analyses could be related to the sentiments, expectations, and anxiety of the users related to ADE (Golder et al., 2019; Lentzen et al., 2022; Clemens et al., 2022; Suragh et al., 2018). They can also be related to some linguistic features validated by clinical experts for detecting ADEs (Lyu et al., 2020) or to a comparison of the ADE related to a given drug with others or evaluating the Complementary and Alternative Medicine (CAM) (Golder et al., 2021; Saha et al., 2021; Nawar et al., 2022). Finally, some analyses are dedicated to evaluating the precision and the accuracy of the ADEs reported on Social media (Zhou et al., 2020).

### 4 Conclusion

In total, we collected 70 papers that we synthesise and summarise (for more detail, refer to the appendix part). These works were classified into 6 different categories: works on classification, detection, normalisation, classification and extraction and normalisation, resources construction and ADEs analysis. To sum up, many challenges are related to extracting data from social media including the proportion of noise, diversity in content, expressions, language and posting formats, non-textual content used as text, and use of symbols, emoticons and jargon (Indani et al., 2020). Finally one of the most important challenges behind working on collected data from social media is to obtain imbalanced corpora. Few studies only focus on these issues and the majority of the works that did,
are focusing on oversampling the data. Many other techniques for balancing a dataset have been proposed. Hence more experiments are required in this part.

References


Kelly S Clemens, Kate Faasse, Winston Tan, Ben Collaguri, Luana Colloca, Rebecca Webster, Lene Vase, Emily Jason, and Andrew Geers. 2022. Social pathways to side-effects: Personal contacts and social media predict covid-19 vaccine side-effect expectations and experience.


Sedigheh Khademi Habibabadi, Pari Delir Haghhighi, Frada Burstein, and Jim Buttery. Mining vaccine adverse events mentions from social media using twitter as a source.


Tianchu Lyu, Andrew Eidson, Jungmi Jun, Xiajie Zhou, Xiang Cui, and Chen Liang. 2020. Data veracity of patients and health consumers reported adverse drug reactions on twitter: Key linguistic features, twitter variables, and association rules. *medRxiv*.


Zeyun Zhou, Kyle Emerson Hultgren, et al. 2020. Complementing the us food and drug administration adverse event reporting system with adverse drug reaction re-