NLP+EHR: Automatic Detection of Adverse Drug Reactions

A group of Spanish researchers present the results of their work on applying natural language processing (NLP) techniques for   electronic health record analysis and specific data extraction.

You might also like: A vendor-neutral, open-source, natural language approach of EHR data is suggested by a team of U.K. researchers, to help providers plan for surges in patients, particularly in the COVID-19 context. Learn more

According to a new study published in Computer Methods and Programs in Biomedicine (Santiso et al. 2021), a group of researchers at   the University of the Basque Country (UPV/EHU) are collaborating with the regional health authorities to develop a system for automatic extraction of adverse drug reactions (ADR) from EHRs written in Spanish.

ADR may cause hospital re-admissions and even death in some cases. Applying machine learning and deep learning techniques for clinical text mining, the researchers aim to pinpoint relations between drug-disease pairs. With such a system these reactions can be identified, summarised and automatically reported, assisting with clinical decision-making.

The flow comprises recognising relevant clinical entities (e.g. drugs and their ingredients, or symptoms) and then evaluating drug-disease candidate pairs as either ADR or non-ADR.

The challenges include the heterogeneity of the EHR lexical data, and the sensitivity of those data which, therefore, are difficult to obtain for research purposes. Also, at the entity recognition stage errors are possible which affects the ADR detection. As such, the system and the sensitivity of the ADR detection stage were assessed against different levels of noise through exposure to different corpora of documents (e.g. from different hospitals and of different size).

The finding show that deep learning techniques were more efficient in detecting ADR, and "having a larger corpus helps the system learn the examples contained in it more effectively, thereby giving rise to better results". The authors therefore stress the need to expand the EHR data samples as much as possible to improve the ADR extraction. Overall, the proposed system was reported as being able to cope with cross-hospital predictions, with the f-measure in the ADR detection stage varying from 75.2 to 68.6 when errors happened at the entity recognition step. The researchers underscored the importance of extracting implicit, inter-sentence information contained in EHR, for drug-disease pairs detection.

Source: UPV/EHU

Image credit: metamorworks via iStock

Published on : Sun, 7 Mar 2021