
Evaluation of BERT-Based Models on Patient Data from French Social Media

Emma Le Priol^{*1,2}, Manissa Talmatkadi³, Stéphane Schuck³, Nathalie Texier³, and Anita Burgun^{4,5,6,7}

¹Université Paris Cité – Université Paris Cité – France

²Health data- and model- driven Knowledge Acquisition (HeKA) – Inserm U1138, Centre de Recherche des Cordeliers, L’Institut National de Recherche en Informatique et en Automatique (INRIA) – France

³Kap Code – Kap Code – 75002, France

⁴Université Paris-Cité – Université Paris Cité – France

⁵Health data- and model- driven Knowledge Acquisition (HeKA) – INSERM UMR1138, L’Institut National de Recherche en Informatique et en Automatique (INRIA) – France

⁶Clinical Bioinformatics Group, Institut Imagine – Inserm UMR 1163, Institut Imagine – France

⁷Department of Medical Informatics, Necker Hospital, AP-HP – AP-HP Hôpital Necker - Enfants Malades [Paris] – France

Résumé

Background

Our study aimed to extract new insights on Developmental and Epileptic Encephalopathies (DEE), a group of rare and severe diseases, from French social media messages by evaluating three BERT-based models on a Named Entity Recognition (NER) task. This task involved identifying phenotypes and treatments from social media posts. The pre-trained models evaluated were CamemBERT, CamemBERT-bio, and DrBERT.

Materials & Methods

They were trained on a dataset containing social media messages discussing DEE and more common diseases, in which treatment and phenotype entities were annotated. The precision, recall and F1-score were calculated in strict mode for each entity, and for micro, macro, and weighted averages. They were averaged on 10 runs.

Results

Our results revealed that CamemBERT and CamemBERT-bio performed similarly, slightly outperforming DrBERT. Both models showed a balanced performance in precision and recall, though their performance was lower on social media data compared to more "traditional" health datasets.

Discussion

Our study highlighted the promise of deep learning methods, particularly transformer-based models, in analyzing medical content from social media. However, the limitations of this

*Intervenant

study include a narrow focus on NER performance and a dataset-specific evaluation, which calls for further research to assess the models on larger and more diverse datasets. Another limitation is the annotation process, which was not the same for the DEE messages and the others, and with a low inter-annotator agreement. Other pre-trained models (SocBERT or BERTweet) could also be explored, as well as generative models. Being able to extract phenotype and treatment entities from social media messages could significantly enhance the understanding of the challenges faced by patients and caregivers, especially in the context of DEE. Future research will include a new annotation process, as well as exploration of other pre-trained and generative models. The long-term objective being to analyze posts from a French patients' organization Facebook group.

Mots-Clés: rare disease, developmental and epileptic encephalopathies, bert, transformers, deep learning, health, social media