A comparison of neural network models for adverse drug reactions identification

Ilseyar Alimova
alimovailseyar@gmail.com
Kazan Federal University
Kazan, Russia

Tutubalina Elena
tutubalinaev@gmail.com
Kazan Federal University
Kazan, Russia

ABSTRACT

This paper presents our experimental work on exploring the potential of neural network models developed for aspect-based sentiment analysis in the adverse drug reaction (ADR) classification. We conducted extensive experiments on various sources of text-based information, including social media, electronic health records, and abstracts of scientific articles from PubMed. The results show that Interactive Attention Neural Network outperformed other models on four corpora in terms of macro f-measure.

KEYWORDS

text mining, natural language processing, health social media analytics, machine learning, deep learning

1 INTRODUCTION

Detection of adverse drug reactions (ADRs) in the post-marketing period is becoming increasingly popular, as evidenced by the growth of ADR monitoring systems. Information about adverse drug reactions can be found in the texts of social media, health-related forums, and electronic health records. There are two steps of ADR detection: named entity recognition and entity classification. Thus, at the first step, all information related to a state of health are extracting using named entity recognition systems. Then all obtained entities are classified in order to distinguish ADRs. In this article, we focused on the task of binary classification for extracting ADRs. Inspired by recent successful methods in the field of sentiment analysis, we investigate state-of-the-art deep neural network models developed for aspect-based sentiment analysis for entity-level ADR classification task.

2 CORPORA

CADEC (CSIRO Adverse Drug Event Corpus) consists of annotated user reviews written about Diclofenac or Lipitor on askapatient.com. **MADE** corpus consists of de-identified electronic health record notes from 21 cancer patients. **TwiMed** corpus consists of sentences extracted from PubMed and tweets. This corpus contains annotations of diseases, symptoms, and drugs, and their relations. **Twitter** corpus include tweets about drugs. More information is provided in [1].

3 MODELS

We started with the basic LSTM-based model and then extend the model by attention mechanism and explicit memory. **LSTM**: The basic model which models the semantic representation of a sentence without considering an entity. **TD-LSTM**: the target-dependent model utilizes two LSTMs to model the left context with an entity and right context with an entity. **IAN**: the interactive attention network consists of LSTMs for handling an entity and its context and cross attention block where the final concatenation of attentions' outputs is fed to softmax for prediction. **MemNet**: The deep memory network applies attention multiple times on the word embeddings (multi-hop property), and the last attentionâĂŹs output is fed to softmax for prediction. **RAM**: recurrent attention on memory is the more complex model than MemNet. RAM utilizes LSTMs for handling an entity and its context and multi-hop attention on memory. A more detailed description can be found in https://github.com/songyouwei/ABSA-PyTorch.

4 RESULTS

The models were evaluated for 5-fold cross-validation using standard classification quality metrics: accuracy (P), completeness (R), and macro F-measure. The results are presented in Table 1. The results show that IAN outperformed other models on all corpora except Twitter. IAN obtained the most significant increase in results compared to other models on Twimed-Twitter and Twimed-Pubmed corpora with 81.9% and 87.4% of the macro F-measures, respectively. On the Twitter corpus, the RAM model achieved the best results with a macro F-measure of 83.4%.

Table 1: Macro F-measure classification results of the compared methods for each datasets.

Corpora	Twitter	Cadec	MADE	Twimed Twitter	Twined Pubmed
LSTM	.613	.784	.771	.700	.839
TD-LSTM	.758	.772	.750	.730	.709
IAN	.794	**.815**	**.786**	**.819**	**.874**
RAM	**.834**	.734	.761	.780	.789
MemNet	.763	.758	.760	.795	.811

5 CONCLUSION

We explored the potential of state-of-the-art neural network models on entity-level ADR classification task. Neural attention models are considered to be state-of-the-art for aspect-based sentiment analysis. Based on our experimental results, we can note that the application of attention mechanism improves the performance of a model. IAN model performs the best results for entity-level ADR classification task in most of our experiments.

REFERENCES

[1] Ilseyar Alimova and Valery Solovyev. 2018. Interactive Attention Network for Adverse Drug Reaction Classification. In *Conference on Artificial Intelligence and Natural Language*. Springer, 185–196.