Drug-drug Interaction Extraction via Transfer Learning

Yuqing Mao, PhD[1], Kin Wah Fung, MD[1], Dina Demner-Fushman, MD, PhD[1]
[1]Lister Hill National Center for Biomedical Communications, National Library of Medicine, National Institutes of Health, Bethesda, MD, USA

Introduction

Drug-drug interaction (DDI) is an unexpected modification in the effect of a drug when taken in combination with another drug. It has the potential to cause significant harm to the patient. The U.S. Food and Drug Administration (FDA) and the National Library of Medicine (NLM) have worked collaboratively on transforming the content of Structured Product Labeling (SPL) documents for prescription drugs into machine-readable data. To inform future FDA efforts at automating important safety processes, the Text Analysis Conference (TAC) 2018 DDI track[1] has provided unique datasets for testing the performance on extracting DDI information from SPL documents of various natural language processing (NLP) approaches.

Recently, deep learning approaches have made significant progress in many NLP tasks. The most important advantage of deep learning approaches is that they do not need manually defined lexical features. More breakthroughs in NLP area were achieved via transfer learning in last year. Transfer learning for NLP tasks includes two steps: the first step is to pre-train a model on large amounts of unlabeled texts; the second step is to transfer the learned general language knowledge to a specific NLP task through fine-tuning. Pre-trained model can improve performances on NLP tasks also because the context can be encoded in its textural representations. In this work, we focus on building a transfer learning framework with BERT[2], one of the best publicly available, large-scale, pre-trained models, which we used for extracting DDI information from SPL documents.

Methods and Results

The TAC 2018 DDI track includes four shared tasks, and our model is designed to tackle the first two tasks. Task 1 is to extract mentions of interacting drugs/substances, interaction triggers and specific interactions at sentence level. Task 2 is to identify interactions, including the interacting drugs, the specific interaction types: pharmacodynamics (PD) pharmacokinetic (PK), or unspecified, and the outcomes of PD and PK interactions. Since the second task requires the output of the first task, we model them jointly by encoding the associated interaction type for each precipitant drug. Our framework contains three major components: an encoder for contextual embeddings, a BiLSTM-CRF to recognize all entities including precipitants, triggers and effects, and a CNN with two separate dense output layers (one binary classifier for PD interactions and one multiclass classifier for PK interactions) to predict outcome. We utilize the pre-trained BERT model as the context encoder, and compare its performance with one of the state-of-the-art models in TAC 2018 DDI track, which has the same architecture but uses a Bi-LSTM encoder that composes a context representation and character-CNN composed representations. The results of evaluation on TAC 2018 DDI track datasets are shown in Table 1. It is clear that the integration of the pre-trained BERT model can improve the performance of DDI extraction tasks.

Table 1. F1-scores of our BERT-based model and the SOTA model based on Bi-LSTM + CNN.

Model	Test Set 1, Task 1	Test Set 1, Task 2	Test Set 2, Task 1	Test Set 2, Task 2
Bi-LSTM + CNN	33.00	21.59	38.28	23.55
$BERT_{BASE}$	35.58	24.32	39.94	25.76
$BERT_{LARGE}$	**41.03**	**28.10**	**44.76**	**30.93**

Conclusion

Transfer learning shows improvements in DDI extraction tasks. The performance can potentially be further improved if the pre-trained model could learn biomedical domain specific knowledge through training with a large set of unlabeled texts such as SPL documents and PubMed Central full-text articles.

References

1. Demner-Fushman D., Fung K.W., Do P., Boyce, R.D., Goodwin T. Overview of the TAC 2018 Drug-Drug Interaction Extraction from Drug Labels Track. In Proceedings of Text Analysis Conference (TAC). 2018.
2. Devlin J, Chang MW, Lee K, Toutanova K. BERT: Pre-training of deep bidirectional transformers for language understanding. Oct. 2018, arXiv preprint arXiv:1810.04805.