

# Filtering Clinical BERT (FC-BERT): 증상과 약물 이상 반응 구분을 위한 약물 이상 반응 탐지 모델

이채연<sup>1</sup>, 김현희<sup>1</sup><sup>1</sup> 동덕여자대학교 정보통계학과[chland23@naver.com](mailto:chland23@naver.com), [heekim@dongduk.ac.kr](mailto:heekim@dongduk.ac.kr)

## Filtering Clinical BERT (FC-BERT): An ADR Detection Model for distinguishing symptoms from adverse drug reactions

Chae-Yeon Lee, Hyon Hee Kim

Dept. of Statistics and Information Science, Dongduk Women's University

### 요 약

최근 소셜미디어 리뷰 데이터를 활용한 약물 이상 반응 탐지 연구가 활발히 진행되고 있지만, 약물을 복용하기 전 증상과 약물 이상 반응을 구분하지 못한다는 한계가 있다. 본 논문에서는 약물 이상 반응 탐지에서 약물 복용 전의 증상을 구분할 수 있는 Filtering Clinical BERT(FC-BERT) 모델을 제안하였다. FC-BERT는 약물 복용 전 증상과 다른 약물에 대한 부작용 표현을 제거하기 위해 약물명이 나오기 전 모든 문장을 제거하는 필터링과 약물-부작용 쌍을 추출하는 모델을 사용했다. 성능 평가 실험을 위해 문장에 대한 ADE(Adverse Drug Event) 여부가 들어있는 ADE Corpus V2 데이터를 활용하였고 SPARK NLP 라이브러리에서 제공하는 ADE Pipeline 모델과 비교하여 성능 평가를 실시하였다. 실험 결과 필터링을 활용한 FC-BERT 모델이 기존 모델보다 정확도, 평균 정밀도, 평균 재현율, 평균 F1-score가 모두 높은 결과를 보여주었다. 본 논문에서 제시한 모델은 기존 연구의 한계점을 보완하여 보다 정확한 약물 부작용 시그널을 탐지하는데 기여할 수 있을 것이다.

### 1. 서론

의약품의 약물유해사례에 대한 관리를 위해 우리나라에서는 1988년부터 자발적 부작용 보고제도를 시행하고 있으나 환자가 직접 부작용을 보고해야 하기 때문에 자발적 약물 부작용 보고 건수는 낮은 편이다 [1]. 이런 문제점 때문에 최근 소셜 미디어 사용자들이 플랫폼에 올리는 리뷰 데이터를 활용한 약물 부작용 시그널을 탐지하는 연구들이 주목받고 있다.

[2]의 연구에서는 HAN, FastText, CNN 등의 특징 추출 네트워크를 통해 Twitter 데이터를 교차 검증 및 학습하여 데이터를 ADR과 nonADR로 분류하였다. [3]의 연구에서는 단어를 벡터로 변환하는 Word2Vec 모델 [4]을 이용하여 문헌 데이터에서 언급되는 약물과 부작용을 벡터로 매핑하고, 이 벡터를 이용하여 약물-부작용 쌍의 관계 점수를 구하여 부작용을 예측했다. 최근에는 BERT 임베딩을 적용한 약물 부작용 추출 연구도 많은 관심을 받고 있다. [5]의 연구에서는 BiLSTM-CNN-CRF 구조 [6]의 개체명 인식 모델을 이용하여 약물 리뷰 데이터에서 부작용 표현을 찾아, 해당 표현을 LSTM을 이용해 MedDRA 용어로 분류했다.

또한 [7]의 연구에서는 BERT를 활용한 약물 리뷰 레이블링, 감성분석을 이용하여 긍정 리뷰와 부정 리뷰로 분류하였고, MedDRA 기반 부작용 사전을 통해 부정 리뷰에서 약물 부작용의 시그널을 추출했다.

이러한 연구들은 문장에서 약물-부작용 쌍을 찾아내거나 약물 리뷰에서 부작용을 추출하는 방식을 사용한다. 이는 모든 리뷰안에 공통적으로 들어있는 약물 복용 전 증상이나 다른 약물에 대한 부작용 표현을 제거하기에는 한계가 있다. 즉, 약물을 복용하기 전의 증상도 일종의 약물 이상 반응이라고 잘못 판별하는 오류를 범할 수 있다. 예를 들어 ‘두통이 있을 때마다 아세트아미노펜 계열의 진통제를 상용하였더니 간손상이 생겼다.’라는 문장에서 두통은 증상이며, 간손상은 약물 이상 반응이다. 현재의 약물 이상 반응 탐지 모델은 증상과 이상 반응의 구분이 없이 모두 약물 이상 반응으로 탐지하고 있다. 뿐만 아니라 소셜미디어 약물 리뷰 데이터에서는 해당 약물을 언급할 때 다른 약물을 사용했을 때와 비교하여 언급하는 경우가 많으므로 다른 약물의 이상반응도 탐지되는 문제도 발생하고 있다.

본 논문에서는 보다 효과적인 약물 부작용 탐지를 위해 불필요한 표현을 제거하는 필터링을 활용하여 약물 부작용 표현 탐지의 정확도를 높인 FC-BERT 모델을 제시하였다. FC-BERT 모델은 SPARK-NLP[8]에서 제공하는 사전 훈련된 ADE Pipeline, ADE Classifier, ADE Relation Extraction, SentenceDetectorDL(SDDL) 모델을 기반으로 하고, 추가적으로 필터링을 활용하여 약물 이상 반응 으로부터 증상을 분리할 수 있도록 하였다. 또한 제안하는 모델을 MEDLINE 데이터베이스 ADE 말뭉치에 적용하여 ADE 유무를 분류하는 관별하고 성능 평가를 실시하였다.

전처리된 데이터에 SPARK NLP 에서 제공하는 ADE Pipeline 모델을 적용하는 기본 모델과 SPARK NLP 에서 제공하는 사전 훈련된 4 개의 모델과 제안하는 2 단계의 필터링을 적용한 FC-BERT 모델을 비교하여 실험하였다. 비교 실험을 통하여 본 연구에서 사용된 FC-BERT 모델 기반의 필터링을 사용한 방식이 ADE Pipeline 모델을 사용한 방식보다 정확도, 평균 정밀도, 평균 재현율, 평균 F1-score 모두 뛰어난 것을 확인하였다.

본 논문의 구성은 다음과 같다. 제 2 장에서는 데이터 수집 및 전처리를 소개한다. 제 3 장에서는 기존 연구의 한계를 보완하는 필터링 방법을 제시하고, 이를 토대로 ADE Corpus 데이터의 ADE 유무를 예측하는 FC-BERT 모델을 제시한다. 마지막으로 제 4 장에서는 두가지 모델을 비교하는 성능평가 방법을 고안하여 성능을 수치화하고, 본 논문의 결론과 향후연구를 제시하는 과정으로 이루어진다.

## 2. 데이터 수집 및 전처리

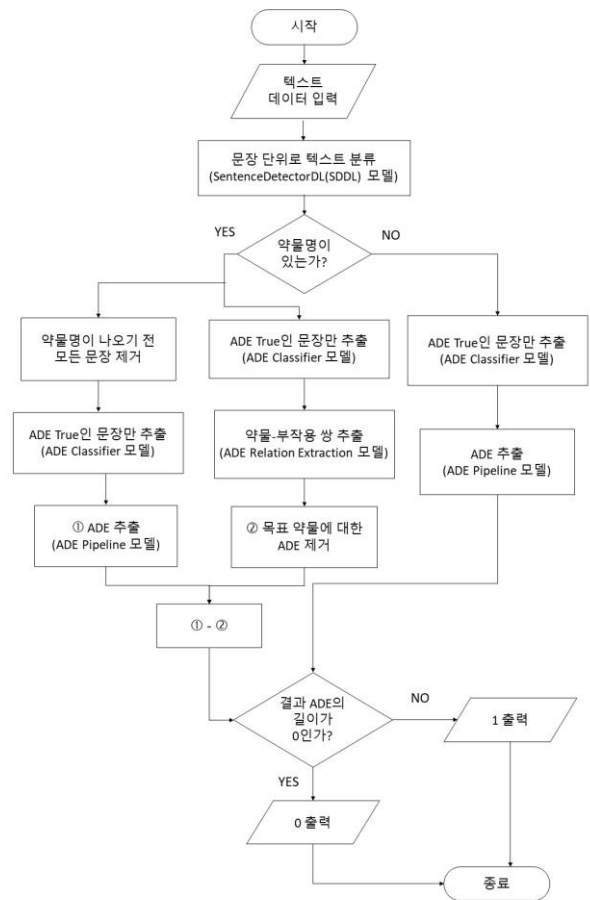
제안하는 모델을 생성하고 성능을 검증하기 위한 데이터로 ADE Corpus v2 데이터를 활용하였다[9]. ADE Corpus v2 데이터는 MEDLINE 데이터베이스에 접근가능한 자유 검색 엔진인 PubMed 에서 수집한 23,516 개의 ADE 말뭉치로 구성되어있으며 github 를 통해 다운받을 수 있다[10]. ADE 가 포함된 말뭉치는 6,821 개로 약물명이 포함되어 있지만, ADE 가 포함되지 않은 말뭉치는 16,695 개로 약물명이 포함되어 있지 않다. 따라서 ADE 가 포함된 말뭉치는 문장과 약물명을 사용하였으며, ADE 가 포함되지 않은 말뭉치는 문장만 사용하였다. 최종적으로 모델에 의해 추출된 ADE 가 있으면 1, 없으면 0 으로 분류를 진행하였다.

ADE 추출이 원활하게 이루어지도록 전처리는 정규 표현식을 활용하여 불필요한 문자를 제거했다. 불필요한 문자에는 숫자, 영어, 마침표, 물음표, 느낌표, 콤마, 따옴표, 슬래시, 괄호 등이 있다. 문장의 끝을 나타내는 마침표, 물음표, 느낌표 등의 문자와 강조를 나타내는 괄호나 슬래시 등의 문자는 문장 구분이 원활하게 이루어지도록 하기 위해서 제거하지 않았다.

## 3. FC-BERT 모델

<그림 1>의 FC-BERT 모델은 SPARK NLP 에서 제공하는 사전 훈련된 4 개의 모델과 제안하는 2 단계의 필터링으로 이루어져 있다. Spark NLP 의 NER 모델을 위한 심층 신경망 아키텍처는 BiLSTM-CNNChar 프레임워크이며, 하이브리드 양방향 LSTM 및 CNN 아키텍처를 사용하여 단어 및 문자 수준 기능을 자동으로 감지한다[11].

먼저 공통적으로 SentenceDetectorDL(SDDL) 모델을 이용해 데이터를 문장 단위로 분류한다. SDDL 모델은 문장 경계 검출을 위한 범용 뉴럴 네트워크 모델을 기반으로 문장 경계 감지 작업을 DL CNN 아키텍처를 사용한 분류 문제로 처리한다.



<그림 1> FC-BERT 모델 순서도

### 3.1 증상을 제거한 ADE 추출

문장 단위로 분류된 전체 문장에서 해당 약물명이 언급되기 전 모든 문장을 제거하는 필터링을 거친다. 이는 ‘약물을 복용하기 전 증상을 말하는 문장은 약물명을 언급하기 전에 나온다.’ 를 전제로 하기 때문이다. 약물명이 언급이 된 이후에 약물에 대한 효과나 부작용을 언급할 것이라고 가정한 것이다. 약물명이 언급되지 않았거나 ADE 가 포함되지 않은 말뭉치는 약물명을 찾을 수 없으므로 어떤 문장도 삭제되지 않는다. 이후 ADE Classifier 모델을 적용하여 ADE 가 포함된 문장들을 추출한다. ADE Classifier 모델은 회화(짧은) 문장으로 훈련된 BioBERT 기반의 ADE 분류기이다. 약물명이 언급되기 전 모든 문장들이 삭

제되고 ADE 에 대한 어떠한 정보도 포함하지 않은 문장들을 제거한 말뭉치에 ADE Pipeline 모델을 적용하여 전체 문장에 대한 부작용 표현들을 추출한다. ADE Pipeline 모델은 ner\_ade\_healthcare 및 classifierdl\_ade\_biobert 를 사용하여 약물 부작용 표현을 추출하는 모델이다.

### 3.2 다른 약물에 대한 ADE 추출

맨 처음 문장 단위로 분류된 전체 문장 각각에 ADE Classifier 모델을 적용하여 ADE 가 포함된 문장들을 추출한다. ADE 에 대한 어떠한 정보도 포함하지 않은 문장들을 제거한 말뭉치에 ADE Relation Extraction 모델을 적용하여 약물명-부작용 쌍들을 추출한다. ADE Relation Extraction 모델은 먼저 ner\_ade\_clinical NER 모델을 사용하여 DRUG 및 ADE 엔티티를 검출한 다음, re\_ade\_clinical 모델을 사용하여 이들 사이의 관계를 찾는다. 추출된 약물명-부작용 쌍 중에서 목표 약물-부작용 쌍을 제거하면 목표 약물이 아닌 다른 약물에 대한 부작용 표현들이 추출된다. 이는 ADE Relation Extraction 모델이 목표 약물에 대한 모든 부작용 표현을 검출하지는 못하지만 목표 약물이 아닌 다른 약물에 대한 부작용 표현들은 비교적 정확하게 검출하기 때문이다. 약물명이 언급되지 않는 말뭉치는 이 과정을 거치지 않는다.

### 3.3 최종 결과 ADE 추출 후 분류

첫번째 결과는 말뭉치에서 약물을 복용하기 전 증상들을 제거한 모든 ADE 이고, 두번째 결과는 목표약물이 아닌 다른 약물의 ADE 이다. 최종적으로 첫번째 결과 중에서 두번째 결과에 해당하는 부작용 표현을 제거하는 필터링을 거치면 약물 복용 전 증상과 다른 약물에 대한 ADE 를 제거한 부작용 표현들이 추출된다. 모델을 통해 추출된 부작용 표현이 있으면 1 로 분류하고, 추출된 부작용 표현이 없으면 0 으로 분류한다. Algorithm 1 은 FC-BERT 모델에 대한 유사 코드를 나타낸다.

```

Algorithm 1 Pseudo code of FC-BERT
Input: X: drug review text data
Output: S: a list of strings that split corpus into sentences
RESULT1: ADE extracted from a corpus removed from symptoms
RESULT2: ADE for other drugs extracted from a corpus
RESULT: The final result obtained by subtracting RESULT2 from RESULT1
// STEP 1 Split text into sentences
Read X
Set S to []
Perform Sentence classification using SDDLModel
Add classification result to S
Return S

// STEP 2 ADE extraction with symptomatic ADE removed
Read S
if S is a sentence with a drug name then
  for k in S:

```

```

    if drug name in k then
      Delete the previous sentence of k
    endif
  endif
endif

Set s_true to []
for i in S:
  Predict ADE status using ADE Classifier model
  if the predicted value is 'True' then
    Add S[i] to s_true
  endif
endif
Perform convert a list of sentences to a single string
Perform ADE df extraction using ADE pipeline model
if 'entities' in df is 'ADE' then
  RESULT1 <- df
endif
Perform Extracting 'Side Effect Representation' and 'classification Entity' from RESULT1
Return RESULT1

// STEP 3 ADE extraction for other drugs
Read S
if S is a sentence with a drug name then
  Set s_true to []
  for i in S:
    Predict ADE status using ADE Classifier model
    if the predicted value is 'True' then
      Add S[i] to s_true
    endif
  endif
  Perform convert a list of sentences to a single string
  Perform ADE DF extraction using the ADE RE model
  Perform Changing the order of ADE DF columns
  RESULT2 <- df
  Perform Extracting 'Side Effect Representation' and 'classification Entity' from RESULT2
else
  RESULT2 <- empty df
endif
Return RESULT2

// STEP 4 Final Result ADE Extraction and Classification
Read RESULT1, RESULT2
RESULT <- RESULT1 - RESULT2
if len(RESULT) is 0 then
  Return 0
else
  Return 1
endif

```

## 4. 실험결과

본 논문에서 제안한 약물 부작용 탐지 모델을 검증하기 위해 구글 Colab 과 SPARK NLP 라이브러리인 sparknlp 3.4.1 버전을 사용하여 실험을 진행하였다. 구글 Colab 은 Intel(R) Xeon(R) CPU @ 2.20GHz 사양이며, 파이썬(python)은 3.7.13 버전을 사용하였다.

Dataset	ADE Corpus V2	
ADE Pipeline	Accuracy	0.611
	Average Precision	0.712
	Average Recall	0.725
	Average F1-Score	0.718
FC-BERT	Accuracy	0.804
	Average Precision	0.767
	Average Recall	0.738
	Average F1-Score	0.752

<표 1> 모델에 따른 분류 실험결과

<표 1>을 통해 기존 모델보다 정확도가 약 19.3% 상승하는 것을 확인할 수 있다. 성능 지표는 0 에 대한 Precision, Recall, F1-score 와 1 에 대한 Precision, Recall, F1-score 의 평균인 Average Precision, Average Recall, Average F1-score 를 사용하였다. 정확도뿐만 아니라 평균 정밀도, 평균 재현율, 평균 F1-Score 모두 개선되는 것을 확인할 수 있었다.

Review	I experienced whole body pain due to dancing practices for several hours of the week. Tried taking ibuprofen, it gave me eye allergy. My eye bulge and I also experienced some chest pain.
ADE Pipeline	whole body pain(증상), eye allergy (ADE), eye bulge(ADE), chest pain(ADE)
FC-BERT	eye allergy (ADE), eye bulge(ADE), chest pain(ADE)

<표 2> 2 가지 모델의 약물 리뷰 문장 적용 결과  
<표 2>는 약물 리뷰 사이트에서 수집한 ‘ibuprofen’ 의 대한 약물 리뷰를 2 가지 모델에 적용한 결과이다. ADE Pipeline 은 약물을 복용하기 전 증상인 ‘whole body pain’ 이 같이 추출되는 반면 FC-BERT 는 증상이 제거되어 ADE 만 추출되는 것을 확인할 수 있다.

## 5. 결론

본 논문에서는 약물 복용 전 증상과 다른 약물에 대한 부작용 표현을 제거하지 못하는 기존 ADE 탐지의 한계점을 필터링을 활용하여 보완한 모델을 제안하였다. FC-BERT 모델은 BiLSTM-CNNChar 구조의 Clinical BERT 를 기반으로 필터링을 더하여 약물 부작용 표현 탐지를 더 정확하게 추출하게 하였다. 부작용 탐지 결과에 따라 ADE 와 nonADE 로 분류를 진행하였다. 비교 실험을 통해 ADE Pipeline 모델만을 사용한 기존 모델보다 필터링을 더한 FC-BERT 모델이 정확도, 평균 정밀도, 평균 재현율, 평균 F1-Score 가 더 높은 것으로 나타났다.

본 연구는 사전학습된 딥러닝 모델과 필터링을 기반으로 약물 리뷰 데이터로부터 효과적으로 부작용을 탐지한다는 점에서 기존 부작용 탐지 연구의 문제점을 해결할 수 있을 것으로 보인다. 본 연구에서는 사전학습된 모델만을 활용하여 분류하였으며, 향후 전

이학습을 통해 약물 부작용 탐지의 정확도를 높일 계획이다. 또한 제안하는 모델을 트위터, 블로그 등 다양한 약물 리뷰 데이터에 적용한다면, 보다 새로운 약물 부작용 시그널을 탐지하는데 기여할 수 있을 것이다.

## <참고 문헌>

- [1] So Hyeon Ahn, Sooyoun Chung, Sun-Young Jung, Ju-Young Shin, Byung-Joo Park, “Awareness of Adverse Drug Reaction Reporting System in General Population”, Health Policy and Management, Vol. 24, No. 2, pp. 164-165, 2014
- [2] Rezaei Z, Ebrahimpour-Komleh H, Eslami B, Chavoshinejad R, Totonchi M, “Adverse Drug Reaction Detection in Social Media by Deep Learning Methods.”, Cell journal, pp. 319-324, 2020
- [3] Seungsoo Lim, Hayon Lee, and Youngmi Yoon, “Prediction of New Drug-Side Effect Relation using Word2Vec Model-based Word Similarity”, Journal of KIIT, Vol. 18, No. 11, pp. 25-33, Nov. 30, 2020
- [4] Yeong-Hwi Ahn, Jin-Young Jung, Koo-Rack Park, “A Study on the Product Planning Model based on Word2Vec using On-offline Comment Analysis”, Proceedings of the Korean Society of Computer Information Conference, pp 79-80, 2021
- [5] Hyeon-jeong Jeong, Hyon Hee Kim, “Detecting and classification ADRs using Named Entity Recognition on social media”, Proc. of KIPS, 28(1), pp. 443-446, 2021
- [6] Yun Jeong Jang, Tae Hong Min, Jae Sung Lee, “Named-Entity Recognition Using Stacked Bi-LSTM-CRF Ensemble Model.”, Proc. of KCC, pp 2049-2051, 2018
- [7] Eun Yeong Heo, Hyun Jeong Jeong, Hyon Hee Kim, “Detection of Adverse Drug Reactions Using Drug Reviews with BERT+ Algorithm”, KIPS Trans. Sofw. Data Eng. Vol. 10, No. 11, pp. 465-472, 2021.
- [8] Veysel Kocaman, David Talby, “Spark NLP: Natural Language Understanding at Scale”, Software Impacts, VoL 8, 2021, <https://doi.org/10.1016/j.simpa.2021.100058>
- [9] Harsha Gurulingappa, Abdul Mateen Rajput, Angus Roberts, Juliane Fluck, Martin Hofmann-Apitius, Luca Toldo, “Development of a benchmark corpus to support the automatic extraction of drug-related adverse effects from medical case reports”, Journal of Biomedical Informatics, Vol. 45, pp. 885-892, 2012, <https://doi.org/10.1016/j.jbi.2012.04.008>.
- [10] <https://github.com/trunght/AdverseDrugReaction/tree/master/ADE-Corpus-V2>
- [11] Veysel Kocaman, David Talby, “Biomedical Named Entity Recognition at Scale”, Kocaman, Veysel and David Talby. “Biomedical Named Entity Recognition at Scale.”, ICPR Workshops (2020), <https://arxiv.org/abs/2011.06315>