

# 생성적 사전학습 언어모델 기반의 판결문 문장 생성에 관한 파일럿 연구

소광섭<sup>1</sup>, 김호중<sup>1</sup>, 박노섭<sup>2</sup>, 원동옥<sup>1\*</sup>

<sup>1</sup>한림대학교 인공지능융합학과

<sup>2</sup>한림대학교 글로벌학부

so@hallym.ac.kr, hojungkim@hallym.ac.kr, rspark@hallym.ac.kr, dongok.won@hallym.ac.kr

## A Pilot Study on the Generation of Legal Document Sentence based on Generative Pre-trained Transformer

Kwangsub So<sup>1</sup>, Ho-Jung Kim<sup>1</sup>, Ro-Seop Park<sup>2</sup>, Dong-Ok Won<sup>1\*</sup>

<sup>1</sup>Department of Artificial Intelligence Convergence, Hallym University

<sup>2</sup>Division of Global Studies, Hallym University

### 요 약

인공지능 기술이 발전함에 따라 경찰의 범죄수사 분야에서도 인공지능 기술을 적용하고자 하는 연구가 활발하다. 범죄수사의 결과물인 수사결과 보고서 작성에 있어 판결문은 중요한 데이터가 될 수 있다. 그러나 판결문은 공개된 데이터의 이미지화로 인해 정형화된 데이터의 확보가 까다롭고, 소수의 법조계 전문가가 아닌 일반인이 생성해내기 어려워 데이터 확보가 쉽지 않은 현실이다. 이에 본 연구에서는 생성적 사전학습 언어모델을 이용한 판결문 문장 데이터 생성을 제안하였다. 카카오의 KoGPT를 활용하여 실제 판결문장 일부를 제시한 결과 판결문과 유사한 형태의 문장을 생성한 것을 확인하였다. 향후 판결문 데이터를 활용하기 위한 인공지능 기술 기반 범죄수사 연구에 있어, 생성된 판결문 데이터를 활용할 수 있을 것으로 기대된다.

### 1. 서론

최근 형사소송법이 개정됨에 따라 경찰의 수사환경이 변화되고 있으며, 경찰의 수사결론에 대한 검증이 중요해지고 있다. 이와 더불어 인공지능 기술이 발전함에 따라 경찰의 범죄수사 내 수사결과보고서 작성 지원 분야에서도 인공지능 기술을 적용하고자 하는 움직임을 보이고 있다.[1]

수사결과보고서는 형식과 내용면에서 판결문과 유사하다. 판결문은 주장과 반대주장이 명확하고, 이에 대한 논증들이 분명히 나타나는 구조를 가진다. 이러한 판결문의 명확한 특성은 수사결과보고서 작성 방향의 지향점이 될 수 있다. 그렇기에 판결문은 인공지능 기술을 이용한 수사결과보고서 작성 및 지원 연구에 있어 중요한 데이터가 될 수 있다. 판결문을 범죄수사 분석 분야에 활용하여 인공지능 모델을 학습시키기 위해서는 대용량의 판결문 데이터 확보가 필수적이다. 문제는 학습에 필요한 판결문 데이터의 확보가 어렵다는 점이다. 법원을 통해 판결문 원문 파일이 공개되고 있으나, 최근 인공지능 모

델을 학습하기 위한 측면에서는 여전히 어려움이 있다. 또한 판결문은 문장의 호흡이 길고, 끊어짐이 적으며, 법률 용어, 한자어 등이 혼용되어 있다는 특성이 있다.[2] 그리고 판결문은 기존 언어들과 다르게 특유의 문체가 있다. 이러한 판결문의 특성들 때문에 인공지능 기술에 활용하기 위한 판결문 데이터 확보가 어렵다. 좋은 데이터를 확보하기 위해서는 비슷한 품질의 판결문을 생성해야 하는데, 판결문 데이터를 만들 수 있는 사람은 소수 법조계 전문가이며, 이들이 직접 생성한 데이터의 양은 제한적이다. 일반인이 인공지능기술 활용을 위한 고품질의 판결문 데이터를 만들기는 쉽지 않다. 이와 더불어 공개된 판결문 데이터의 절대적인 양이 부족하기에 데이터 확보에는 한계가 있다. 판결문 데이터를 활용한 원활한 인공지능 데이터 학습을 위해서는 판결문 학습데이터가 추가로 필요하다고 할 수 있다.

이에 본 논문에서는 생성적 사전학습 언어모델을 활용하여 판결문 데이터 생성에 관한 연구를 제안한다. 제안하는 시스템에서는 생성적 사전학습 언어모델을 활용하여 판결문 데이터와 유사한 특성을 갖는

문장을 생성해낸다.

본 논문의 구성은 다음과 같다. 2장에서는 배경 및 관련연구에 대해 살펴본다. 3장에서는 본 논문에서 제안하는 생성적 사전학습 언어모델을 활용한 판결문 데이터 생성 예시를 제안하고, 마지막에서는 결론 및 향후과제를 제시한다.

## 2. 배경 및 관련연구

### 2.1 데이터 증강

데이터 분석에 있어서, 확보된 데이터의 좋은 품질과 더불어 데이터의 절대적인 양의 확보가 중요하다. 기존의 딥러닝 연구에서는 데이터 증강기법을 이용하여 부족한 데이터를 확보할 수 있었다. 이미지 분석에 관한 연구에서 데이터의 증강은 다양하게 발전되어 왔다. 반면, 텍스트 데이터 분석에 관한 연구에서의 데이터의 증강은 이미지 분석 분야와 비교해 난이도가 높으며 한계가 있다. 이미지 데이터는 방향, 색상 등의 변화만으로도 데이터를 쉽게 변형, 생성할 수 있는 반면에, 텍스트는 문장의 의도를 보존하면서 표현을 변화시켜야 하기 때문에 데이터 증강이 어려운 점이 있다. 기존에 소개된 텍스트 데이터 증강기법으로는 특정 단어의 유의어 교체, 임의의 단어 삽입, 문장 내 임의의 두 단어 위치 변경, 임의의 단어 삭제 등의 간단한 텍스트 편집 기법을 활용한 EDA(Easy Data Augmentation) 방법[3], 번역기를 사용하여 Monolingual 데이터로부터 인공 데이터를 생성하는 방법을 연구한 Back Translation 방법[4], 사전 학습된 트랜스포머 모델 기반의 데이터 생성 방법[5] 등이 있다. 텍스트 데이터 증강기술에서 영어가 아닌 한글이라는 점은 또 다른 어려움을 갖는다. 한글은 영어와 비교하면 어순이 중요하지 않고, 교착어이며, 띄어쓰기가 제대로 지켜지지 않는 등의 특성을 갖는다.

### 2.2 생성적 사전 학습 언어모델

#### (Generative Pre-trained Transformer, GPT)

본 논문에서는 사용한 GPT는 BERT와 함께 트랜스포머 기반으로 발전된 언어 생성모델이다. BERT 모델은 트랜스포머의 인코더 부분을 활용하고, GPT는 트랜스포머의 디코더 부분을 활용한다. 인코더는 입력 데이터의 차원을 축소하며 DNN(Deep Neural Network)의 내부적 표현으로 변환하는 작업에 사용된다. 그래서 데이터 구조 인식에 자

주 사용되며, 디코더는 내부적 표현으로 변환된 데이터를 원래 상태로 복원하는 작업에 유리하다. 그리고 복원 과정 중 일부 잡음을 섞어 새로운 데이터의 생성이 가능하다. 그렇기에 BERT 모델은 문장 속의 문맥을 잘 이해하고, 중요 문장을 색출해내는 역할에 적합하며, GPT 모델은 텍스트 생성에 적합하다. 2020년 발표된 OpenAI의 GPT-3는 총 1,750억 개의 매개변수로 학습하였으며, 우리나라에서는 SKTbrain KoGPT-2, 네이버 하이퍼클로바, 카카오 KoGPT 등을 발표해왔다. 그러나 OpenAI의 GPT3 및 네이버의 하이퍼 클로바는 학습 모델이 공개되어 있지 않기 때문에, 본 연구에서 활용한 언어생성모델은 OpenAI의 GPT-3 모델을 한국어에 특화시킨 카카오 KoGPT이다.

### 3. 제안

본 연구에서는 생성적 사전학습 언어모델 중 카카오의 KoGPT를 이용한 판결문 데이터 생성을 제안한다. 기존 연구에서는 사전학습 언어모델을 활용하여 판결문서 요약[6]에 대해 진행하였다면, 본 연구에서는 데이터 증강을 위해 언어생성모델을 이용해 판결문 데이터를 생성한 점이 특징이다.

<표 1> KoGPT 기반 문장 생성 예시 1-1

제시	E마트로부터 '2008년 9월 경찰관 주석 구입비'로 1,367,000원 상당의 물품을 구입한 것으로 공문서인 지급요청서 1매와 지급결의서 1매를 각각 허위로 작성한 것을 비롯하여,
생성	E마트로부터 '2008년 9월 경찰관 주석 구입비'로 1,367,000원 상당의 물품을 구입한 것으로 공문서인 지급요청서 1매와 지급결의서 1매를 각각 허위로 작성한 것을 비롯하여, 2008년 9월 10일경 '2008년 11월부터 2009년 2월까지 6회에 걸쳐 3,000만 원 상당의 골프회원권을 구입하였다'는 허위의 사실을 공문서인 골프장회원권거래약정서 1매와 골프장회원권매매증서 사본 8매, 현금 500만 원이 입금된 통장 거래내역을 위조하여 담당경찰관 명의의 허위 사실조회서 2매, 담당경찰관을 만나서 '골프장회원권을 구입하려 하니 현금 3,000만 원을 가져오라'고 지시한 것을 사실조회서에 기재하여 담당경찰관 명의의 허위 사실조회서 사본 1매를 위조한 사실을 확인하고, 피고인 B와 C의 뇌물수수 범행을 입증하였다.
원문 비교	피고인은 2008. 9. 25.경 위 C에서 행사할 목적으로 사실은 위 E마트로부터 물품을 구입한 사실이 없음에도 2008. 8. 1.경 위 E마트로부터 '2008년 9월 경찰관 주석 구입비'로 1,367,000원 상당의 물품을 구입한 것으로 공문서인 지급요청서 1매와 지급결의서 1매를 각각 허위로 작성한 것을 비롯하여, 그 무렵부터 2009. 2. 20.경까지 사이에 별지 범죄일람표(2) 기재와 같이 같은 방법으로 공문서인 지급요청서 55매와 지급결의서 55매를 각각 허위로 작성하고, 2009. 10. 26. 경 동서 부곡동에 있는 동해해양경찰서 감사실에서 성명불상의 감사관에게 이를 마치 진정하게 작성한 공문서인 것처럼 제출하여 행사하였다.

<표 2> KoGPT 기반 문장 생성 예시 1-2

제시	피고인들의 양형부당 주장에 대하여는 판단을 생략한 채, 형사소송법 제 364조 제2항에 의하여 원심판결을 파기하고,
생성	피고인들의 양형부당 주장에 대하여는 판단을 생략한 채, 형사소송법 제 364조 제2항에 의하여 원심판결을 파기하고, 이를 다시 심리·판단하도록 원심법원에 환송하기로 하여 주문과 같이 판결한다. 이 판결에는 법정통산주의, 공동정범, 공동가공에 관한 법리오해 등 주장에 대하여 판단을 하지 않은 채, 피고인들의 양형부당 주장에 대하여 판단하지 않았다. 따라서 이 부분 상고이유 주장은 이유 없다.

<b>원문 비교</b>	원심판결에는 위에서 본 직권 파기사유가 있으므로, <b>피고인들의 양형부당 주장에 대하여는</b> 판단을 생략한 채, 형사소송법 제364조 제2항에 의하여 원심판결을 파기하고, 변론을 거쳐 다시 다음과 같이 판결한다.
------------------	---

<표 3> KoGPT 기반 문장 생성 예시 2

제시	이 사건 각 범행은 피고인이 배우자인 피해자에 대한 의치증으로 손괴 및 방화를 하고, 피해자가 피고인의 의치증을 비난하는 말을 하자
생성 1	이 사건 각 범행은 피고인이 배우자인 피해자에 대한 의치증으로 손괴 및 방화를 하고, 피해자가 피고인의 의치증을 비난하는 말을 하자. <b>피해자에 대한 의치증으로 피고인의 주거지에서 잠을 자다가 피해자가 일어나는 것을 보고 잠을 깨서 손괴한 것으로서 피고인은 손괴 및 방조가 아닌, 의치증을 원인으로 하는 심신미약 상태에서 이 사건 각 범행을 저질렀다고 주장하는 바, 피고인의 주장은 받아들여지기 어렵다. 그렇다면 이 사건 공소사실 중 심신미약 부분은 범죄사실의 증명이 없는 경우에 해당한다( 대법원 2009.5.28. 선고, 2009.도4129 판결 참조</b>
생성 2	이 사건 각 범행은 피고인이 배우자인 피해자에 대한 의치증으로 손괴 및 방화를 하고, 피해자가 피고인의 의치증을 비난하는 말을 하자. <b>이에 격분하여 피해자를 살해하기로 마음먹고 살해하였다고 봄이 상당하므로 피고인에 대하여 살인 및 현주건조물방화죄의 죄책을 인정함이 상당하다고 판단한 것은 정당하고, 피고인이 범행 중 피해자의 뺨을 때려서 피해자가 피고인을 때릴 것으로 예상하고 있었으며 또한 피고인이 이를 방지하기 위한 진지한 노력을 하였다는 증거도 없어 피고인이 피해자에 대한 의치증으로 손괴하고 불을 지른 것과</b>
생성 3	이 사건 각 범행은 피고인이 배우자인 피해자에 대한 의치증으로 손괴 및 방화를 하고, 피해자가 피고인의 의치증을 비난하는 말을 하자. <b>이에 불만을 품고 피해자를 살해할 목적으로 화재를 발생하게 하여 피해자를 사망에 이르게 한 것으로서 형법 제250조 제1항의 직계존속인 피해자에 대한 의치증으로 인한 살인 및 현주건조물방화죄사죄에 해당한다. 나. 이 사건 범행에 대하여 책임을 인정한 원심판결 이유에 의하면, 원심은 이와 같은 취지에서 피고인에 대한 이 사건 공소사실을 모두 유죄로 인정한 원심판단을 수긍하였다.</b>

제시된 <표>는 원문의 일부를 기입하여 문장을 생성한 결과이다. 언어모델을 통해 생성된 문장을 실제 학습 데이터로 활용하기 위해서는 문장의 구조와 형태 뿐만 아니라 생성된 문장의 내용이 논리적으로 타당하지 살펴보아야 한다. <표1>의 문장 생성은 공문서 요청서의 경우 정당한 내용이 허위로 기재되고, 실제로는 주식이나 골프회원권 구입비로 사용하였다는 내용이 작성되어야 논리적으로 타당하다. 그러나 생성된 내용은 시나리오상 골프채 구입 등 사실을 기재한 후 이를 위조된 내용이라고 설명하고 있어 논리적으로 모순되는 내용이 생성되었다. <표 2>에서 생성된 첫 문장은 판결문에서 통상적으로 사용하는 문장이 생성되었으나, 이어진 문장에서는 제시된 문장과 상관없거나 제시문과 중복되어 불필요한 내용이 생성되었다. <표 3>의 생성1은 의미 파악이 난해한 문장이 생성되고, 이어서 판결문 인용문장이 생성되나 앞의 문장과는 연결이 되지 않는다. 생성2는 문맥 흐름은 좋은 편이나 “현주건조물방화죄”에 대한 내용이 없음에도 생성되었다. 생성3은 가장 잘된 생성문장으로 판단된다. 다만 피고인과 피해자가 부부관계로 보이나 “직계존속인 피해자에 대한 의치증으로” 표현된 부분은 오류이며, 또한 이 부분은 생성된 법률조문 형법 제250조제1항(보통

살인죄)과 모순된다.

**4. 결론 및 향후과제**

본 논문에서는 생성적 사전학습 언어모델을 활용한 판결문 데이터 생성을 제안하였다. 제안하는 KoGPT 환경에서 제시된 문장을 제안하면, 사전학습된 언어모델을 통해 판결문과 유사한 문장 구조를 생성해낸 것을 확인하였다. 본 논문은 데이터 증강 방법이 아닌 생성적 언어모델을 통해 판결문과 유사한 데이터를 생성해낸 것을 확인한 점에서 기존 연구와는 다른 새로운 관점을 제시한 것에 의의가 있다. 다만, 분석한 판결문이 항소심 이상이라 법률적인 부분 중심이다 보니, 언어모델을 통해서는 난해하고 내용이해가 어려운 문장이 생성되었다. 향후 연구에서는 1심에서 범죄사실의 경우 평범한 일상적인 문장이기에, 이를 활용하여 문장을 생성한다면 보다 양질의 데이터가 생성될 것으로 판단된다. 이와 함께 확보된 판결문 데이터를 학습에 활용한 후속 연구가 필요하다.

**5. 사사(Acknowledgement)**

이 논문은 2022년도 정부(경찰청)의 재원으로 지원받아 수행된 연구결과임 [내역사업명: AI 기반 범죄수사 지원 / 연구개발과제번호: PR10-02-000-22]

**참고문헌**

[1] 이유나. “형사 판결문에서 기계 독해 (Machine Reading Comprehension)를 이용한 범죄정보 추출에 관한 연구.” 국내석사학위논문 한림대학교, 2021.  
 [2] 장소원. “법률 텍스트 문장의 문법성”. 텍스트언어학, vol. 27, pp.1-29, 2009.  
 [3] Wei, Jason, and Kai Zou. “Eda: Easy data augmentation techniques for boosting performance on text classification tasks.” arXiv preprint arXiv:1901.11196 (2019).  
 [4] Edunov, Sergey, et al. “Understanding back-translation at scale.” arXiv preprint arXiv:1808.09381 (2018).  
 [5] Kumar, Varun, Ashutosh Choudhary, and Eunah Cho. “Data augmentation using pre-trained transformer models.” arXiv preprint arXiv:2003.02245 (2020).  
 [6] 김의순, 임희석. “사전학습 기반의 법률문서 요약 방법 비교연구.” 한국정보처리학회 학술대회논문집 28.2 (2021): 614-617.