

공감형 대화 생성 인공지능 모델에 관한 연구

민치혁¹, 송혜진¹, 유보민¹, 이장안¹, 정호성¹, 조홍식¹, 유석봉^{2*}전남대학교 소프트웨어공학과 학부생¹전남대학교 인공지능융합학과 교수²

asdf7270@jnu.ac.kr thd3566@jnu.ac.kr 201780@jnu.ac.kr lja3723@jnu.ac.kr

wjdghtjd06@jnu.ac.kr 201037@jnu.ac.kr sbyoo@jnu.ac.kr

A Study on the Artificial Intelligence Model for
Empathy Conversation GenerationChi-Hyeok Min¹, Hye-Jin Song¹, Bo-Min Lyu¹, Jang-An Lee¹, Ho-Seong
Jeong¹, Hong-Sik Cho¹, Seok-Bong Yoo²Dept. of Software Engineering, Chonnam National University¹Dept. of Artificial Intelligence Convergence, Chonnam National University²

요 약

이 연구는 KoGPT2(Korean Generative Pre-trained Transformer)를 활용해 아동 및 청소년의 의사소통을 지원하는 공감형 대화 모델을 개발하는 것을 목표로 한다. 이를 위해 소아청소년 심리 데이터를 기반으로 KoGPT2를 파인튜닝하고, 향후 음성 및 이미지 데이터를 포함한 멀티모달 모델로 확장할 계획이다. 이로써 아이들의 정서와 의사소통 능력을 다각도로 지원하는 AI(Artificial Intelligence) 시스템을 구축하고자 한다.

1. 서론

NLP(Natural language processing) 분야는 트랜스포머(Transformer) 모델의 도입으로 급격한 변화를 맞이했다. 2017년에 이 기술이 처음 제안된 이후, 트랜스포머는 NLP 전반에 걸쳐 폭넓게 활용되며, 해당 분야에 지대한 영향을 미쳤다. KoGPT2는 GPT2[2] 아키텍처를 기반으로 하며, 주어진 텍스트의 다음 단어를 예측하는 데 최적화된 디코더 기반 언어 모델이다. 한국어 성능을 강화하기 위해 40GB 이상의 한국어 텍스트로 학습되었으며, 대규모 데이터에서 자연스러운 문장 생성을 가능하게 한다. BERT(Bidirectional Encoder Representations from Transformers)[3]보다 더 긴 학습 시간과 더 많은 데이터로 훈련되어, 문장의 깊은 의미와 문맥을 파악하는 능력이 뛰어난 RoBERTa(Robustly Optimized BERT Pretraining Approach)는 감정 분류를 가능하게 한다. 이 두 모델을 결합함으로써, KoGPT2는 자연스러운 대화형 문장 생성을, RoBERTa는 문맥 내에서의 의미 이해를 강화하여, 상담 데이터를 더욱 효과적으로 처리하고, 아동 및 청소년의 의사소통 능력을 발달시키는 데 중요한 역

할을 할 것으로 기대된다. 따라서 본 연구의 목표는 한국어에 특화된 KoGPT2와 RoBERTa 모델을 함께 활용하여 아동 및 청소년 상담 데이터를 기반으로 파인튜닝을 수행함으로써, 이들의 의사소통 능력이 원활하게 발달할 수 있도록 지원하는 것이다.

2. 제안하는 연구 주제의 필요성

2022년 맞벌이 가구 비율이 46.1%에 달하면서, 아이를 돌보는 방식이 다양해졌다. 많은 부모들이 가사도우미나 학습용 태블릿 등 다양한 방법을 사용하지만, 특히 디지털 기기에 의존하는 경우가 많다. 이로 인해 "디지털 기기가 과연 아이에게 친화적인가?"라는 의문이 제기되었다. 일부 사례에서는 아이가 기본적인 표현을 기계적 언어로만 하는 극단적인 상황이 발생하기도 한다. 이러한 문제를 해결하기 위해, 양방향 소통이 가능한 대형 언어 모델, 음성인식, 음성 합성 기술을 활용하여 아이들의 언어 발달에 도움을 줄 수 있는 대화형 인형을 개발하는 것이 필요하다고 판단했다.

3. 기존 관련 연구

3-1. 트랜스포머 기반 언어모델

트랜스포머 아키텍처는 자연어 처리 분야에서 혁신

* 교신저자

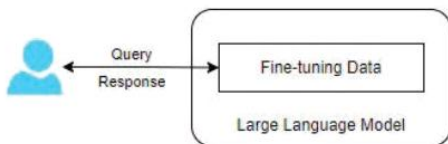
적인 변화를 이끌어낸 모델로, 병렬 처리가 가능하다는 강점 덕분에 대규모 데이터셋을 효과적으로 학습할 수 있다. 특히, 트랜스포머 기반 언어모델은 자기 회귀적 방식으로 텍스트의 다음 토큰을 예측하거나, 입력된 문장의 전체적인 의미를 파악하는 데 뛰어난 성능을 보인다. 이러한 특성은 자연어 생성, 기계 번역, 문서 요약 등 다양한 NLP 응용 분야에서 탁월한 결과를 가져왔으며, 최근 연구는 이 모델을 더욱 고도화하여 언어 간 편향 해소, 문맥 이해력 증대와 같은 문제 해결에 집중하고 있다.

3-2. 한국어 특화 언어모델

한국어의 고유한 언어적 특성(어순의 유연성, 조사 사용, 어휘의 다의성 등)을 반영한 한국어 특화 언어모델에 대한 연구는 최근 활발히 이루어지고 있다. 특히, KoGPT, KorBERT 등의 모델은 한국어 문법 구조와 맥락 이해에 최적화된 사전 학습된 모델로, 한국어 자연어 처리의 성능을 크게 향상시키고 있다. 이러한 모델들은 뉴스, 웹 문서, 대화 데이터 등 다양한 도메인의 데이터를 학습하여 텍스트 분류, 감정 분석, 기계 번역 등에서 매우 높은 정확도를 보이고 있다. 나아가, 한국어 특화 모델은 문화적, 사회적 맥락을 반영한 대화 시스템이나 맞춤형 언어 서비스 개발에도 필수적인 요소로 자리잡고 있다.

4. 제안하는 설계 방향

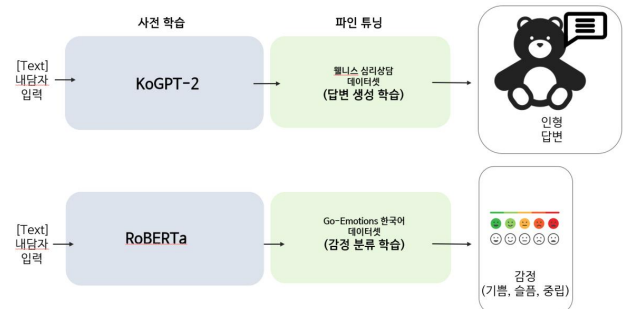
본 연구에서는 KoGPT2를 활용한 공감형 대화 모델을 설계하고자 한다. 소아청소년 심리검사 및 뇌영상 데이터를 통합하여, 공감형 대화 모델을 구축하는 데 중점을 두고 있다. 이를 위해 Hugging Face 라이브러리를 활용하여 KoGPT2 모델을 파인튜닝할 계획이다. 심리 상담 상황에서 아이들과의 정서적 교감을 원활히 이끌어낼 수 있도록, 대화의 맥락을 이해하고 공감할 수 있는 대화형 AI를 구현하는 것이 목표이다.



(그림 1) Fine-tuning[1]

또한, RoBERTa 모델을 활용하여 아이들이 입력한 문장을 분석하고, 그 문맥에서 나타나는 정서적 표현을 파악하는 기능을 추가할 예정이다.

RoBERTa는 아이들의 대화 내용을 깊이 있게 이해하고, 문장 내에 포함된 정서적 신호를 분석하는 데 사용된다. 이를 통해 아이들의 감정 상태를 실시간으로 평가하고, 해당 감정에 적절하게 반응하는 공감형 대화를 생성함으로써, 더 나은 상담 경험을 제공할 수 있는 AI 시스템을 설계하는 것이 연구의 또 다른 방향이다.



5. 향후 계획

텍스트뿐만 아니라 음성 및 이미지 같은 다양한 데이터를 처리할 수 있는 멀티모달 모델로 확장할 계획이다. 이를 통해 음성 및 시각적 정보를 종합하여 더 풍부한 상호작용이 가능하도록 하며, 아이들의 정서와 의사소통 능력을 다각도로 지원하는 포괄적인 솔루션을 제공할 예정이다.

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 소프트웨어중심대학사업, 인공지능융합혁신 인재양성사업, 대학ICT연구센터사업의 연구결과로 수행되었습니다. (2021-0-01409, 2023-00256629, 2024-00437718)

참고문헌

- [1] Park, D., An, G. T., Kamyod, C., Kim, C. G., "A Study on Performance Improvement of Prompt Engineering for Generative AI with a Large Language Model," Journal of Web Engineering, vol. 22, no. 8, pp. 1187-1206, 2023
- [2] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever, "Language models are unsupervised multitask learners," 2019.
- [3] Devlin, J. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv 2018 arXiv:1810.04805.