

청소년 진로 상담을 위한 LLM 기반 챗봇 설계에 관한 연구
- 진로문장완성검사 데이터를 기반으로

이원태, 윤수연*

국민대학교, *국민대학교

lwt1025@kookmin.ac.kr, *1104py@kookmin.ac.kr

A Study on LLM-based Chatbot Design for Adolescent Career Counseling

Won Tae Lee, Soo Yeon Yoon*

Kookmin Univ., *Kookmin Univ.

요약

본 연구는 고등학교 진로상담 현장에서 학생의 진로성숙도를 효율적으로 파악하고 진로계획수립을 지원하기 위한 챗봇 설계 가능성을 탐색하였다. 기반한 데이터는 진로문장완성검사(Career-SCT)의 불완전구조 자유응답 데이터이며, 이를 통해 고등학생의 진로성숙도를 파악하고 LLM의 후속 상담 응답을 위한 기초 입력을 생성할 수 있는지 실험하였다. 벡터 유사도 기반 검색과 기초적인 챗봇의 구현을 위해 BM-K/KoSimCSE-RoBERTa 임베딩 모델과 GPT-4 모델을 활용했으며, 이러한 구현을 통해 진로성숙도 분석과 후속 상담 응답 생성을 위한 기반 기술로서의 활용 가능성을 모색하였다. 진로 상담 분야에서 챗봇 도입을 통한 개별 진로 지원을 탐색한 기초 연구로서, 향후 추가 연구의 기반이 될 것으로 기대한다.

I. 서론

고등학생은 진로 관련 고민이 가장 큰 시기임에도 불구하고 교육 현장에서 충분한 진로상담자를 받기 어려운 경우가 많다. 특히 코로나19 이후 비대면 상황에 익숙한 현 학생 세대에게는, 모바일로 접근 가능하고 대화형 상호작용이 가능한 챗봇이 유용한 진로 탐색 도구가 될 수 있다. 최근 범용 대규모 언어모델(LLM)의 발전으로 진로 관련한 간단한 질의는 누구나 가능하게 되었다. 하지만 이런 범용 모델은 개별 진로 배경을 구조적으로 반영하기 어렵다는 한계가 존재한다.

본 연구는 진로문장완성검사(Career-SCT) 데이터를 기반으로 챗봇이 개별 학생의 진술을 구조화하고 진로성숙도를 추론하여 효율적인 상담을 제공하는 가능성을 실험한다. 구조화된 자유 입력을 활용함으로써, 개인 특성에 대해 전문적 상담 인력의 해석과 유사한 프롬프트를 생성하여 상담의 기초입력으로 사용한다. 이는 개별 직접 상담이 어려운 환경에서도 기본적인 진로지원을 가능케 하는 챗봇의 기초가 될 수 있다.

II. 관련 연구

2.1 진로문장완성검사

진로문장완성검사(Career-SCT)는 “내가 생각하기에 나는...” 과 같은 불완전 문장을 제시하고, 피검자가 이를 자유롭게 완성토록 유도하는 진로성숙도 평가 도구이다. 학생의 진로에 대한 생각, 감정 등을 구조화된 형태로 서술토록 함으로써 진로 관련 내적 요소를 파악할 수 있다. 상담자는 완성문장에 투사된 피검자의 진로 인식과 심리적 맥락을 상담 현장에서 유의미한 자료로 활용한다.

2.2 LLM 기반 챗봇 기술 선행 연구

최근 GPT4.5, Gemini 등 대규모 언어모델(LLM)의 발전과 함께 다양한 분야에서 LLM을 접목한 챗봇 연구가 활발해지고 있다. 특히 교육, 의료,

상담 등 대상 개인의 맥락과 특성의 파악이 중요한 분야에서 입력 텍스트의 의미적 유사성을 벡터 기반으로 검색하여 종합하고, 이를 LLM의 입력으로 활용하여 응답의 정확성을 높이는 연구가 다양해지고 있다[1][2]. 이러한 접근법은 단순히 LLM의 생성 능력만을 이용하는 방식에 비해, 미리 구축된 신뢰 가능한 데이터의 활용을 통해 응답의 신뢰성을 확보할 수 있다. 상담 및 진로 영역에서도 학생들의 서술형 응답이나 상담 사례를 임베딩하여 유사 사례를 검색하고, 이를 바탕으로 맞춤형 상담 응답을 생성하는 사례들이 시도되고 있다[3].

2.2.1 벡터 기반 검색 및 LLM 활용

본 연구에서 제안하는 챗봇은 학생의 진로문장 응답을 임베딩하여 벡터화하고, 이를 기반으로 사전에 구축된 벡터 저장소 내에서 의미적으로 유사한 응답을 검색한다. 이 과정에서 검색된 문장과 연결된 전문가 평가와 진로성숙도 정보를 추출하여 사용자의 진로 상태를 진단한다. 최종적으로는 이 정보를 초기 맥락으로 활용하여, LLM 모델 GPT-4가 학생 개개인의 상태에 적합한 맞춤형 상담 응답을 생성하도록 설계하였다. 이러한 접근법은 비정형 텍스트 데이터에서 의미적 유사성을 기반으로 신뢰할 수 있는 정보를 효과적으로 활용할 수 있어, 개인화된 상담 챗봇 구축에 유리하다.

III. 실험 및 결과

3.1 실험 환경 및 데이터셋 구성

본 연구는 AI-Hub에서 제공하는 진로문장완성검사(Career-SCT) 텍스트 데이터셋을 기반으로 했다. 해당 데이터셋은 [표 1] 과 [표 2]에서 살펴볼 수 있듯이, 진로문장완성검사 응답 데이터와 각 문항영역별 진로성숙도 분류(상-중-하), 전문가 해석 어노테이션으로 구성된다. 이 중 고등학생의 데이터만 추출한 응답은 15,764 건이며, 실험 환경은 [표 3] 과 같이 구성하였다.

[표 1] 데이터셋 개요

데이터 종류	형태	데이터 셋 규모
진로문장완성검사 응답 데이터	json	40,422
피검자 메타정보 데이터	json	1,322
진로문장완성검사 카테고리 메타데이터	json	59
진로문장완성검사 문항별 해석 데이터	json	40,422

[표 2] 데이터셋 분포

문항 영역(카테고리) 분포		
비고	기록 수	비율
자기이해 및 긍정적 자아상	8,415	20.82%
대인관계 및 의사소통 역량	6,958	17.21%
진로 의사결정능력	1,908	8.30%
... 외 11개 영역
총합	40,422	100.00%

[표 3] 실험 환경 구성

항목	내용
임베딩 모델	BM-K/KoSimCSE-RoBERTa (768차원 벡터)
프롬프트 생성 로직	벡터 기반 유사도 검색(코사인 유사도)
LLM 모델	GPT-4 (OpenAI API 사용)
하드웨어 환경	Apple M1 MacBook Air (RAM 8GB)
개발 환경	macOS Sonoma, Python 3.10
사용 라이브러리	Hugging Face Transformers, PyTorch(MPS)

3.2 실험 내용

챗봇은 크게 다음 두 단계로 구성된다. 첫째, 사용자의 진로문장완성검사 응답을 받아 미리 임베딩한 벡터 저장소의 유사도를 기반으로 진로성숙도 수준과 전문가 코멘트를 추론하고, 이를 종합한다. 둘째, 이러한 추론 결과를 기반으로 후속 진로 관련 질의에 대해 LLM 기반 상담을 제공한다. 이를 구성한 임베딩 모델과 유사도 기반 검색 구성, LLM 프롬프트 입력 설계는 아래와 같다.

3.2.1 임베딩 모델

문장을 벡터화 하기 위해 임베딩에 BM-K/KoSimCSE-RoBERTa 모델을 사용하였다. 해당 모델은 한국어 문장 임베딩에 특화되었으며, KLUE STS 데이터 등 한국어 문장 유사도 데이터셋으로 학습되었다. 타 모델에 비해 문장간 의미 파악에 강점이 있기 때문에 진로문장완성검사의 완성문장에 대한 임베딩을 생성하는데 적합하다. 이 모델을 통해 개별 문항의 완성문장과 피검자 개인의 문항영역별 전문가 코멘트를 임베딩하고 벡터로 저장하였다.

3.2.2 벡터 기반 유사도 검색

사용자 입력 문장은 완성문장 형태로 변환된 후 동일한 768차원 벡터로 임베딩 후 벡터 저장소의 완성문장과 코사인 유사도를 비교한다. 이 중 상위 유사도 완성문장의 전문가 코멘트와 진로성숙도 라벨을 추출하고, 이들의 진로성숙도 수준 및 해석 코멘트를 종합하여 진로성숙도 초기 진단 정보를 생성한다.

3.2.3 LLM 프롬프트 입력 설계

초기 진단 정보는 사용자 메타데이터 및 진로성숙도 수준 요약 및 대표 문항의 해석 코멘트로 구성되며, 이는 LLM에게 사용자 상태에 대한 맥락 정보를 제공하는 프롬프트로 변환하여 사용된다. 이후 사용자가 입력하는 진로 관련 질문에 대해 LLM은 초기 진단 결과를 반영한 상담 응답을 생

성한다. 본 연구에서는 LLM 모델로 GPT-4를 사용하였다.

3.3 실험 결과 분석

본 연구에서 설계한 진로문장완성검사 데이터 및 LLM 기반 진로 상담 챗봇의 초기 구현으로 몇 가지 유의미한 결과를 확인할 수 있었다. 첫째, 챗봇은 KoSimCSE-RoBERTa 임베딩 모델과 유사도 기반 검색 구조를 활용하여 학생들의 자유 서술형 응답과 의미적으로 유사한 기존 사례를 추출하여 전문가 평가와 유사한 진로성숙도 초기 진단을 생성하였다. 챗봇의 진로성숙도 추론 결과는 [표 4]에서 알 수 있듯이, 원 데이터의 전문가 평가 등급(상·중·하)과 대체로 일치하는 경향을 보였으며, 특히 극단적 수준(상·하)의 응답에서 보다 정확한 등급 일치성을 나타냈다.

[표 4] 진로성숙도 추론 결과

항목	precision	recall	f1-score	support
상	0.84	0.81	0.82	540
중	0.67	0.71	0.69	540
하	0.75	0.72	0.73	190

둘째, GPT-4를 활용한 후속 상담 응답 생성은 초기 진단 정보를 적절히 반영하여 학생 개인의 특성을 고려한 자연스러운 상담 문장을 생성할 수 있음을 확인하였다. 특히 학생의 고민과 관심 분야가 명확히 드러난 입력으로 생성된 프롬프트에 대해서는 후속 상담 응답 초기부터 구체적 주제를 제시하는 등 적극적 반응을 보였으나, 입력이 비교적 추상적인 입력을 통해 생성된 프롬프트에 대해서는 생성된 초기 응답의 구체성과 적합성이 다소 떨어지는 경향을 보였다.

IV. 결론

본 연구는 진로문장완성검사(Career-SCT)의 완성문장 응답 데이터를 활용하여 벡터 유사도 검색과 LLM 기반의 진로상담 챗봇을 설계하고, 그 가능성을 초기적으로 확인하였다. 상담 자원 부족과 개인화된 진로 지원 수요에 대응하여, 비대면에 익숙한 세대에게 접근성 높은 챗봇 형태를 구현한 것에 연구의 의의가 있다. 또한 본 연구는 진로 영역 외의 다른 SCT에도 확장 가능하며, 상담 분야를 넘어 개인화된 맥락이 중요한 다양한 분야에 적용될 수 있는 방법론적 가능성을 제시한다. 의료, 교육, 영업 등 개별 객체의 맥락이 핵심적인 분야라면, 적절히 구조화된 데이터셋을 바탕으로 챗봇과 같은 즉시성 있는 도구를 설계할 수 있는 출발점이 될 수 있다. 이와 같이 본 연구는 학생 개별의 상담 이력과 상호작용 데이터를 장기 기억 구조에 통합함으로써, 향후 보다 연속적이고 적응적인 상담 흐름을 갖춘 지능형 상담 에이전트로의 발전 가능성을 제시한다.

참 고 문 헌

[1] Gao, L. et al. (2023). "Dense Passage Retrieval for Open-Domain Question Answering". ACL.

[2] Reimers, N., & Gurevych, I. (2019). "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks". EMNLP.

[3] Rasool, A., et al. (2024). "Emotion-Aware Response Generation Using Affect-Enriched Embeddings with LLMs". arXiv preprint arXiv:2410.01306.