

# 사용자 의도를 고려한 생성형 인공지능 서비스 모델 개발

최유빈, 최승원, 김승후, 김현빈, 김태건, 윤태복  
서일대학교 AI융합콘텐츠학과  
e-mail:tbyoon@seoil.ac.kr

## Development of generative artificial intelligence service models with consideration of user intent

Youbin Choi, Seungwon Choi, Seunghoo Kim, Hyeonbeen Kim,  
TaeGeon Jeong, and Taebok Yoon  
Dept. of AI Convergence Contents, Seoil University

### 요약

최근 생성형 인공지능의 관심이 급격하게 높아지고 있다. 데이터, 콘텐츠, 이미지 생성 등 다양한 산업 분야에 서비스 되고 있으나 사용자의 질의에 따라 답변 결과가 상이하여 질의 문장을 잘 구성하는 것이 중요한 상황이다. 하지만, 생성형 인공지능 서비스를 이용하는 이용자는 배경지식이 다양하고, 서비스 경험이 많지 않기 때문에 원하는 결과를 얻기 위하여 어떻게 문장을 구성해야 하는지 이해하는 데는 한계가 있다. 본 연구는 생성형 인공지능 서비스 모델을 생성하는 과정에 사용자 의도를 고려한 모델을 적용함으로써 서비스 이용자들이 일상적인 용어를 이용하여 질의하더라도 그 의도에 근접한 답변을 할 수 있는 방법을 제안한다. 이를 위하여 사용자의 다양한 질문에 의도를 분류할 수 있는 모델을 개발하고 그 결과를 미세조정(Fine-tuning)에 활용하였다. 실험에서는 커뮤니티 데이터를 이용하여 의도 분류 모델을 생성하였으며 질의에 따라 의도를 파악하고 적절하게 서비스하는 결과를 확인하였다.

### 1. 서론

챗봇(Chatbot)이란 음성이나 문자를 통한 인간과의 대화를 통해서 특정 작업을 수행할 수 있도록 제작된 인공지능(Artificial Intelligence)을 의미한다. 최근 인공지능 기술과 자연어 처리(Natural Language Processing) 기술의 발전함에 따라 다양한 분야에서 많이 사용되고 있다. 기업이나 공공기관, 그리고 학교에서 업무 자동화를 위한 챗봇을 도입하는 사례가 증가하고 있다. 몇 가지 사례를 확인해보면 챗봇을 사용하기 전에 사용자가 궁금한 내용이 있으면 담당자에게 이메일 또는 유선통화로 해결했어야 한다면 최근에는 사용자가 궁금한 내용을 챗봇을 통해 빠른 해결이 가능하게 되었다. 이러한 챗봇 서비스를 만들기 위하여 이용하는 모델 중 하나인 GPT(Generative Pre-trained Transformer)는 사전 훈련된 트랜스포머(Transformer) 모델을 기반으로 한 자연어 처리 모델이며, 사용자와의 대화를 처리하고 응답을 생성하는 데에 특화되어 있다. 사용자가 궁금한 내용이나 알아보고 싶은 내용을 GPT 모델에게 전달하면 사용자의 질의에 따라 원하는 답변을 내놓아서 만족도가 높다. 하지만 GPT 모델에게 질문을 할 때 사용자마다 의도한 바와 질문하는 문장이 상이하여 이를 해결하기 위한 방안이 필요하다. GPT 미세 조정을 진행할

때 사용자의 질문 유형을 모두 학습시킬 수 없어 의도를 파악할 수 있도록 사용자 의도를 분류할 수 있는 모델을 제안한다.

### 2. 관련 연구

챗봇은 시간과 공간의 제약 없이 실시간으로 소통하며 개인 맞춤형 서비스가 가능하다. 최근 공공기관, 민간 등에서 고객 서비스를 위해 챗봇을 도입하고 있는 사례가 많이 등장하고 있다. 별도의 플랫폼을 사용하지 않고 기존의 SNS, 웹사이트 내에서 진행하고 있다[1]. 대학에서는 반복되는 학사일정들이 매년 진행되는 데, 신입생이나 재학생들은 학사 일정에 관한 질문을 하게 된다. 반복되는 질문으로 인하여 발생하는 교직원들의 업무 부담을 줄이고, 학생들에게 필요한 정보를 제공할 수 있는 챗봇을 도입하고 있다[2]. 하지만, 챗봇의 응답이 항상 정답을 제공하는 것은 아니다. 종종 논리적으로 보이긴 하지만 틀린 정보나 의미 없는 답변을 제공할 수 있다. 이는 답을 찾는 것이 아니라 데이터 간의 관계를 분석하고 답을 생성하는 형태로 이루어져 있기 때문이다[3]. 이를 보완하기 위해서는 안전하고 신뢰성 있는 학습 데이터를 제공하는 것이 중요하다[4].

### 3. 사용자 의도를 고려한 데이터 생성

제안하는 방법은 사용자의 의도를 분류할 수 있는 모델과 적용 기관과 관련된 의도를 파악할 수 있는 모델을 만들고, 미세 조정된 GPT 모델에게 질문하였을 때 사용자가 원하는 답변 받을 수 있도록 설계하는 것을 목표로 한다.

#### 3.1 모델 개발

모델을 제작하기 위해서는 초기 데이터가 필요하다. 예를 들어 학교의 학생들은 소속 학교에 관하여 궁금한 점이 있다면 주로 커뮤니티나 학교 네이버 카페, 카카오톡 오픈채팅방에서 질문하는데, 이러한 데이터를 수집하고 사용하게 된다면 모델 제작에 의미 있는 데이터라고 볼 수 있다. 본 연구에서는 학생들이 주로 이용하는 커뮤니티 게시물과 네이버 카페 게시물 데이터들을 스크래핑(Scraping)하였다. 약 42만 데이터 수집하였으며 이 중 필요 없거나 의미 없는 데이터들을 전처리 및 데이터 불균형을 해결한 후 모델을 개발하였다.

##### 3.2.1 데이터 전처리

무의미한 데이터들을 제거하기 위해서 다음과 같은 절차를 걸쳤다. 첫째, 링크 또는 이모티콘만 존재하는 데이터를 제거하였다. 이를 위하여 링크와 이모티콘의 고유패턴을 분석하여 불필요한 링크와 이모티콘을 제거했다. 둘째, 초성만 존재하는 데이터를 제거했다. 한글 음절의 유니코드 범위 밖에 있는 데이터들은 모두 제거했다. 앞에서 소개한 데이터 전처리 과정을 거쳐 기존 42만 라인에서 20만 라인으로 감소하였다.

##### 3.2.2 레이블 생성과 데이터 불균형

분류 모델을 만들기 위해서는 전처리 데이터를 이용하여 레이블 분류해야 한다. n개의 모델을 만들기 위해선 레이블 분류하는 작업도 n가지로 나뉘어서 작업해야 한다. 학교 관련 분류 모델에서의 레이블은 학생들이 주로 궁금해하는 핵심 단어들로 추렸다. 예를 들어 “성적 확인하고 싶으면 어디서 봐?”라는 질문을 받았을 때 학교 관련 분류 모델에서는 “성적”이라는 단어를 추출하고, 사용자 의도 분류 모델에서는 “확인”이라는 단어를 추출한다. 학교 관련 분류 모델에서는 총 13가지의 레이블(휴학, 전과 등)을 구성하였고, 학교 관련 모델에서는 총 9가지의 레이블(장소, 번호 등)을 구성하였다.

레이블 추가된 데이터들의 균형이 이뤄지지 않아 업샘플링(UpSampling)을 하였다. 클래스의 데이터를 중복으로 추출하여 샘플(Sample) 수를 가장 많은 레이블과 동일하게 만들고, 재현할 수 있는 랜덤 시드(Random Seed)를 설정하여 재샘플링(ReSampling) 작업도 수행했다.



[그림 1] 미세 조정 작업에 필요한 데이터 세트 예시

#### 3.2.4 모델 훈련과 미세 조정

레이블 데이터를 기반으로 Naive Bayes 분류기를 사용하여 모델을 훈련했다. 먼저 텍스트 데이터를 TF-IDF 피쳐(Feature)로 변환하였다. TF-IDF는 텍스트 데이터의 특성을 추출할 때 쓰는 방법이며, 텍스트 데이터를 벡터 형태로 변환할 때 사용한다.

GPT 미세 조정 작업을 하려면 먼저 GPT에 학습시킬 데이터 세트(Dataset)를 준비해야 한다. 본 연구에서 사용한 데이터 세트는 학교 학칙을 기반으로 테스트하였다. 데이터 세트는 그림 1과 같이 질문과 답변을 하나의 세트로 구성하여 나뉜다.

### 4. 결론

본 연구에서는 생성형 인공지능 서비스 모델의 질을 향상시키기 위해 사용자 의도를 고려한 모델을 적용하는 방법을 제안하였다. 이를 위하여 사용자의 다양한 질문에 의도를 분류할 수 있는 모델을 개발하고 그 결과를 미세조정에 활용하였다. 실험 결과, 커뮤니티 데이터를 이용하여 생성한 의도 분류 모델은 질의에 따라 의도를 파악하고 적절하게 서비스를 수행하는 것으로 나타났다. 향후 연구로는 더욱 다양한 서비스가 가능하도록 의도 레이블링의 범위를 확장하고 제안하는 방법의 유효성을 판단할 수 있는 추가 실험이 요구된다.

#### 참고문헌

- [1] 박준호, 윤경일, 민성태, “Korea information processing society review”, KIPS Transactions 2019.07
- [2] Tai-Sung Hur, Gi-Yun Mok, Jun-Ho Kim, Sang-Heong Baek, Jee-Hoon Le, “University Bulletin Chat-bot Systembased on Natural Language Processing”, KSCI, 2020. 1
- [3] 김태원, “챗GPT가 촉발한 생성형 AI 시대, 미래 모습과 대응 방안”, FUTURE HORIZON 2023.03
- [4] Jang Mook Kang, Cho Min Je, Jo Dongbaig, Kim NamSun, “Reflections on Hallucination Security in an Autonomous Vehicle Service (in terms of preventing incorrect answers and biased judgments from AI Hallucination)”, JKIT 2023.06