

BERT 모델을 활용한 고객 질의 응답 챗봇 구현

사수진, 윤성철*, 전명근**

충북대학교 산업인공지능학과, *충북대학교 산업인공지능연구센터, **충북대학교 지능로봇공학과

sasujin@gmail.com, *steveyun@chungbuk.ac.kr, **mgchun@chungbuk.ac.kr

Implementing a Customer Q&A Chatbot Using the BERT Model

Su-Jin Sa, Sung-Chul Yun*, Myung-Geun Chun**

Department of Industrial AI, Chungbuk Univ., *Industrial AI Research Center, Chungbuk Univ.,

**Department of Intelligent Systems and Robotics, Chungbuk Univ.

Abstract

This paper describes the development of a customer query-response chatbot using BERT (Bidirectional Encoder Representations from Transformers). The main objectives are to reduce the workload of customer support teams, provide 24/7 assistance, and ensure consistent and accurate responses to user inquiries. By fine-tuning BERT to understand customer queries, the chatbot delivers fast and personalized responses, enhancing user satisfaction. Additionally, this study emphasizes the effectiveness of specialized models such as KoBERT for handling Korean query processing in customer service.

Key words : Chatbot, BERT, Multilingual BERT, KoBERT, Natural Language Processing

I. 서론

고객 문의의 증가로 인해 효율적인 고객 지원이 조직에 점점 더 어려운 과제가 되고 있습니다. 최근 1년간 고객 문의 건수가 30% 증가하였으며, 이로 인해 평균 응답 시간이 20% 늘어났습니다. 이러한 증가는 기존 고객 서비스 지원팀의 업무 부담을 가중시키고 있으며, 인력 확충 없이 효율성을 높이는 방안이 시급합니다. 대규모 고객 지원 팀을 유지하는 것은 비용이 많이 들며, 사람에 의한 응답은 종종 일관성이 부족합니다. 자연어 처리(NLP)의 발전을 활용한 챗봇은 고객 지원 인력의 업무 부담을 줄이고, 효율성을 높이며, 응답 시간을 개선할 수 있는 잠재적인 해결책을 제공합니다.[1]

본 논문은 BERT 모델을 사용하여 제품 매뉴얼 내용을 활용한 사용자 질의응답을 자동화하는 챗봇을 개발하는 것을 목표로 합니다. 기존의 고객 서비스 지원은 주로 고객의 요구에 맞춘 인력 배치와 비대면 상담에 의존해 왔습니다. 그러나 디지털 전환과 비대면 서비스의 수요가 증가함에 따라, 이러한 전통적인 방법은 한계에 부딪히고 있습니다. 이에 따라, 더 나은 고객 경험을 제공하기 위해 AI 기반 챗봇의 도입이 점차 중요해지고 있습니다. 본 연구는 BERT 모델을 활용한 챗봇 개발이 이러한 문제 해결에 중요한 역할을 할 수 있음을 시사한다.

II. 본론

본 논문은 기존 고객 서비스 운영의 비효율성을 해결하는 것을 목표로 한다. 현재 고객 문의 처리 시간이 길어짐에 따라 고객 불만이 증가하고 있으며, 숙련된 인력의 부족으로 인해 응답 품질의 일관성 저하 문제가 발생하고 있다. 또한, 업무량 증가, 팀원의 경험 수준에 따른 응답의 일관성 부족, 대규모 지원 팀 유지에 따른 높은 비용 등의 문제에 직면해 있다. 챗봇은 24시간 일관되고 효율적인 응답을 제공함으로써 운영 비용을 절감할 수 있다. 실시간으로 고객이 원하는 정보를 제공함으로써 고객 만족도를 향상시키고, 반복적인 문의 처리를 자동화하여 직원들이 보다 복잡한 문제 해결에 집중할 수 있도록 한다. 이는 고객 서비스 품질 향상과 직원의 스트레스 감소로 이어지며,

전체적인 운영 효율성을 증대시킨다. 특히, AI 기반 챗봇은 인간 에이전트가 처리하기 어려운 대량의 요청을 신속하게 처리할 수 있어, 고객 질의응답의 접근성과 효율성을 극대화할 수 있다.

본 논문에서는 자연어 처리(NLP) 분야에서 우수한 성능을 보이는 BERT 모델을 활용하였다. BERT는 트랜스포머(Transformer) 아키텍처를 기반으로 하며, 양방향 문맥을 이해할 수 있는 특징을 통해 사용자 문의 의도를 정확하게 파악하는 데 유리하다. 사전 학습된 대규모 코퍼스를 바탕으로 다양한 언어적 패턴과 의미를 학습한 BERT는 특정 도메인에 맞게 파인튜닝(Fine-tuning)함으로써 고객 질의응답에 최적화된 성능을 발휘한다.[2] 특히, BERT의 문장 이해 능력은 복잡한 문의 사항에 대한 정확한 응답 생성 및 관련 정보의 신속한 추출에 기여하여 챗봇 응답 품질을 크게 향상시킨다. 이러한 특성은 높은 정확도와 신뢰성을 바탕으로 고객 질의응답의 효율성을 극대화할 수 있게 한다. 본 연구에서는 BERT를 제품 매뉴얼에서 추출한 맞춤형 데이터셋으로 파인튜닝하여 고객 문의에 정확히 응답할 수 있도록 하였다. 또한, 다국어 BERT 모델과 더불어 한국어 문법과 어휘 처리를 개선하기 위해 BERT의 한국어 버전인 KoBERT를 사용하였다.[5]

```
{
  "context": "문서나 글의 일부, 질문의 배경이 되는 내용.",
  "question": "질문 내용",
  "id": "고유한 id값",
  "is_impossible": "문맥 내에서 질문에 대한 답변이 존재 유무(true/false)",
  "answers": {
    "text": "[답변 내용]",
    "answer_start": [문서 내에서 답변이 시작되는 위치 (index)]
  }
}
```

그림 1 데이터 포맷 예시(KorQuad 한국어 질의응답 데이터셋) [4] 챗봇 학습에 사용된 데이터셋은 제품 매뉴얼에서 추출한 질문과 답변으로 구성된다. 총 1,200개의 Q&A 쌍을 JSON 형식으로 생성하였으며, 이 데이터셋은 학습용과 검증용으로 8:2 비율로 분할하여 모델 학습에 사용하였다. 데이터 전처리 과정에서는 질문과 답변의 길이 분석, 불필요한 공백 제거,

오타 수정 등의 작업을 수행하였다. 각 질문에는 고유한 ID가 부여되었으며, 답변의 시작 위치를 명확히 지정하여 모델이 문맥을 정확히 이해할 수 있도록 하였다. 데이터셋은 다양한 유형의 질문을 포함하여 모델이 광범위한 시나리오에서 효과적으로 작동할 수 있도록 설계되었다.

본 연구에서 성능 평가는 EM(Exact Match)과 F1 스코어와 같은 지표를 사용하여, 챗봇의 응답이 예상 답변과 얼마나 일치하는지를 측정하였다.

표 1 성능 평가 환경

Table 1 Performance Evaluation Environment

구분	내용
환경	Laptop
프로세서	Apple M3 Pro칩(11코어 CPU)
메모리	36GB
그래픽	14코어 GPU, 16코어 Neural Engine
사용언어	Python
라이브러리	Transformers, Python-3.12.2 Pytorch-2.2.2

평가된 챗봇 모델은 다국어 BERT 모델과 KoBERT를 포함하며, 각 모델은 배치 크기 16으로 설정하여 30 에포크 동안 Adam 옵티마이저를 사용해 학습되었다. 또한, 학습 과정에서 과적합을 방지하기 위해 드롭아웃(dropout) 기법을 적용하였으며, 하이퍼파라미터 튜닝을 통해 최적의 성능을 달성하고자 하였다. 실험 결과, 다국어 BERT 모델은 45%의 정확도를 기록한 반면, KoBERT는 60%의 정확도를 보이며 한국어 질의응답에서 15% 이상의 성능 차이를 나타냈다.

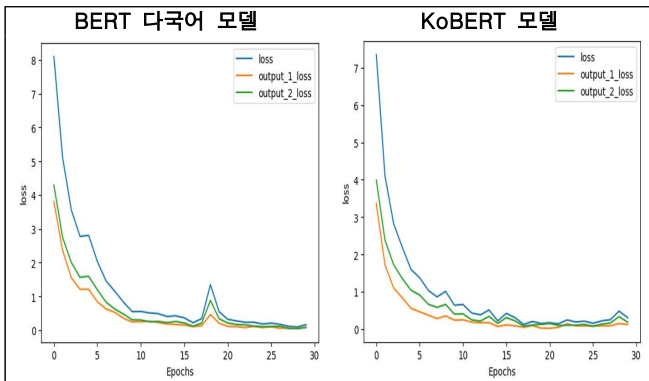


그림 2 BERT 모델별 학습 곡선

각 모델의 학습 곡선을 비교한 결과, KoBERT는 더 빠른 수렴 속도를 보였으며, 정확도와 F1 스코어에서도 우수한 성능을 나타냈다. 이러한 결과는 한국어와 같이 복잡한 언어에서는 도메인 특화 모델의 중요성을 시사한다.

질문 : 선택 결재하기 두 개 이상의 문서를 선택하여 결재할 경우, 선택한 문서를 순서대로 창을 띄워 원하는 문 정답 : 참조문서는 처리구분에 어떻게 [UNK]? 정답 : [UNK] 참조 [UNK] 로 표시 예측 : [UNK] 참조 [UNK] 로 표시되고 [전자결재] 창에 참조확인 버튼이 나타납니다.
질문 : 메일 일반설정하기 일반 설정 메뉴에서는 메일 수신자의 별칭, 메일의 보관 기간, 메일 읽기 화면, 보낸 정답 : 메일의 보관 기간이 경과된 메일은 어떻게 [UNK]? 정답 : 자동으로 [UNK]. 예측 : 자동으로 [UNK].

그림 3 다국어 BERT 모델 검증 결과

질문 : 결재 처리 방법 결재 기안문서의 안전에 대한 의사 결정 권한이 있는 결재자가 그 의사를 결정하는 행위임 정답 : 검토는 어떤 행위입니까 정답 : 중간결재자의 결재 행위 예측 : 중간결재자의 결재 행위
질문 : 투표 참여기간 변경하기 진행 중인 투표의 참여 기간을 변경할 수 있습니다 1 설문 메뉴에서 진행중인 투표 정답 : 투표 종료는 누가 할 수 있습니까 정답 : 투표 작성자 또는 투표 삭제 권한이 있는 사용자 예측 : 투표 작성자 또는 투표 삭제 권한이 있는 사용자

그림 4 KoBERT 모델 검증 결과

다국어 BERT 모델은 종종 [UNK] (unknown token)을 출력하는데, 이는 사전에 없는 단어나 처리할 수 없는 특수 문자를 만났을 때 해당 단어를 대체하는 방식으로 나타난다.[3] 이로 인해 다국어 BERT는 한국어 처리에 한계를 보였다. 반면, KoBERT는 한국어 문법과 도메인 특화 어휘를 더 잘 이해하여 고객 질의응답에서 높은 정확도를 기록하였다.

III. 결론

본 연구에서는 다국어 BERT 모델이 종종 [UNK] (unknown token)을 출력하여 한국어 처리에 한계가 있음을 보였다. 특히, KoBERT를 적용한 결과 정확도가 15% 향상되며, 한국어 질의응답에서 KoBERT의 우수성이 두드러졌다. 이는 KoBERT와 같은 도메인 특화 언어 모델이 한국어 고객 서비스의 품질 향상에 중요한 역할을 한다는 점을 시사한다.

향후 연구에서는 데이터셋 구축 과정에서 매뉴얼에 포함된 복잡한 기술 용어를 일반 사용자가 이해하기 쉬운 표현으로 변환하여 정확도를 더욱 강화할 계획이다. 또한, 데이터의 정확성을 확보하기 위해 각 데이터셋을 수동으로 검증하여 데이터 품질을 유지하고, 모델의 오차 학습을 방지할 것이다. 더 나아가, Ko-Sentence-BERT를 활용하여 지문을 학습시킴으로써 질문만 입력해도 정확한 답변을 제공할 수 있도록 연구할 예정이다.[6] 실제 고객 서비스 환경에서 파일럿 테스트를 수행하여 실용성을 평가하고, 이를 통해 더욱 자연스럽고 인간적인 대화를 실현하며 사용자 피드백을 반영하여 챗봇의 정확성과 유용성을 지속적으로 개선할 것이다.

ACKNOWLEDGMENT

이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 지역지능화혁신인재양성 사업임 (IITP-2024-2020-0-01462).

참 고 문 헌

- [1] Adamopoulou, E., & Moussiades, L. (2020). An Overview of Chatbot Technology. In Artificial Intelligence Applications and Innovations. Springer, Cham.
- [2] Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv preprint arXiv:1810.04805.
- [3] Sangah Lee, Hansol Jang, Yunmee Baik, Suzi Park, Hyopil Shin. (2020). KR-BERT: A Small-Scale Korean-Specific Language Model. arXiv preprint arXiv:2008.03979.
- [4] Lim, S., Kim, M., & Lee, J. (2019). KorquAD1.0: Korean QA dataset for Machine Reading Comprehension. arXiv.org. <https://arxiv.org/abs/1909.07005>
- [5] SK Telecom. (2019). KoBERT: BERT for Korean. GitHub repository. <https://github.com/SKTBrian/KoBERT>
- [6] SK Telecom. (2019). Ko-Sentence-BERT-SKTBERT. GitHub repository. <https://github.com/BM-K/KoSentenceBERT-SKT>