

GPT-3와 KoBERT를 활용한 감정 분석 기반

AI 챗봇 시스템

김준현^o, 문미경^{*}

^o동서대학교 소프트웨어학과,

^{*}동서대학교 소프트웨어학과

e-mail: 20181495@g.dongseo.ac.kr^o, mkmoon@dongseo.ac.kr^{*}

Emotion Analysis-Based AI Chatbot System Using GPT-3 and KoBERT

Junhyeon Kim^o, Mikyeong Moon^{*}

^oDept. of Software, Dongseo University,

^{*}Dept. of Software, Dongseo University

● 요약 ●

최근 챗봇 시스템은 급격한 발전과 함께 사용자와 자연스러운 대화를 할 수 있는 인공지능 기술의 필요성이 대두되고 있다. 기존의 챗봇 시스템은 대화 상황을 충분히 이해하지 못하거나, 학습된 데이터를 벗어나는 문장에 대한 일관성 있는 응답을 제공하지 못하는 한계가 있다. 본 논문에서는 GPT-3와 KoBERT를 활용하여 사용자의 감정 상태를 파악하고 해당 감정을 고려한 일관성 있는 대화를 제공하는 감정 분석 기반 챗봇 시스템을 제안한다. 이를 바탕으로 긍정적인 대화를 이어 나가는데 초점을 두어 자연스러운 대화가 가능할 것으로 기대된다.

키워드: 챗봇(Chatbot), KoBERT(Korean BERT), GPT(Generative Pre-trained Transformer), 감정 분석(Emotion Analysis), BERT(Bidirectional Encoder Representations from Transformers)

I. Introduction

최근 챗봇 시스템은 급격한 발전을 이루어 왔으며, 이와 함께 인공지능(Artificial Intelligence) 기술을 활용한 대화 시스템에 관한 관심이 크게 증가하고 있다. 인공지능 기술의 발전과 자연어 처리 기술의 진보로 인해 챗봇 시스템은 이전보다 더욱 놀라운 성능을 보여주고 있다 [1].

그러나 기존의 챗봇 시스템은 여전히 몇 가지 한계점을 가지고 있다. GPT-3는 다양한 문맥과 문장을 이해하고 생성할 수 있는 강력한 자연어 처리 모델이지만, 그 자체로는 감정 분석 기능이 부족하다. GPT를 활용한 기존의 챗봇 시스템은 대화 상황을 충분히 이해하지 못하거나, 사용자의 질문이나 요청에 대해 일관성 있는 응답을 제공하지 못하는 경우가 많다.

본 논문에서는 이러한 한계를 극복하기 위해 GPT-3에 KoBERT 감정 분석 기능을 결합하여 적절한 응답을 생성하는 감정 분석 기반 AI 챗봇 시스템에 관한 연구 내용을 기술한다.

II. Preliminaries

연구[2]는 감정 어휘 사전을 활용하여 감정 말뭉치를 구축하고, KcBERT를 학습시켜 감정 분석의 성능을 평가했다. 해당 연구는 감정 어휘 사전의 다양성이 제한되고 어휘의 양이 제한적으로 구축되는 한계가 있다. 본 연구는 이러한 한계점을 극복하기 위해 GPT-3 모델을 활용하여 더 효과적인 자연어 처리 및 생성 작업을 수행하고자 한다.

연구[3]에서는 부정적인 감정을 완화하기 위해 BERGPT-chatbot을 제안했다. 그러나 현재 공개된 감정 데이터셋의 한계로 인해 일상 대화 데이터셋을 활용한 모델에서는 모든 질문에 대한 응답에 한계가 있었다. 이에 관해 본 연구는 KoBERT를 활용하여 한국어 감정 데이터셋을 구축하고, GPT-3와의 조합으로 다양한 질문에 대한 응답을 개선하고자 한다.

본 연구에서는 사전 연구의 한계점을 극복하여 GPT-3와 KoBERT를 결합한 감정 분석 기반 AI 챗봇 시스템을 제안하여 보다 자연스러운 대화와 감정에 맞는 응답을 구현하는 것을 목표로 한다.

III. The Proposed Scheme

1. System Architecture

본 논문에서 제안하는 시스템은 다음과 같다. 사용자가 채팅을 입력하면, 해당 문장이 사전 학습된 KoBERT 모델에 삽입되어 문장의 감정을 추출한다. 추출된 감정 정보와 함께 입력 문장은 데이터베이스에 저장되며, 이는 이전 대화 기록을 읽어오는 과정에서 활용된다. 전처리된 데이터는 GPT-3 모델에 입력되어 Transformer Decoder Layer를 통해 예측 연산을 수행하고, 이를 통해 응답이 생성된다. 생성된 응답은 데이터베이스에 저장되거나 필요에 따라 출력될 수 있다.

이러한 시스템 구조를 통해 자연스러운 대화를 구현하고, 생성된 응답은 이후 대화에 활용될 수 있다.

2. Sentiment Analysis

본 연구에서는 KoBERT 모델의 Fine-Tuning을 위해 AI_Hub의 세브란스 상담 데이터를 활용하였다. 해당 데이터셋은 실제 상담 대화에서 추출된 문장으로 총 4,738개의 문장과 359개의 감정 카테고리 리로 이루어져 있다.

Table 1은 KoBERT 모델을 사용하여 입력 문장에 대한 감정 분석의 정확도를 나타내는 지표이다.

Table 1. Sentiment Analysis Accuracy

Sentence	Sentiment	Accuracy
처음에는 내 미래가 너무 막연하게 걱정돼서 그랬어	걱정/미래	0.994854
오늘 친구랑 점심 같이 먹기로 했어	배경/친구	0.790377
비가 오지는 않을까 걱정 돼	걱정	0.715363

Table 2는 KoBERT를 사용하여 감정 분석을 적용한 GPT 응답과 적용하지 않은 GPT 응답을 비교한 결과를 나타낸다. 첫 번째 행은 입력 문장을 나타내고, 두 번째 행은 감정 분석을 적용하지 않은 GPT의 응답을 나타내며, 세 번째 행은 감정 분석을 적용한 GPT의 응답을 나타낸다.

Table 2. Comparison of Responses GPT with and without Sentiment Analysis

Input Text	예전에 비해서 인내심이 너무 짧아진 거 같아.		
Sentiment Analysis	걱정	Accuracy	0.81249
GPT	그럴 수도 있습니다. 오히려 이런 과정을 거친 사람은 "내가 원해서 하는 일이라서 참 좋다." 이런 말을 더 많이 해줍니다.		
GPT with Sentiment	그렇게 느껴지시는군요. 이전에는 어떤 일이 있어도 참을 수 있었는데, 요즘은 그렇지 않으시다는 건가요? 저는 당신이 감정을 잘 조절할 수 있도록 조언을 해드리겠습니다.		

감정 분석을 적용한 GPT의 응답은 입력 문장에 대한 감정을 파악하고, 이를 바탕으로 보다 적절하고 감정에 맞는 응답을 생성하는 것을 확인할 수 있다.

IV. Conclusions

본 논문에서는 KoBERT와 GPT-3를 활용하여 입력 문장의 감정을 추출하고 이를 기반으로 자연스럽게 의미 있는 응답을 생성할 수 있는 감정 분석 기반 AI 챗봇 시스템의 개발 내용에 관하여 기술하였다.

본 시스템은 사용자들에게 더욱 풍부하고 맞춤형 대화 경험을 제공하며, 챗봇 기반 서비스 분야에서의 차별화된 응용 가능성을 제시한다.

ACKNOWLEDGEMENT

본 연구는 2023년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음 (2019-0-01817)

REFERENCES

- [1] Park Junho, yoon Kyungil, Min Sungtae. "Technical trends in chatbot systems based on artificial intelligence." Journal of Information Processing 26.2: 39-46. (2019)
- [2] Yeonji Jang, Jiseon Choi, Hansaem Kim. "KcBert-based Movie Review Corpus Emotion Analysis Using Emotion Vocabulary Dictionary." Journal of KIISE,49(8),608-616. (2022)
- [3] Yun-Gyeong Song, Kyung-Min Jung, Hyun Lee. "A BERGPT-chatbot for mitigating negative emotions." Journal of the Korea Society of Computer and Information 26.12 :53-59.(2021)