

감정분석 기반 심리상담 AI 챗봇 시스템에 대한 연구*

안세훈** · 정옥란***

A Study on the Psychological Counseling AI Chatbot System based on Sentiment Analysis*

Se Hun An** · Ok Ran Jeong***

■ Abstract ■

As artificial intelligence is actively studied, chatbot systems are being applied to various fields. In particular, many chatbot systems for psychological counseling have been studied that can comfort modern people. However, while most psychological counseling chatbots are studied as rule-based and deep learning-based chatbots, there are large limitations for each chatbot. To overcome the limitations of psychological counseling using such chatbots, we propose a novel psychological counseling AI chatbot system. The proposed system consists of a GPT-2 model that generates output sentence for Korean input sentences and an Electra model that serves as sentiment analysis and anxiety cause classification, which can be provided with psychological tests and collective intelligence functions. At the same time as deep learning-based chatbots and conversations take place, sentiment analysis of input sentences simultaneously recognizes user's emotions and presents psychological tests and collective intelligence solutions to solve the limitations of psychological counseling that can only be done with chatbots. Since the role of sentiment analysis and anxiety cause classification, which are the links of each function, is important for the progression of the proposed system, we experiment the performance of those parts. We verify the novelty and accuracy of the proposed system. It also shows that the AI chatbot system can perform counseling excellently.

Keyword : Electra, GPT-2, Sentiment Analysis, AI Chatbot, Psychological Counseling

Submitted : January 31, 2021

1st Revision : March 17, 2021

Accepted : June 9, 2021

* 이 논문은 한국 과학기술정보통신부의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(NRF2019 R1A2C1008412).

** 가천대학교 AI·소프트웨어학부 학사

*** 가천대학교 AI·소프트웨어학부 부교수, 교신저자

1. 서 론

2010년대 이후에는 사회적 불안 장애와 같은 정신 질환이 있음에도 대부분의 사람들이 시간적 여유 부족, 재정적인 어려움, 사회적 시선의 두려움 등과 같은 이유로 병원에 방문하여 심리 상담을 받지 않는 경우가 많다(Erica, 2013).

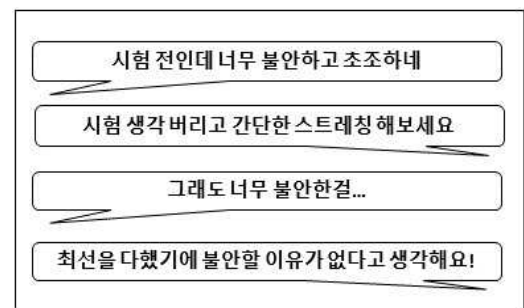
이런 이유로 대면 심리 상담이 아닌 컴퓨터를 통해서 상담하는 방식의 효율성이 입증되면서 비대면으로 심리 상담을 하는 앱, 프로그램 등이 많이 제공된다(Gary et al., 2008). 하지만 이런 프로그램들은 사람들에게 상담과 조언을 통해 정서적 안정을 주려는 목적을 가지지만 일시적인 대화를 통해서 정서적 안정을 취하기는 힘든 일이다. 그렇기에 일시적이지 않고 지속적으로 감정을 보살펴 줄 수 있는 심리 상담 서비스의 필요성이 커지고 있다. 따라서 상담가가 일시적으로 상담을 하는 것이 아닌 인공지능이 지속적으로 감정을 보살펴 기존의 심리 상담을 대체할 수 있을 것이라는 점이 현실로 다가오고 있다(이아라 외, 2019).

이에 규칙 기반 챗봇과 딥러닝 기반 챗봇을 사용하여 [그림 1], [그림 2]와 같은 대화를 목표로 심리 상담 연구가 제시되어 왔다. 하지만 규칙 기반 챗봇은 [그림 3]의 사례 1처럼 기존에 입력되지 않은 내용에 대해서 대답이 불가능하다는 한계점이 있다. 그리고 딥러닝 기반 챗봇은 아직 한국어를 기반으로 하는 딥러닝 기반 챗봇 개발이 많이 이루어지지 않았을뿐더러 [그림 3]의 사례 2 대답 2처럼 사용자의 고민에 대해 충분한 해결책을 제시하지 못할 가능성이 있다(이아라 외, 2019).

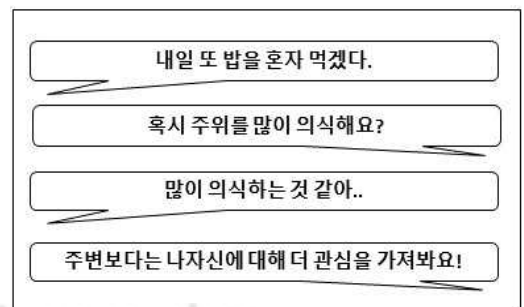
이 두 가지 종류와 같은 기존 챗봇의 한계점을 해결할 수 있는 새로운 챗봇 시스템의 필요성이 대두되고 있다. 딥러닝 기반 챗봇의 한계를 극복하기 위해서는 감정 분석에 따른 심리 검사와 불안 원인 분류 기반 집단 지성의 기능을 제시한다. 이러한 기능들로 설계된 모델에서는([그림 1], [그림 2], [그림 3]의 사례 3)처럼 자유로운 주제를 바탕으로 상담할 수 있으며, 감정 분석과 불안 원인 분류에 따

른 심리 검사와 집단 지성을 통해 공감할 수 있는 해결책을 제시할 수 있다. 본 연구에서는 한국어 기반의 GPT-2(Radford et al, 2019)를 통해서 자유로운 주제를 바탕으로 상담할 수 있으며, 한국어 기반의 Electra 모델을 활용하여 감정 분석에 따른 심리 검사와 불안 원인 분류를 통한 집단 지성 기능을 제공하는 심리 상담 AI 챗봇 시스템을 제안한다. 시스템은 사용자의 인터페이스 역할을 하는 한국어 기반 GPT-2 심리 상담 챗봇, STAI 불안 척도(Charles et al., 1984)를 기반으로 하는 불안 정도 검사, SSCT(Sacks et al., 1950)와 한국어 기반 Electra 모델을 기반으로 하는 불안 원인 검사 및 집단 지성 기능으로 구성된다.

또한, 이 논문에서는 챗봇 상담, 심리 검사 그리고 집단 지성의 연결고리인 감정 분석 및 불안 원인 분류가 가장 중요하기에 한국어 기반 Electra 모델(Kevin et al., 2020)을 바탕으로 한 감정 분석과 불안 원인 분류에 대한 성능을 실험한다.



[그림 1] 심리 상담 예시1



[그림 2] 심리 상담 예시2

입력 문장

회사상사때문에 너무 스트레스받고 매일 매일이 불안해

사례 1. 규칙 기반 챗봇

규칙이 있을 때

상사와진심 어린 대화를 나눠보는 건 어때요?

규칙이 없을 때

무슨 말인지 모르겠습니다.

사례 2. 딥러닝 기반 챗봇

대답 1

상사와이야기를 나눠보는 건 어때?

대답 2

요즘 일이 너무 많아.

사례 3. AI 챗봇 시스템

상사와이야기를 나눠보는 건 어때?

불안 점수
63점, 매우 높은 수준

불안 원인
관계 - 권위자
나는 누군가가 나에게 다가올 때 너무 두려워. 너무
떨려서 평범한 생각을 할 수가 없어.
...

집단 지성
나는 내 상사와 함께 식사를 하면서 내가 그동안 힘든
점에 대해서 이야기를 나눴어. 상사는 나에게
진심으로 그 동안 일에 대해 부담을 준 점에서 사과를
하고 충고를 해줬어.
:
:

[그림 3] 동기 부여 예시

2. 관련 연구

2.1 챗봇

챗봇이란 텍스트 기반 대화 인터페이스를 통해 사용자와 상호작용하는 소프트웨어 에이전트를 의미한다. 챗봇은 최근에 사용되는 분야가 매우 넓은데 대표적으로 헬스케어, 호텔 예약, 기업 등(Sakthivel et al., 2018)에서 사용되고 있으며 최근에는 심리상담 분야에서도 진단, 탐색, 평가 과정을 바탕으로 한 시스템이 연구되었다. 실제 내담자들의 경험을 바탕으로 하나의 프로세스를 조직하여 문제 해결 진단을 내리는 시스템으로 긍정적인 결론에 도달한 바 있다(이아라 외, 2019). 하지만, 수많은 내담자의 경험을 바탕으로 시스템이 구성되었다 해도 사람마다 심리상담을 받는 이유가 다를 수 있기 때문에 모든 사람에게 해결책을 제공하는 것은 불가능할 수 있다.

2.2 임베딩 기반 모델

2.2.1 GPT-2

GPT-2(Radford et al, 2019)는 Open AI에서 Webtext의 40GB의 코퍼스와 크롤링한 데이터를 합쳐서 훈련시킨 언어 모델이다. BERT(Devlin et al., 2019) 모델과는 다르게 Encoder Stack은 사용하지 않고 Transformer의 Decoder Stack을 통해 Masked Self-Attention을 사용하여 이전의 출력값이 다음 입력값으로 변하는 모델로 다음 출력값을 예측하는 능력이 뛰어나 문장 생성을 하는데 주로 사용된다. GPT-2 모델에 문장을 입력하면 Byte Pair Encoding(BPE)를 통해 문자들을 토큰으로 나누고 이 토큰들을 임베딩 행렬로 변환한 뒤 Decoder Stack에 들어간다. 그리고 각 입력 벡터는 Decoder Stack 안의 각각의 Decoder cell에서 Self-attention 과정을 통해서 문맥을 이해하는 과정을 거치고 임베딩 벡터를 통해 확률 값을 구한 뒤 확률 값이 가장 높은 문장을 출력하는 형태로 진행된다. 하지만 GPT-2 모델은 영어를 기반으로 개발이 되었기 때문에 한국어 기반 시스템을 제안한 이 논문에서는 아파존 웹 서비스

(AWS)와 SKT가 협력해 개발한 KoGPT-2를 사용한다. KoGPT-2 모델은 117M(12 layers)로 구성되어 있으며 20GB의 문장을 사용하여 학습된다.

2.2.2 ELECTRA

ELECTRA(Kevin et al., 2020)는 BERT(Devlin et al., 2019)의 장점을 유지하고 효율적으로 학습할 수 있는 언어 사전 훈련 방식을 제안한 모델이다. 동일한 자원을 가지고 Generator와 Discriminator를 통해서 토큰 분류를 정확히 할 수 있는 인코더를 효율적으로 학습한다. Generator는 MASK로 대체된 입력값에 대해서 원래 단어와는 다른 단어들로 대체하고 Discriminator는 Generator의 출력값을 입력으로 받아 실제 토큰인지 아닌지를 판별하는 방법으로 학습한다. 이 논문에서는 한국어를 기반으로 하는 시스템을 제안했기 때문에 34GB의 한국어 말뭉치를 통해 학습한 한국어 기반 ELECTRA 모델인 KoELECTRA-small을 사용한다.

2.3 심리 상담을 위한 메커니즘

2.3.1 STAI 불안 척도

STAI는 State-Trait Anxiety Inventory의 약자로 상태 불안 척도를 의미한다(Charles et al., 1984). 1970년 Spielberg에 의해서 개발되었으며 현재 가지고 있는 상태 불안을 20개의 문항을 통해 측정한다. 정상인을 대상으로 한 불안의 정도를 측정하는 도구로 개발되었으며, 최근에는 임상 집단의 불안을 측정하는데 유용성이 밝혀지면서 많이 사용되고 있다.

STAI 불안 척도는 자기 자신이 보고하는 방식으로 진행이 되며 각 문항에 대해서 얼마나 현재 느끼고 있는지를 4점 척도로 평가한다. 4점 척도는 1점은 전혀 그렇지 않다, 2점은 조금 그렇다, 3점은 보통으로 그렇다, 4점은 대단히 그렇다로 구성된다. 이와 같은 항목이 아래와 같이 총 20개 문항이 있으며 점수를 합산하여 불안의 정도를 판단한다.

<표 1>의 항목에서 1, 2, 5, 8, 10, 11, 15, 16, 19, 20번 항목은 역방향으로 채점이 되어야 하며 점수의

<표 1> STAI 불안 척도 20개 질문

문항	내용
1	나는 마음이 차분하다.
2	나는 마음이 든든하다.
3	나는 긴장되어 있다.
4	나는 후회스럽고 서운하다.
5	나는 마음이 편하다.
6	나는 당황해서 어찌할 바를 모르겠다.
7	나는 앞으로 불행이 있을까 걱정하고 있다.
8	나는 마음이 놓인다.
9	나는 불안하다.
10	나는 편안하게 느낀다.
11	나는 자신감이 있다.
12	나는 짜증스럽다.
13	나는 마음이 조마조마하다.
14	나는 극도로 긴장되어 있다.
15	내 마음은 긴장이 풀려 푸근하다.
16	나는 만족스럽다.
17	나는 걱정하고 있다.
18	나는 흥분되어 어쩔 줄 모른다.
19	나는 즐겁다.
20	나는 기분이 좋다.

범위는 20~80점 사이가 될 수 있다. 51점 이하는 정상인 상태, 52점~56점은 불안 수준이 약간 높은 상태, 57점~61점은 불안 수준이 상당히 높은 상태, 62점 이상은 불안 수준이 매우 높은 상태를 나타낸다.

2.3.2 문장 완성형 검사

The Sacks Sentence Completion Test(SSCT)는 Joseph M.Sacks에 의해서 개발되었으며 문장 완성형 검사의 일종으로 가족, 성, 자기 개념, 대인 관계라는 4가지 영역에 대한 미완성 문장을 피검자가 자신의 생각을 바탕으로 완성하도록 하는 검사이다(Sacks et al., 1950). 자신 개념과 주변 사람들과의 관계를 파악하기 위해서 사용되며 문장 완성을 통해서 다양한 감정들을 표면으로 밝힐 수 있으며 구체적인 문제에 대해서 살펴볼 수 있다. SSCT는 총 50개의 문항으로 현재 많이 사용계가 모호하여 전에는 하지 않았던 야근을 하게 되고 있으며 4가지 영역으로 구분되는 문항은 다음과 같다.

<표 2> SSCT(문장 완성형 검사) 질문

문항	내용
아버지	2. 내 생각에 가끔 아버지는 19. 대개 아버지들이란 29. 내가 바라기에 아버지는 50. 아버지와 나는
어머니	13. 나의 어머니는 26. 어머니와 나는 39. 대개 어머니들이란 49. 나는 어머니를 좋아했지만
가족	12. 다른 가정과 비교해서 나의 가족은 27. 우리 가족이 나에게 대해서 42. 내가 아는 대부분의 집안은 57. 내가 어렸을 때 나의 가족은
여성	9. 내가 바라는 여인상은 25. 내 생각에 여자들이란
남성	8. 남자에 대해서 무엇보다 좋지 않게 생각하는 것은 20. 내 생각에 남자들이란 36. 완전한 남성상은
결혼 생활	10. 남녀가 같이 있는 것을 볼 때 23. 결혼 생활에 대한 나의 생각은
성적 관계	37. 내가 성교를 했다면 47. 나의 성생활은
친구	6. 내 생각에 참다운 친구는 22. 내가 싫어하는 사람은 32. 내가 제일 좋아하는 사람은 44. 내가 없을 때 친구들은
권위자	3. 우리 위사람들은 31. 위사람이 오는 것을 보면 나는
두려움	5. 어리석게도 내가 두려워하는 것은 21. 다른 친구들이 모르는 나만의 두려움은 40. 내가 잊고 싶은 두려움은 43. 때때로 두려운 생각이 나를 휩쓸 때
죄책감	14. 무슨 일을 해서라도 잊고 싶은 것은 17. 어렸을 때 잘못했다고 느끼는 것은 27. 내가 저지른 가장 큰 잘못은 46. 무엇보다도 좋지 않게 여기는 것은
자신의 능력	1. 나에게 이상한 일이 생겼을 때 15. 내가 믿고 있는 내 능력은 34. 나의 가장 큰 결점은 38. 행운이 나를 외면했을 때
과거	7. 내가 어렸을 때는 33. 내가 다시 젊어진다면 45. 생생한 어린 시절의 기억은
미래	4. 나의 장래는 11. 내가 늘 원하기는 16. 내가 정말 행복할 수 있으려면 18. 내가 보는 나의 앞날은 28. 언젠가 나는
목표	30. 나의 야망은 41. 나의 평생 가장 하고 싶은 일은 42. 내가 늙으면

<표 2>의 각 항목을 통해서 문항에 해당하는 개념에 대해 어떻게 생각하는지 해석하고 이를 요약한 후에 심리상담을 할 때 핵심으로 사용한다.

2.3.3 집단 지성

집단 지성이란 ‘어디에나 분포하고, 지속적으로 가치가 부여되며, 실시간으로 조정되고, 역량의 실제적 동원에 이르는 지성’을 의미한다. 개인 스스로 문제에 대한 해답을 찾는 것이 아닌 다수의 사람들과 정보 교류 및 공유를 해서 성찰을 통해 새롭게 해답을 창출하는 효과를 지닌다.

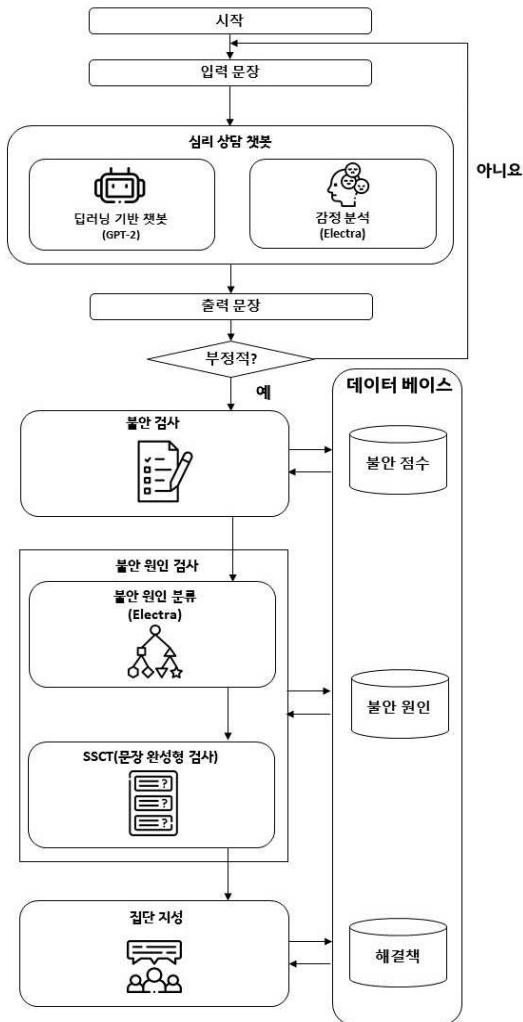
‘대중의 지혜’에 따르면 집단 지성의 형성 조건은 다양성, 독립성, 분산화와 통합이 이루어져야 한다고 한다. 첫 번째로 다양성은 집단이 가질 수 있는 해답의 범위를 확장함으로써 창의적인 방법으로 문제를 해석할 수 있게 하는 장점이 있다. 두 번째로 현명한 사고를 위한 개인의 독립성이 확보되어야 한다. 개인의 판단으로 인해서 집단을 파괴하지 않아야 하며 편향하지 않게 해야 한다. 세 번째로 집단 내에서 개인들이 가진 전문적인 분야에 대해서 지식을 얻음과 동시에 이런 지식을 집단적 형태로 종합할 수 있어야 한다. 이를 통해 얻어진 집단적 해법이 다른 해법보다 탁월한 해결책을 제시 해준다(양미경, 2010).

본 논문에서는 SSCT를 통해 얻은 개인의 문제와 그에 대한 해결책을 가지는 커뮤니티를 구현해서 집단 지성의 3가지 조건을 충족할 수 있는 환경을 조성해 개인의 문제를 공유하는 집단에서 해결할 수 있는 시스템을 설계 및 구현한다.

3. AI 챗봇 시스템

본 논문에서는 대면 심리 상담을 대체할 수 있는 심리 상담 시스템을 제안한다. 시스템은 크게 심리 상담 챗봇, 불안 정도 검사, 불안 원인 검사, 집단 지성 커뮤니티 순서로 이루어진다(그림 4 참조). 상호 지향성 모델의 객관적 일치도(Agreement)는 ‘대상 심리 상담 챗봇은 딥러닝 기반 챗봇을 통해 심리

상담을 할 수 있고 지속적인 감정 분석을 통해서 사용자의 감정을 파악한다. 만약 감정 분석을 통해 사용자의 감정이 부정적이라고 판단되면 이를 해결할 수 있도록 도와주는 형식으로 진행된다.



[그림 4] AI 챗봇 시스템 구조

3.1 심리 상담 챗봇

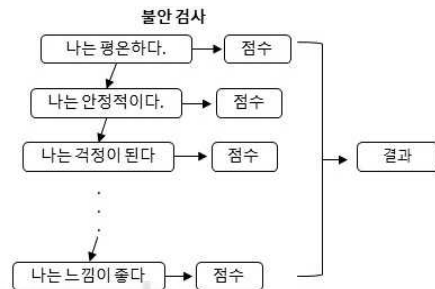
심리 상담 챗봇은 한국어 기반의 딥러닝 챗봇인 GPT-2를 사용한다. AI Hub에서 제공하는 2가지 데이터셋인 “웰니스 대화 스크립트 데이터셋(AI Hub, 2019)”과 “감성대화 말뭉치(AI Hub, 2019)”와 songys

이 제공하는 대화 데이터셋(Github, 2020)을 사용한다. “웰니스 심리 상담 데이터셋”은 신촌 세브란스 병원에서 정신건강 상담 기록을 가공하여 만들었으며 ‘감성 대화 말뭉치’는 일반인 1,500명을 상대로 인터뷰를 진행하고 우울증 환자를 대상으로 하는 WOZ 대화를 수집함으로써 데이터셋을 가공했다. 또한, songys이 제공하는 대화 데이터셋은 다음 카페 “사랑보다 아름다운 실연”에서 자주 나오는 이야기를 참고하여 제작했다. 3가지 데이터셋을 전처리해서 대화 데이터셋을 총 약 62,000개로 만들었으며 epoch을 10번으로 해서 챗봇 학습을 진행한다.

[그림 4]와 같이 GPT-2 모델은 사용자의 입력 문장에 대해서 문장을 출력함으로써 사용자와 상담을 해준다. 이 GPT-2 모델이 출력 문장을 예측할 때 동시에 Electra 모델을 통해 감정 분석도 이루어진다. 감정 분석을 통해서 긍정인지 부정인지를 판단하는 과정이 일어난다. 감정 분석 결과가 긍정이라면 아무 일도 일어나지 않고 일반적인 챗봇과 같이 다시 사용자가 문장을 입력할 수 있다. 반면에 입력 문장에 대해서 부정 감정이 인식되면 불안 정도 검사, 불안 원인 검사, 집단 지성 커뮤니티와 같은 차례로 이루어지는 형식으로 진행된다.

3.2 불안 정도 검사

불안 정도 검사는 Google에서 제공하는 챗봇 플랫폼인 Dialogflow를 사용한다. STAI 불안 척도의 원리를 바탕으로 하나의 시나리오로 작성하여 챗봇으로 구현한다.

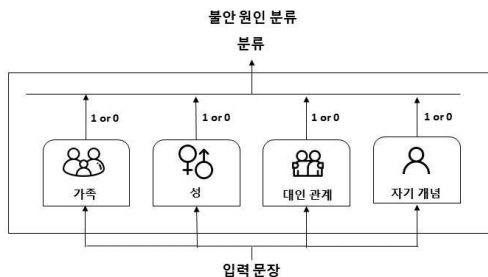


[그림 5] 불안 검사

[그림 5]와 같이 20개의 질문에 대해 현재 느끼고 있는 감정을 점수로 나타낸다. 기존에 STAI 불안 척도는 단순히 설문지를 작성하는 형태로 진행이 되지만 Dialogflow를 활용한 챗봇은 각 항목을 친근한 물음으로 변형하여 제공된다.

3.3 불안 원인 검사

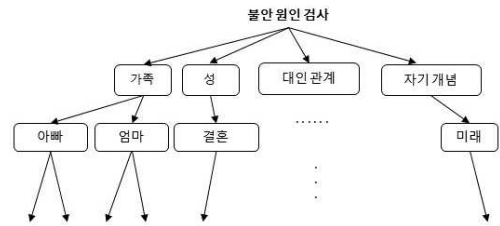
불안 원인 검사는 Electra 모델(Kevin et al., 2020)은 BERT(Devlin et al., 2019)과 Google에서 제공하는 챗봇 플랫폼인 Dialogflow(Dialogflow, 2021)을 사용한다. Electra 모델을 통해서 불안 원인 분류를 한다. 기존 SSCT 검사 원리 기반으로 4가지 분류(가족, 성, 대인 관계, 자기 개념)에 대해 알아낸다. 사용자 입력 문장에 대해 어떤 항목에 속하는지를 알아낼 수 있는데 4가지 항목 중 하나만 해당하는 것이 아닌 다수의 항목에 포함될 수 있기 때문에 [그림 6]과 같은 형식으로 분류한다.



[그림 6] 불안 원인 분류

총 4가지 Electra 모델로 구성이 되며 입력 문장이 각 항목에 해당한다고 판단이 되면 1을 출력하고 그렇지 않다면 0을 출력한다. 그리고 최종적으로 1이 나온 항목에 한해서 불안 원인 검사가 진행된다. 불안 원인 검사는 SSCT 검사 원리를 바탕으로 Dialogflow에 시나리오로 작성하여 챗봇으로 구현한다.

[그림 7]과 같이 Electra 모델을 통해 분류된 카테고리를 바탕으로 선택된 카테고리에서 세부적인 카테고리를 선택한다. 그리고 각 원인에 해당하는 항목에 대해 <표 2>와 같은 질문을 챗봇이 하게 되고 사용자가 입력한 대답을 저장한다.



[그림 7] 불안 원인 검사

3.4 집단 지성 커뮤니티

불안 정도 검사와 불안 원인 검사로부터 얻은 불안 점수와 SSCT로부터 얻은 완성된 문장을 바탕으로 집단 지성 커뮤니티를 제공한다. 또한 집단 지성의 형성 조건 3가지 모두 만족하도록 구성한다.

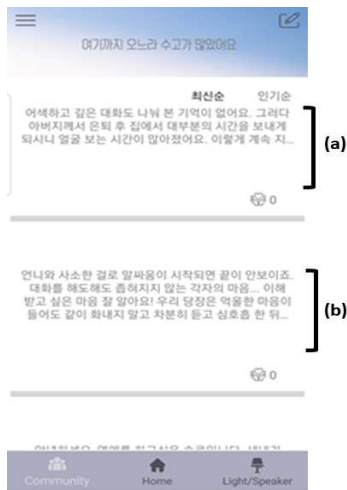


[그림 8] 집단 지성 카테고리

[그림 8]과 같이 4가지(가족, 성, 대인 관계, 자기 개념)로 분류된 사례와 해결책만 모아볼 수 있는 환경을 제시한다. 그림 8에서 원하는 항목을 누르게 되면 [그림 8]과 같이 집단의 다양한 사람들이 자신의 이야기를 서술하고 자신이 제시한 해결책을 볼 수 있도록 구성했다.

4가지 항목에서 많은 사람들이 자신의 이야기를 제한 없이 쓸 수 있도록 한다는 점에서 다양성을

만족하고, 타인과 이야기를 주고받는 환경이 아니라 서로의 이야기에 좋아요를 눌러주는 구조이기에 독립성을 만족한다. 마지막으로 분산화와 통합 역시 좋아요를 통해서 많은 사람들이 공감할 수 있는 이야기나 해결책에 대해서 볼 수 있게 구성되었기 때문에 전문적인 지식을 얻고 집단적으로 통합할 수 있는 환경을 제공한다.



[그림 9] 집단 지성 커뮤니티

[그림 9]의 (a)와 (b)는 집단 지성의 예로 들 수 있다. (a)는 가족 중에 아버지 집단 지성에 속하는 내용이다. 평소에 한 번도 제대로 된 대화를 해본 적 없는 자신의 아버지가 은퇴하시고 아버지와 같이 집에 있는 시간이 많아지면서 어색한 횃수가 많아지고 불안, 불편함 등의 감정을 겪은 사례이다. 하지만 마주칠 때마다 피하지 않고 말 한 번이라도 건네는 노력을 함으로써 현재는 좋은 관계가 되었다는 해결책을 전해준다. 또한 (b)는 가족 중에 연니 집단 지성에 속한 내용이다. 매일 싸우고 이로 인해서 많은 상처를 받는 경우가 생기는데 이럴 때마다 자신의 감정 표출을 자제하고 다른 사람들의 말을 잘 듣는 습관을 기르면 좋은 관계가 될 수 있다는 해결책을 전해준다. (a)와 (b)같은 사례가 집단 지성 커뮤니티에 많아지면 모든 경우의 문제에 대해서도 해결이 가능할 수 있다.

4. 실험

이 장에서는 감정 분석과 불안 원인 분류에 사용된 데이터셋, 실험 환경, 실험 결과에 대해서 설명한다.

4.1 실험 환경

3가지 모델에 대해서 구글 colab의 GPU 환경에서 실험한다. 한국어 기반 BERT 모델은 KoBERT를 사용한다. 이 모델은 SKTBrain에서 제공하며 PyTorch를 통해 구현되었다. 한국어 기반 Electra는 KoElectra를 사용한다. 이 모델 역시 PyTorch를 통해 구현되었다. LSTM(Sepp et al., 1997) 모델은 tensorflow와 keras를 통해 구현되었으며 many-to-one model을 통한 감정분석이 이루어진다. 또한, 4가지 항목에 대한 불안 원인 분류는 한국어 기반의 4개(가족, 성, 대인관계, 자기개념)의 Electra 모델을 통해서 이루어지고 각각의 Electra 모델에 대한 검증이 이루어진다.

4.2 데이터 셋

감정 분석을 위한 3가지 모델 학습을 위한 데이터셋은 AI Hub에서 제공하는 2가지 데이터셋 ‘웰니스 대화 스크립트 데이터셋’, ‘한국어 감정 정보가 포함된 단발성 대화 데이터셋(AI Hub, 2020)’과 songys이 제공하는 대화 데이터셋을 사용한다. ‘웰니스 대화 스크립트 데이터셋’과 songys이 제공하는 대화 데이터셋은 3.에서 이미 다루었고 ‘한국어 감정 정보가 포함된 단발성 대화 데이터셋’은 SNS 글 및 온라인 댓글에 대한 웹 크롤링을 하여 문장을 선정하고 7개의 감정(기쁨, 슬픔, 놀람, 분노, 공포, 혐오, 중립)으로 레이블링을 수행했다.X에 관한 3가지 데이터셋을 바탕으로 각 문장에서 구분된 감정을 바탕으로 긍정 문장 50,000개와 부정 문장 50,000개를 전처리하여 분류했으며 100,000개의 데이터셋에서 90%와 10%의 비율로 랜덤하게 분류하여 Training과 Test 데이터셋으로 나눈다.

다음으로 불안 원인 분류에 대해서는 ‘웰니스 대화 스크립트 데이터셋’을 통해서 이루어진다. 359개의 원인으로 분류된 총 5200개의 데이터를 4개(가족, 성, 대인관계, 자기개념)로 분류했다. 각각의 항목에 대해서 2600개씩 분류했다. 최종적으로 5200개의 데이터셋에서 90%와 10% 비율로 랜덤하게 분류하여 Training과 Test 데이터셋으로 나눈다.

4.3 결과

감정 분석에 대한 실험으로 3가지 모델에 대해서 epoch을 5번으로 동일하게 한 후 학습을 진행했다. 학습을 완료한 후에 테스트 데이터셋을 통해서 예측한 감정과 기존에 분류된 감정을 비교한다.

〈표 3〉 감정 분석 결과

	Precision	Recall	F1-score
BERT	0.81	0.82	0.82
ELECTRA	0.82	0.84	0.83
LSTM	0.38	0.78	0.51

첫 번째로 <표 3>은 각 모델에 대한 감정 분석 정확도 실험 결과이다. 실험의 정확도 측정을 위해서 Precision, Recall, F1 점수를 사용한다. 각 모델의 F1 점수는 약 0.82, 0.83, 0.51로 계산된다. Electra 모델의 F1 점수가 0.82로 제일 높았으며 이는 가장 감정 분석을 잘 한다는 것을 의미한다. 이 실험을 바탕으로 가장 좋은 성능을 가지는 Electra를 제안한 시스템에 적용한다.

두 번째로 불안 원인 분류에 대한 실험으로 한국어 기반의 4가지 Electra 모델에 대해 대해서 epoch을 10번으로 동일하게 한 후 학습을 진행했다. 학습을 완료한 후에 테스트 데이터셋을 통해서 예측한 불안 원인 분류와 기존에 분류된 감정을 비교한다.

〈표 4〉 감정 원인 분류 결과

	Precision	Recall	F1-score
가족	0.85	0.81	0.83
성	0.89	0.77	0.83
대인관계	0.78	0.84	0.80
자기개념	0.87	0.79	0.82

<표 4>는 한국어 기반 Electra 모델을 통한 불안 원인 분류 정확도 실험 결과이다. 실험의 정확도 측정을 위해서 Precision, Recall, F1 점수를 사용한다. 각 항목에 대해서 F1 점수가 약 0.80 이상으로 계산된다. 이는 타 실험과 비교했을 때 높은 성능을 보인다. 황상흠 외(2020)에서는 한국어 기반의 Bert 모델을 통해 국가과제 기술문서를 총 33개의 기술 분류하는 작업을 하였는데 이는 대상이 속한 항목을 예측하고 정확도를 판단하는 실험을 한다는 점에서 우리의 연구와 유사하기에 비교한다. 황상흠 외(2020)에서는 Precision, Recall, F1-score의 평균 값이 각각 약 0.54, 0.55, 0.54가 나온 반면 Electra를 사용한 분류의 Precision, Recall, F1-score의 평균 값은 각각 약 0.85, 0.80, 0.82로 확실히 높은 결과가 나왔다(<표 5> 참조).

〈표 5〉 비교 실험

	Precision	Recall	F1-score
ours	0.85	0.80	0.82
(황상흠 외, 2020)	0.54	0.55	0.54

실험 결과들을 바탕으로 GPT-2 모델에 문장을 입력할 때 Electra에도 역시 그 문장을 입력해 입력 문장마다 감정을 측정하고 부정적인 감정이 판단되면 이를 바탕으로 사용자에게 불안 정도 검사, 불안 원인 검사, 집단 지성 커뮤니티 과정을 제공할 것이며 불안 원인 검사 단계에서는 한국어 기반의 Electra를 통해 충분한 분류가 이루어져 사용자의 불안 문제 해결에 충분한 도움을 줄 것이다.

5. 결 론

본 논문에서는 규칙 기반 챗봇과 딥러닝 기반 챗봇의 한계점을 해결하고자 감정 분석을 활용한 불안 정도 검사, 불안 원인 검사, 집단 지성 커뮤니티를 추가하여 기존에 한계를 극복하고자 한다. 또한, 본 논문에서는 감정이 포함된 문장들에 대해 감정을 가장 잘 분류할 수 있는 모델을 검증하는

실험을 한다. BERT, Electra, LSTM과 같은 3가지 모델을 동일한 데이터셋으로 비교를 한 후에 가장 높은 정확도를 보이는 모델인 Electra를 선택하여 시스템에 적용한다.

본 논문에서 제안한 시스템은 심리적으로 약한 사람들에게 많은 도움이 될 수 있다. 항상 시간적, 공간적 제약에 의해 구애받지 않고 이야기할 수 있는 챗봇이 있으며 사용자가 힘들다고 스스로 판단할 때 심리 검사와 집단 지성 커뮤니티를 이용할 수 있다. 스스로 극복하는 것이 아닌 비슷한 경험을 가지는 사람들과 함께 해결한다는 점이 장점이라고 할 수 있다.

하지만 제안한 시스템은 한계점 역시 존재한다. 아직 한국어 임베딩 기반 모델은 영어 임베딩 기반 모델보다 좋은 성능을 보여주지 못하고 있다. 현재 챗봇과 감정 분석에 대해 정확도를 높이고자 많은 연구가 이루어지고 있고 성능이 좋은 모델을 한국어 기반으로 변형한다면 실제 심리 상담과 같은 느낌을 줄 수 있을 것이다.

또한, 특정 도메인에서만 심리 상담을 제공하는 것이 아닌 앱, 웹, 웨어러블 기기와 같은 도메인에 AI 챗봇 시스템을 제공한다면 부정적 감정을 지속적으로 해결하는데 도움을 줄 것이다.

참고문헌

- 양미경, “집단지성의 특성 및 기제와 교육적 시사점의 탐색”, *열린교육연구*, 제18권, 제4호, 2010, 1-30.
- 이아라, 김효창, 차민철, 지용구, “상담 이론 기반의 심리 상담 챗봇을 활용한 내담자 경험 연구”, *대한인간공학회지*, 제38권, 제3호, 2019, 161-175.
- 황상흠, 김도현, “한국어 기술문서 분석을 위한 BERT 기반의 분류모델”, *전자거래학회지*, 제24권, 제1호, 2020, 203-214.
- AI Hub, “Corpus dataset for AI-based emotional chatbot for depression prevention and mitigation”, 2020, Available at <https://aihub.or.kr/aidata/7978> (Accessed May 1. 2021).
- AI Hub, “Single chat dataset with Korean sentiment information”, 2019, Available at <https://aihub.or.kr/node/274> (Accessed May 02, 2021).
- AI Hub, “Wellness dialogue script datasets”, 2019, Available at <https://aihub.or.kr/node/267#> (Accessed May 1. 2021).
- Charles, D. and R. Peter, “Psychometric Properties of the STAI : A Reply to Ramanaiah, Franzen, and Schill”, *Journal of Personality Assessment*, Vol.48, No.1, 1984, pp. 95-97.
- Devlin, J., M.W. Chang, K. Lee, and K. Toutanova, “BERT : Pre-training of Deep Bidirectional Transformers for Language Understanding” in *Proc. of NAACL, Minneapolis, America*, 2019, 1-16.
- Dialogflow, “Dialogflow essential”, Available at <https://dialogflow.cloud.google.com> (Accessed May 5. 2021).
- Erica, K., “Treatment of Social Anxiety Disorder Using Online Virtual Environments in Second Life”, *Behavior Therapy*, Vol.44, No.1, 2013, 51-61.
- Gary, E. and A., Kerr, “Computer-Mediated Counseling”, *Computers in Human Services*, Vol. 15, No.4, 2008, 13-26.
- Clark, K., M.T. Luong, Q.V. Le, and C.D. Manning, “ELECTRA : Pre-training Text Encoders as Discriminators Rather Than Generators”, *ICLR 2020 Conference, Addis Ababa, Ethiopia*, 2020, 1-14.
- Radford, A., J. Wu, R. Child, D. Luan, and D. Amodei, and I. Sutskever, and “Language models are unsupervised multitask learners”, 2019, Available at <https://openai.com/blog/better-language-models> (Accessed February 14. 2019).

- Sacks, J.M. and S. Levy, "The Sentence Completion Test", In *L.E. Abt & L. Bellak(Eds.), Projective Psychology : Clinical Approaches to the Total Personality*, 1950, 357-402.
- Sakthivel, V., K. Srihari, S. Karthik, and C.D. Anisha, "Intelligent Counselling Bot Using Ranking Algorithm in AI", *International Conference on Soft-Computing and Network Security(ICSNS)*, Coimbatore, India, 2018, 1-7.
- Hochreiter, S. and J. Schmidhuber, "Long Short Term Memory", *Neural Computation*, Vol. 9, No.8, 1997, 1735-1780.
- songys, "AwesomeKorean_Data", 2020, Available at https://github.com/songys/AwesomeKorean_Data (Accessed May 2. 2021).

◆ About the Authors ◆



안 세 훈 (ansehoon1999@gachon.ac.kr)

가천대학교에 AI · 소프트웨어학부 학사과정을 이수하고 있다. 주요 관심 분야는 머신러닝, 딥러닝, 자연어처리를 포함한다.



정 옥 란 (orjeong@gachon.ac.kr)

이화여자대학교 컴퓨터공학과(박사), University of Illinois와 서울대학교 (박사후 연구원), 현재 가천대학교 AI · 소프트웨어학부 부교수로 재직 중이다. 주요 연구분야는 빅데이터 마이닝, 머신러닝, 딥러닝, 인공지능을 포함한다.