

NLP 를 이용한 연애폰치

김형준, 박정민, 유재형, 이성주

1. 서론

1.1. 연구배경

연애는 사람들의 삶에서 매우 중요한 부분을 차지하며, 대다수의 사람들이 행복한 연인 관계를 유지하기 위해 노력하고 있다. 하지만, 연애는 대개 쉽지 않은 일이며, 다수의 사람들이 연인과의 관계에서 어려움을 겪기도 한다. 이러한 어려움은 종종 의사소통의 부재, 자존감의 부족, 불안, 집착 등으로 인해 발생한다.

이러한 문제를 해결하기 위해 많은 사람들은 연애폰치를 이용한다. 연애폰치는 전문적인 코치가 고객들과 함께 대화하고, 그들이 더 행복한 연인 관계를 유지할 수 있도록 지원해주는 서비스이다. 그러나 연애폰치는 전문적인 지식과 시간, 비용 등이 필요한 서비스이기 때문에, 많은 사람들이 이를 이용하는 데 어려움을 겪는다.

이에 대한 해결책으로, 자연어처리(Natural Language Processing, NLP) 기술을 활용한다. 사용자의 언어 데이터를 토대로 감정을 분석하여(semantic analysis) 그들이 직면한 문제를 파악하고, 그에 따른 조언을 제공해주는 인공지능 코치 역할을 할 수 있는 프로젝트를 만들고자 한다. 이를 통해, 연애폰치를 이용하기 어려워했던 사람들의 접근성을 높여, 더 행복한 연인 관계를 유지할 수 있도록 돕고자 한다.

1.2. 연구목표

1. 자연어처리(NLP) 기술을 활용하여 사용자의 언어 데이터로부터 감정을 분석하여, 그들이 직면한 문제를 파악하는 모델을 개발한다.
2. 파악한 사용자의 문제에 대해, 전문적인 연애폰치가 제공하는 조언을 모델이 자동으로 제공할 수 있도록, 인공지능 모델을 구축한다.

3. 이 프로젝트를 통해, 연애키친 서비스를 이용하기 어려웠던 사용자들도 쉽게 접근하여, 더 행복한 연인 관계를 유지할 수 있도록 지원한다.

4. 더 나아가, 이 프로젝트를 통해 수집된 사용자의 언어 데이터를 분석하여, 연애 관련 문제에 대한 인사이트를 얻을 수 있으며, 이를 활용하여 더 나은 연애키친 서비스를 제공할 수 있도록 발전시킨다.

2. 관련연구

2.1. NLP

자연어 처리(NLP)는 기계가 인간의 언어를 이해, 해석 및 생성할 수 있도록 하는 모델 및 알고리즘 개발에 중점을 둔 연구 분야이다. 인터넷에서 사용할 수 있는 텍스트 데이터의 양이 증가함에 따라, NLP 는 기업과 연구원 모두에게 필수적인 도구가 되었다.

최근 몇 년 동안 NLP 연구에서 가장 중요한 혁신은 대규모 사전 훈련 방법의 개발이다. 예를 들어 BERT(Bidirectional Encoder Representations from Transformers) 모델은 언어 모델링에서 질문 응답 및 감정 분석에 이르기까지 광범위한 NLP 작업에서 놀라운 결과를 보여주었다. 사전 학습 방법을 사용하면, 모델이 방대한 양의 텍스트 데이터에서 학습하여 언어의 맥락과 뉘앙스를 더 잘 포착할 수 있다.

그러나, NLP 연구자들이 직면한 과제는 여전히 많다. 가장 큰 장애물은 언어 데이터의 다양성 부족이다. 대부분의 NLP 모델은 영어 데이터로 학습되며, 다른 언어로 학습된 모델은 리소스와 데이터가 제한적인 경우가 많다. 따라서, 다른 언어, 방언 및 악센트의 의미와 뉘앙스를 정확하게 캡처할 수 있는 모델을 만드는 것이 어렵다.

NLP 의 또 다른 과제는 언어의 빈정거림과 아이러니를 이해하는 것이다. 빈정거림과 아이러니가 소셜 미디어에 널리 퍼져 있으며, 기계가 그러한 언어의 의도를 이해하기 어려운 경우가 많다. 예를 들어 "좋아요, 또 다른 월요일!"과

같은 문장은 긍정적으로 보일 수 있지만, 반대 감정을 전달하기 위해 종종 냉소적으로 사용된다.

이러한 문제에도 불구하고 NLP 는 chatbot 및 가상 비서에서 자동 번역 및 소셜 미디어의 감정 분석에 이르기까지 많은 잠재적 응용 프로그램을 가지고 있다. NLP 는 이미 고객 서비스 경험을 개선하고, 온라인 남용 및 증오심 표현을 식별 및 완화하고, 검색 엔진 결과를 개선하는 데 사용되고 있다. 지속적인 연구 개발을 통해 NLP 는 앞으로 우리 일상 생활에서 훨씬 더 중요한 역할을 할 것이다.

결론적으로 NLP 는 최근 몇 년 동안 상당한 발전을 이루었지만, 아직 해야 할 일이 많다. 연구자는 NLP 모델이 인간 언어의 뉘앙스를 정확하게 캡처할 수 있도록 데이터 다양성 및 풍자 이해와 같은 문제를 해결해야 한다. 이러한 어려움에도 불구하고, NLP 는 우리가 기계와 상호 작용하는 방식을 혁신하고 삶을 더 쉽고 효율적으로 만들 수 있는 잠재력을 가지고 있다.

2.2. Transformer

Transformer 모델은 최근 몇 년 동안 자연어 처리(NLP) 분야에서 엄청난 인기를 얻은 일종의 neural network architecture 이다. 이러한 모델은 language translation, text summarization, sentiment analysis 와 같은 다양한 애플리케이션에서 사용되었다.

Transformer 의 가장 중요한 장점은 input sequence 에서 long-range dependency 를 모델링 하는 기능이다. 기존의 순환 신경망(RNN)은 vanishing gradient problem 으로 인해 long sequence 를

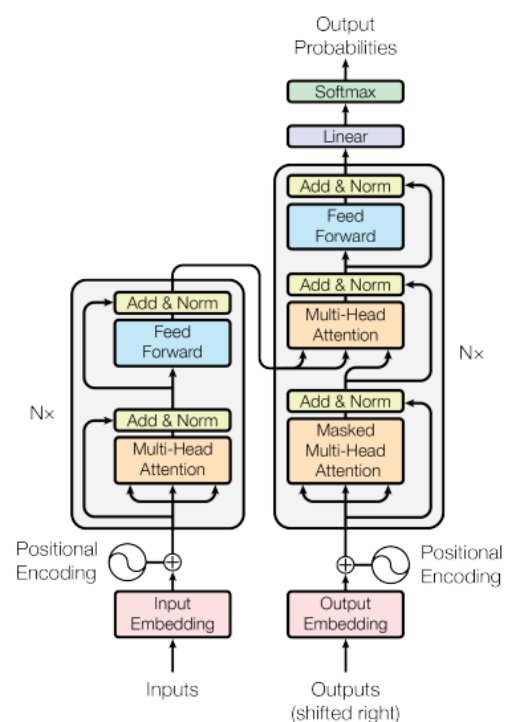


Figure 1: The Transformer - model architecture.

처리하는 능력이 제한된다. Transformers 는 다양한 중요도 수준으로 input sequence 의 다른 부분에 주의를 기울일 수 있는 self-attention mechanism 을 사용하여 이 문제를 해결한다.

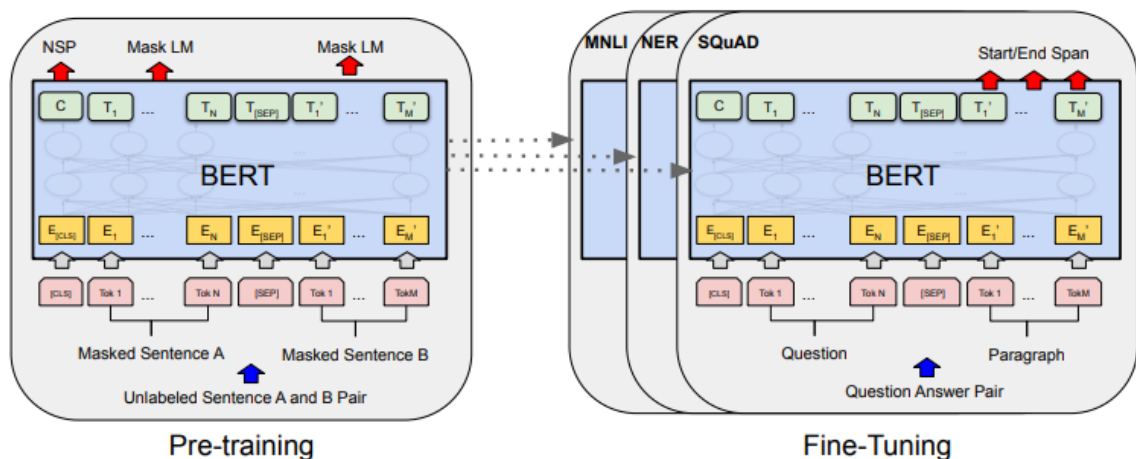
다양한 NLP 작업에서 최첨단 결과를 달성한 여러 transformer-based model 이 개발되었다. 예를 들어 OpenAI 에서 개발한 GPT-3 모델은 놀라운 language generation 기능을 보여줬고, Google 에서 개발한 BERT 모델은 question-answering, sentiment analysis, natural language inference 을 포함한 다양한 NLP 작업에서 뛰어난 성능을 보여주었다.

이러한 큰 pre-trained model 외에도 연구자들은 제한된 양의 training data 로 특정 작업에 대해 이러한 모델을 fine-tuning 하는 방법을 개발했다. 이 transfer learning approach 방식은 다양한 NLP 작업의 성능을 향상시키는 동시에 training 에 필요한 데이터의 양을 줄이는 데 매우 효과적인 것으로 나타났다.

Transformers 는 NLP 분야에 혁명을 일으켰고, 다양한 응용 분야에서 상당한 발전을 이루었다. 새로운 transformer-based model 의 개발과 기존 모델의 개선은 활발한 연구 분야이며, 앞으로 이 분야에서 더 많은 흥미로운 발전을 기대할 수 있다.

2.3. BERT

BERT(Bidirectional Encoder Representations from Transformers)는 Google AI Language 에서 개발한 pre-trained language model 이다. 일반적인 언어 표현을 학습하기 위해 대규모 텍스트 corpus 에서 훈련된 transformer model 을 기반으로 하는 deep neural network architecture 를 사용한다.



2018 년 출시 이후 BERT 는 text classification, question answering, named entity recognition 과 같은 자연어 처리(NLP) 작업에서 가장 인기 있고 널리 사용되는 언어 모델 중 하나가 되었다. BERT 의 핵심 혁신은 왼쪽 및 오른쪽 context 를 모두 고려하여, 문장에서 단어의 context 와 의미를 capture 하는 기능이다.

여러 연구에 따르면, BERT 는 다양한 NLP 작업에서 이전의 최신 모델보다 훨씬 뛰어난 성능을 보인다. 예를 들어 Stanford Question Answering Dataset(SQuAD)에서 BERT 는 93.2%의 점수를 달성하여 이전 최고 점수인 90.9%를 넘어섰다. 마찬가지로 다양한 NLP 작업에 대한 모델의 성능을 측정하는 GLUE benchmark 에서 BERT 는 9 개 작업 모두에서 최첨단 결과를 달성했다.

또한 BERT 는 biomedical text mining, sentiment analysis, text summarization 과 같은 특정 도메인 및 작업에 맞게 adapted, fine-tuned 되었다. 이러한 적응성과 다양성 덕분에 BERT 는 NLP 의 다양한 분야에서 일하는 연구자와 실무자를 위한 강력한 도구가 되었다.

전반적으로 BERT 모델은 언어 이 해 및 처리를 위한 강력하고 유연한 도구를 제공함으로써 NLP 분야에 혁명을 일으켰다. 그 영향은 다양한 NLP 작업에서 최첨단 결과를 달성하기 위해 그 기능을 활용한 수많은 응용 프로그램 및 연구 연구에서 분명하다.

2.4. KoBERT

KoBERT 는 한국어로 된 자연어 처리(NLP) 작업을 위한 pre-trained language model 이다. 한국과학기술정보연구원(KISTI)과 SK Telecom 이 공동으로 개발했다. KoBERT 는 pre-training language 표현을 위해 설계된 neural network architecture 인 인기 있는 BERT 모델을 기반으로 한다.

KoBERT 는 한국어 텍스트 데이터의 대규모 corpus 에 대한 교육을 받았으며, test classification, question answering, language generation 과 같은 다양한 NLP 작업에 대해 fine-tuning 할 수 있다. 한국어 NLP 작업에 대한 여러 benchmark dataset 에서 최첨단 성능을 달성했다.

KoBERT 는 한국에서 NLP 연구 및 응용을 위한 인기 있는 도구가 되었으며, pre-trained model 과 code 는 공개적으로 사용할 수 있다.

2.5. 연애폰치

연애폰치 및 관련 분야에는 다양한 연구들이 진행되고 있습니다. 여기에는 그중 일부 연구들을 소개한다.

1. 연애폰치에 대한 사용자 만족도와 효과성에 대한 연구: 이 연구는 연애폰치 서비스를 이용한 사용자들의 만족도와 효과성에 대해 조사하고 분석한다.
2. 연애 관련 문제와 자존감 간의 상관관계 연구: 이 연구는 연애 관련 문제와 자존감 간의 상관관계를 조사하고, 자존감을 높이는 방법에 대해 연구한다.
3. 의사소통에 대한 연구: 이 연구는 연인 관계에서 의사소통의 중요성과, 의사소통을 향상시키는 방법에 대해 조사하고 분석한다.
4. 행동분석을 이용한 연애 코칭 연구: 이 연구는 사용자들의 행동을 분석하여, 그들이 더 행복한 연인 관계를 유지할 수 있도록 조언을 제공한다.
5. 긍정심리학을 이용한 연애 코칭 연구: 이 연구는 긍정심리학을 이용하여, 사용자들의 긍정적인 감정과 태도를 유지하며, 더 행복한 연인 관계를 유지할 수 있도록 조언을 제공한다.
6. 인공지능을 이용한 감정인식 연구: 이 연구는 자연어처리와 인공지능 기술을 이용하여, 사용자들의 감정을 인식하고, 그에 따른 조언을 제공한다.

7. 연애 관련 빅데이터 분석 연구: 이 연구는 사용자들의 연애 관련 데이터를 수집하고 분석하여, 연애에 대한 인사이트를 얻고, 연애폰칭 서비스를 발전시킨다.

2.6. 연애폰칭 어플리케이션

연애폰칭 어플리케이션에 대한 연구도 많이 진행되어 왔다. 여기에서는 기존의 연애폰칭 어플리케이션에 대한 일부 연구를 소개한다.

1. 연애폰칭 어플리케이션의 사용자 경험 평가 연구: 이 연구는 연애폰칭 어플리케이션의 사용자 경험을 평가하고, 사용자가 어플리케이션을 사용할 때 겪는 문제점을 파악한다.

2. 연애폰칭 어플리케이션의 효과성에 대한 연구: 이 연구는 연애폰칭 어플리케이션의 효과성을 평가하고, 사용자들의 연애 관련 문제를 해결하는데 얼마나 효과적인지를 분석한다.

3. 인공지능 기술을 활용한 연애폰칭 어플리케이션 연구: 이 연구는 인공지능 기술을 활용하여, 사용자들의 연애 관련 문제를 분석하고, 그들에게 맞는 조언을 제공하는 어플리케이션을 개발한다.

4. 빅데이터 분석을 이용한 연애폰칭 어플리케이션 연구: 이 연구는 빅데이터 분석 기술을 활용하여, 사용자들의 연애 관련 데이터를 수집하고 분석하여, 어플리케이션에서 사용자들에게 제공하는 조언을 개선한다.

5. VR 기술을 활용한 연애폰칭 어플리케이션 연구: 이 연구는 가상현실(VR) 기술을 활용하여, 사용자들이 실제 상황에서 연애 관련 문제를 경험하듯이 시뮬레이션하며, 문제 해결 능력을 향상시키는 어플리케이션을 개발한다.

6. 의사소통 훈련을 위한 연애폰칭 어플리케이션 연구: 이 연구는 사용자들이 연인과의 의사소통 능력을 향상시키기 위한 훈련 프로그램을 제공하는 어플리케이션을 개발한다.

2.7. 기존 연구의 문제점 및 해결 방안

기존 연구들도 많은 노력과 시간이 투입되었지만, 몇 가지 문제점이 있었다. 여기에는 몇 가지 문제점과 해결 방안을 소개한다.

1. 적은 데이터 양: 연애폰칭 어플리케이션을 만들기 위해 필요한 데이터 양이 매우 많은데, 기존 연구에서는 그 양이 부족한 경우가 있었다. 이를 해결하기 위해서는 더 많은 데이터를 수집하고, 이를 분석하는 빅데이터 기술 등을 활용해야 한다.

2. 개인화된 조언의 부재: 기존 연구에서는 대부분의 어플리케이션에서 일반적인 조언만을 제공하는 경우가 많았다. 이는 사용자들이 자신의 문제 상황에 따라 맞는 조언을 받기 어렵다는 것을 의미한다. 이를 해결하기 위해서는 인공지능 기술과 데이터 분석 기술을 활용하여, 사용자들에게 개인화된 조언을 제공하는 어플리케이션을 개발해야 한다.

3. 실시간 상담 부재: 일부 연구에서는 어플리케이션에서 제공하는 조언만으로는 사용자들의 문제를 해결하기 어려운 경우가 있다. 이 경우, 사용자들은 실시간으로 상담을 받을 수 있는 기능이 필요하다. 이를 해결하기 위해서는 어플리케이션에서 실시간 채팅 또는 영상 통화 기능을 제공하는 것이 필요하다.

4. 피드백 및 개선 방안 제공의 부재: 어플리케이션을 사용하는 사용자들의 피드백을 수집하고, 이를 바탕으로 어플리케이션을 개선하는 것이 중요하다. 그러나 기존 연구에서는 이에 대한 노력이 부족한 경우가 많았다. 이를 해결하기 위해서는 사용자들의 피드백을 수집하고, 이를 바탕으로 어플리케이션을 지속적으로 개선하는 시스템을 마련해야 한다.

3. 제안하는 프로젝트 소개

3.1. 시나리오

이 프로젝트는 개인의 연애 실력 향상과 이상적인 파트너 찾기를 목표로 새롭게 제안된 연애 코칭 서비스이다. 우리는 기계 학습 및 자연어 처리의 힘을 활용하여 고객에게 개인화된 코칭 조언을 제공하고자 한다. 이를 위해 Google 에서 자연어 처리 작업을 위해 개발한 최신 언어 모델인 BERT 모델을

사용할 계획이다. BERT 모델은 질문 답변, 감정 분석, 언어 번역 등 다양한 NLP 작업에서 인상적인 결과를 얻었다. BERT의 감정 분석 작업을 사용하여, 고객에게 특정 요구 사항과 선호도에 맞는 정확하고 관련성 높은 코칭 조언을 제공하고자 한다.

3.2. 요구사항

1. Data collection: 연애 프로필, 채팅 로그, 온라인 데이트 메시지와 같은 데이트 관련 텍스트의 대규모 dataset을 수집해야 한다. 이 dataset은 연애 context에서 사용되는 언어를 이해하도록 BERT 모델을 교육하는데 사용된다.

2. BERT model training: 수집된 dataset에 적절한 training algorithm을 사용하여 BERT model을 training해야 한다. 이 단계에서는 pre-trained BERT model을 fine-tuning하여 감정분석이라는 특정 작업에 맞게 조정한다.

3. Coaching advice generation: client가 제공한 입력을 기반으로 개인화된 코칭 조언을 생성할 수 있는 시스템을 개발해야 한다. 이 시스템은 trained BERT model을 사용하여 client의 입력을 이해하고 관련 코칭 조언을 생성한다.

4. User Interface: 고객이 코칭 서비스와 상호 작용할 수 있도록 직관적이고 User-friendly한 interface를 만들어야 한다. Interface는 고객이 연애 관련 질문을 입력하고, 명확하고 간결한 방식으로 코칭 조언을 받을 수 있도록 해야 한다.

5. Quality assurance: 시스템에서 생성된 코칭 조언이 고객에게 정확하고 적절하며 유용한지 확인해야 한다. 코칭 조언의 정확성과 효과를 개선하기 위해 정기적인 품질 점검을 실시하고 고객으로부터 피드백을 구해야 한다.

6. Data privacy: 고객으로부터 수집된 데이터가 데이터 보호법에 따라 안전하게 처리되도록 해야 한다. 적절한 데이터 보안 조치를 구현하고, 데이터 수집 및 처리에 대해 고객으로부터 정보에 입각한 동의를 얻는다.

4. 향후 일정 및 역할 분담

4.1. 향후 일정

1. Data Collection: 연애 프로필, 메시지 및 조언글과 같은 데이트 관련 텍스트의 대규모 dataset 을 수집해야 한다. 여기에는 web scrapping 및 manual curation 이 포함된다.
2. Preprocessing: 노이즈 및 관련 없는 정보를 제거하려면 dataset 을 정리하고 전처리(preprocess)해야 한다.
3. Training BERT: data 가 정리되고 전처리되면, unsupervised learning 을 사용하여 BERT model 을 training 한다.
4. Fine-tuning BERT: 모델을 training 한 후, supervised learning 을 사용하여 모델을 fine-tuning 한다. 이를 통해 연애폰칭의 특정 작업에 대한 감정 분석 성능을 최적화한다.
5. User interface development: 사용자가 연애폰칭 관련 텍스트를 입력하고 BERT 모델로부터 피드백 및 코칭 제안을 받을 수 있는 연애폰칭 시스템을 위한 UI 를 개발한다.
6. Testing and evaluation: 시스템이 효과적으로 User-friendly 한지 확인하려면, 시스템을 엄격하게 테스트하고 평가해야 한다.

4.2. 역할 분담

1. Data Scientist: Data Scientist 는 BERT 모델에 필요한 데이터를 수집, 정리 및 분석할 책임이 있다. 또한, 모델의 training 및 평가 작업을 수행하여 모델이 예상대로 작동하는지 확인한다.
2. NLP Programmer: NLP Programmer 는 BERT 모델을 만들고 연애폰칭 시스템의 특정 요구 사항에 맞게 fine-tuning 하여 감정 분석을 시행한다. 모델의 출력이 정확하고 의미 있는지 확인하는 일도 담당한다.
3. UI/UX Designer: UX Designer 는 연애폰칭 시스템의 UI Design 을 담당한다. 프로젝트 관리자 및 다른 팀 구성원과 긴밀히 협력하여, UI/UX 가 직관적이고 매력적이며 효과적이 되도록 한다.

4. Content Developer: Content Developer 는 시스템을 통해 전달될 연애폰칭 콘텐츠를 작성을 담당한다.

All. Quality Assurance: 시스템이 올바르게 작동하는지 확인하고, 유저에게 정확하고 유용한 코칭을 제공하기 위해 시스템을 테스트한다. 여기에는 품질 보증과 데이터 분석이 포함되어야 한다.

5. 결론 및 기대효과

연애폰칭 어플리케이션은 다양한 연애 상황에서 사용자에게 조언을 제공하여, 사용자의 연애 능력 향상에 도움을 주는 서비스이다. 이를 통해, 사용자는 더 건강하고 성숙한 연애를 할 수 있으며, 자신의 연애 능력을 개선하면서 자신감도 향상된다. 또한, 사용자들 간의 정보 공유를 통해, 서로의 연애 능력을 공유하고 발전시키는 커뮤니티가 형성된다.

기대효과로는, 연애폰칭 어플리케이션을 이용한 사용자들은 더욱 건강하고 성숙한 연애를 할 수 있어, 연애와 관련된 문제에서 비롯된 스트레스를 줄이고 행복한 삶을 살 수 있다. 또한, 사용자들 간의 정보 공유를 통해 연애 능력이 개선되면, 이를 통해 사회 전반적으로 건강한 인간관계와 사랑의 문화가 확산된다. 마지막으로, 개인화된 조언을 제공하는 연애폰칭 어플리케이션은 사용자들의 성공적인 연애를 돕는데, 이를 통해 사랑의 만족도가 높아져, 개인의 삶 뿐만 아니라 사회 전반의 행복도 증진된다.

6. 참고문헌

Park, S., & Kim, K. J. (2018). A Study on the Development of Relationship Coaching System. Journal of the Korea Society of Computer and Information, 23(12), 23-30.

Lee, J. A., & Jeong, M. (2020). A Study on the Development of AI-based Couple Counseling System. Journal of Digital Convergence, 18(5), 225-233.

Kim, S. Y., & Shin, D. (2019). A Study on the Design of Relationship Coaching App Based on the Analysis of User Needs. *Journal of Digital Convergence*, 17(10), 99-106.

Lee, S., Lee, J., & Park, Y. (2020). Analysis of Relationship Coaching Apps and Development of Recommendation System. *Journal of Digital Convergence*, 18(8), 389-397.

Moon, S., Kim, Y., & Lee, M. (2021). Design of Personalized Relationship Coaching Service Based on Machine Learning. *Journal of Digital Convergence*, 19(1), 115-122.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin. (2017). Attention Is All You Need.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding