# 채팅에서 사용자의 입력을 AI로 분석하여 피드백을 제공하는 연구

이승헌[1], 김도연[2], 김태윤[2], 조성호[2], 유석봉[3]
[1]전남대학교 수학과 학부생
[2]전남대학교 소프트웨어공학과 학부생
[3]전남대학교 소프트웨어공학과 교수
aaalsh13@gmail.com, kehdus18@gmail.com, xodbs507@gmail.com, tjdgh5463@naver.com,
sbyoo@jnu.ac.kr

# A study on providing feedback by analyzing user input in chat using AI

Seungheon Lee[1], Doyeon Kim[2], Taeyun Kim[2], Seongho Joe[2], Seokbong Yoo[3,]
[1]Dept. of Mathemetics, Cheonnam National University
[2]Dept. of Software Engneering, National University
[3]Dept. of Software Engneering, National University

요 약

온라인에서의 비격식적적 소통은 다양한 언어 문제를 발생시킨다. 사용자의 언어를 분석하여 교정해 주는 프로그램을 통해 언어생활의 문제를 완화할 수 있다. 이 논문에서는 트렌스포머(Transformer)를 이용한 거대 언어 모델(LLM)을 통하여 잘못된 맞춤법을 검사하는 방법을 살피고자 한다. 기존의 방법과 달리, 언어 모델의 자체적인 사고력을 바탕으로, 학습되지 않은 분야에서의 맞춤법도 보다 효과적으로 검사할 수 있다. 적절한 사전 학습 모델을 선택하고, 정교한 프롬프트를 구성함으로써 프롬프트가 있기 전 0.09에서 0.27로 약 3배 가량의 정확도 향상을 얻을 수 있다. 이러한 결과를 바탕으로 보다 신뢰성 있는 언어 사용 교정 애플리케이션을 개발 할 수 있다.

## 1. 서론

디지털 시대에 사회적 상호작용의 많은 양은 온라인 소통으로 이뤄진다. 그러나 온라인 소통의 비격식성은 철자나 구조가 잘못된 문장 및 비속어를 사용하는 문장을 비롯하여 언어적 문제를 지닌 글을 발생시키고 있다. 이런 언어에 자주 노출되어 생긴 잘못된 습관은 의사소통의 질에 영향을 미칠 뿐만 아니라 부정적인 채팅 문화를 조장한다.

본 논문에서는 잘못된 온라인 언어 사용을 바로잡는 애플리케이션 및 맞춤법 검사에 LLM(거대 언어 모델)을 사용하는 방법을 제안한다.

제안하는 애플리케이션은 잘못된 문장 구조 및 철자를 교정하는 능력을 갖추고 있으며, 이는 사용자의 스마트폰에 통합되어 잘못된 언어 사용을 즉각적으로 교정하게 된다. 또한 프로그램은 사용자의 입력을 자동으로 분석하여 자주 사용하는 비속어나, 자주 틀리는 맞춤법을 사용자에게 제공하여 이용자가 스스로 언어 습관을 개선하도록 돕는다.

한편, 이 논문에서는 이러한 애플리케이션을 개발하는 과정에서 LLM을 어떻게 적용할 수 있는지를 분석한다. 최근 LLM은 사고력[2]이 필요한 문제를 풀 수 있을 만큼 강력한 성능을 보인다. 그러나 거대한 모델을 바로 사용하는 것은 스마트폰의 연산 성능상 불가능할 뿐만 아니라, 충분한 정확도를 확보하지 못한다. 이러한 문제를 해결하려는 방법으로 본 논문에서는 적절한 사전 학습 모델 선정 및 few-shot[3]을 비롯한 프롬프트 엔지니어링을 제시한다. 이 논문에서 제안하는 프로그램을 통하여 이용자들이 온라인상에서 건전한 언어생활을 유지할 수 있으리라 사료된다.

## 2. 사용자 맞춤형 피드백 제공

사용자의 언어 습관을 효과적으로 교정하기 위해서는 언어 사용의 습관을 개별로 분석하여야 한다. 사전에 비속어에 대해, 사용자가 입력하는 빈도를 통계 내야 한다. 또한, 틀린 맞춤법과 비문을 종류별로 분류하여, 자주 틀린 내용을 바탕으로 맞춤 설명을 제공해야 한다. 이러한 프로그램의 설명은 사용자가 직관적으로 이해할 수 있으면서도, 언어 습관을 개선할 수 있을 만큼 자세하여야 한다.

## 3. 자연어 처리 모델의 효과적 이용

임의의 사용자 입력에 대해서 효과적으로 맞춤법을 검사하기 위해서는 강력한 언어 모델이 필요하다. 특히 RNN, CNN보다 언어 모델링에서 높은 성능을 보이는 트랜스포머(Transformer)모델을 사용함이 적절하다. 그러나 강력한 언어 모델만으로는 실용적인 정확도를 얻을 수 없다. 본 논문에서는 적절한

사전 학습 모델의 선정 및 few-shot을 비롯한 프롬프트 사용이 맞춤법 검사 정확도에 어떠한 영향을 미치는지를 확인하고자 한다. 이를 위해 한국어 데이터셋[4, 5]에서 미세 조정한 Llama-3.2 3B[6] 모델과 Qwen 2.5 3B[7]에 대해, few-shot 프롬프트 사용 유무에 따른 성능 변화를 실험하였다. 검증에는 데이터셋[8]의 validation 분할을 사용하였다.

| 모델 | few-shot | 정확도 |
|---|---|---|
| Llama | O | 8.36% |
| | X | 0.00% |
| Qwen | O | 27.19% |
| | X | 9.51% |

우선 Llama를 사용한 결과 대비 Qwen을 사용한 결과가 일관적으로 좋음을 알 수 있다. 이는 Llama는 한국어에서 사전 학습되지 않았지만, Qwen은 한국어에 대해서 사전 학습되었기 때문이다. 또한 두 모델 전부에서 few-shot 프롬프트를 사용한 쪽이 성능이 높았다. 이는 모델이 맞춤법 검사라는 특별한 목적 및 입출력 양식에 학습되지 않았기 때문으로, 예시를 제공하면 모델의 실수가 줄어듦을 볼 수 있다.

## 4. 언어 모델의 반응 속도 최적화

채팅이나 소셜 미디어 등에서 본 애플리케이션을 활용하기 위해서는 모델이 사용자의 입력에 즉시 반응해야 한다. 이를 위해서는 언어 모델의 압축 및 최적화가 필요 한데, 본 논문에서는 양자화를 사용할 것을 제안한다. 모델의 양자화는 단정밀도나 반정밀도 부동소수점으로 구성된 모델의 파라미터를 8비트나 4비트의 등 보다 적은 용량의 정수로 나타내 사용하는 것을 의미한다. 특히, GPTQ[9]를 비롯한 최신 양자화 기술들은 언어모델의 당혹감(Perplexity)를 크게 악화시키지 않으면서도 필요한 메모리 및 연산의 양은 단정밀도에 비하여 1/4로 줄이도록 돕는다.

## 5. 결론

디지털 시대가 본격적으로 대두되면서, 문자를 통한 비격식적 의사소통은 어느새 언어생활의 대다수를 차지하게 되었다. 이 과정에서 철자 및 문법적 오류를 가진 문장 또는 비속어를 남용하는 문장의 활용도 늘고 있다. 본 논문에서는 이러한 언어생활의 문제를 효과적으로 대처할 수 있는 프로그램을 제안하였고, 해당 프로그램에서 사용할 수 있는 맞춤법 확인 모델을 제안하였다.

최신의 언어 모델에 본 논문에서 제시한 모델 최적화 및 프롬프트 엔지니어링을 이용하면 정확한 맞춤법 검사 능력을 갖출 수 있다. 또한, 양자화를 비롯한 최적화를 통한다면 채팅을 비롯한 즉각적인 언어 활용에서 활용할 수 있는 속도를 얻을 수 있다. 한편, 프로그램이 사용자의 언어 사용 습관을 분석하고, 개선할 점을 이용자에게 알려준다면 사용자가 바람직한 언어 습관을 갖도록 도울 수 있다.

이렇게 제안한 프로그램은 아직까지 긴 길이의 글을 완전히 분석할 수 있는 능력을 갖추지는 못한다. 이러한 기술적 문제는 추후 연구를 통하여 보완될 수 있으리라 기대된다.

## 감사의 글

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 소프트웨어중심대학사업, 인공지능융합혁신인재양성사업, 대학ICT연구센터사업의 연구결과로 수행되었습니다. (2021-0-01409, 2023-00256629, 2024-00437718)

## 참고문헌

[1] Vaswani, A. "Attention is all you need." Advances in Neural Information Processing Systems (2017).

[2]Bubeck, Sébastien, et al. "Sparks of artificial general intelligence: Early experiments with gpt-4." arXiv preprint arXiv:2303.12712 (2023).

[3]Laria Reynolds and Kyle McDonell. 2021. Prompt Programming for Large Language Models: Beyond the Few-Shot Paradigm. In Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems (CHI EA '21). Association for Computing Machinery, New York, NY, USA, Article 314, 1 – 7.

[4]https://huggingface.co/datasets/maywell/korean_textbooks

[5]https://huggingface.co/datasets/jojo0217/korean_rlfh_dataset

[6] Dubey, Abhimanyu, et al. "The llama 3 herd of models." *arXiv preprint arXiv:2407.21783* (2024).

[7]Qwen Team "Qwen2 Technical Report " arXiv preprint arXiv:2407.10671, 2024

[8]https://aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&aihubDataSe=data&dataSetSn=71560

[9] Frantar, Elias, et al. "Gptq: Accurate post-training quantization for generative pre-trained transformers." *arXiv preprint arXiv:2210.17323* (2022).