

캡스톤디자인 I 계획서

제 목	국문	LLM 기반 한국어 문법 오류 교정 서비스			
	영문	LLM-based Korean Grammatical Error Correction Service			
프로젝트 목표 (500자 내외)	<p>본 프로젝트의 핵심 목표는 한국어의 복잡한 문법과 문맥을 깊이 있게 이해하고, 이를 정밀하게 교정할 수 있는 한국어 특화 LLM 기반의 문법 오류 교정 웹 서비스를 개발하는 것이다. Meta의 최첨단 Transformer 기반 디코더 모델인 LLaMA를 파인튜닝하여, 사용자가 작성한 글의 문법 오류를 자동으로 감지하고 수정 제안을 제공하는 모델을 개발한다. 이를 통해, 사용자는 글쓰기의 정확성을 높이고, 효과적인 의사소통을 할 수 있다.</p> <p>또한 MLOps 시스템을 구축하여 모델 연구를 가속화하고, 모델 모니터링, 자동화된 학습 및 배포 파이프 라인을 구성한다. 이를 통해, 모델 연구를 가속화하고 지속적인 서비스 개선을 가능케 한다.</p> <p>[논문 투고]</p> <p>KCC(컴퓨터 종합 학술대회) 수상, SCI 이상 학회를 목표로 한다.</p>				
프로젝트 내용	<p>본 프로젝트의 주요 수행 내용은 다음과 같다</p> <p>1. 모델 개발: LLaMA 모델을 파인튜닝하여 한국어에 특화된 GEC를 수행하는 모델을 개발한다. 두 단계에 걸쳐 학습이 진행되며 각 단계에서 GEC, 글쓰기 평가를 수행할 수 있도록 모델을 학습한다.</p> <p>3. MLOps 인프라 구축: 효과적인 모델 관리와 자동화된 배포를 위해 MLOps 인프라를 설계 및 구현한다. 이는 지속적 통합(CI), 지속적 배포(CD), 모델 모니터링 및 로깅을 포함한다.</p> <p>4. 웹 어플리케이션 개발: 최종 사용자가 모델을 쉽게 사용할 수 있도록 사용자 친화적인 웹 서비스를 개발한다. 사용자는 웹상에서 실시간으로 모델의 도움을 받아 글의 품질을 높일 수 있다.</p>				
중심어(국문)	거대 언어 모델	문법 오류 교정	MLOps	웹 어플리케이션	
Keywords (english)	LLM	GEC (Grammatical Error Correction)	MLOps	Web Application	
멘토	소속	서울과학기술대학교	이름	임경태	
팀 구성원	학년/반	학 번	이 름	연락처(전화번호/이메일)	
	4	20191120	이용빈	010-6560-9717/20191120@edu.hanbat.ac.kr	
	4	20191780	육정훈	010-4713-9845/20191780@edu.hanbat.ac.kr	
<p>컴퓨터공학과와 캡스톤디자인 관리규정과 모든 지시사항을 준수하면서 본 캡스톤디자인을 성실히 수행하고자 아래와 같이 계획서를 제출합니다.</p> <p>2024 년 3월 6일</p> <p>책 임 자 : 이용빈 (인)</p> <p>희망 지도교수 : 장한얼</p>					

1. 캡스톤디자인의 배경 및 필요성

글로벌 커뮤니케이션과 디지털 콘텐츠의 중요성이 증가함에 따라, 정확하고 효과적인 글쓰기가 어느 때보다 중요해졌다. 이러한 시대적 요구를 반영하여, 다양한 언어의 문법 및 맞춤법을 검토하고 문맥까지 고려하여 적절한 표현을 제안하는 인공지능(AI) 기반의 Grammatical Error Correction (GEC) 서비스들이 출시되었다. 대표적으로 Grammarly는 영어 글쓰기를 지원하는 인공지능 기반 GEC 서비스를 제공하지만, 그 기능이 영어에 국한되어 한국어 사용자에게는 적용 범위가 제한적이다.

부산대 맞춤법 검사기 등 한국어에 특화된 서비스가 존재하지만, 맞춤법 교정에만 기능이 한정되어 있다. 또한, 한국어는 그 구조와 문법이 영어와 크게 다르며, 특히 조사 사용, 어미의 변화, 복잡한 띄어쓰기 규칙 등 한국어만의 독특한 문법적 특성을 가지고 있다. 따라서 한국어의 복잡한 문법 구조와 언어적 특성을 이해하고, 이를 바탕으로 한국어 작문을 지원하는 특화된 인공지능 모델 개발이 필요하다.

한국어 특화 GEC 모델은 한국어 문서 작성에 있어 그 품질을 향상시키고, 효과적인 커뮤니케이션을 지원하는 역할을 할 수 있다. 특히, LLM 기술이 발전함에 따라, 한국어의 특성을 이해하고, 사용자에게 실질적인 글쓰기 도움을 제공할 수 있는 기술적 가능성이 확대되고 있다. 본 프로젝트는 이러한 기술적 발전을 바탕으로 한국어 사용자들에게 GEC를 제공하는 LLM 모델을 개발하고, 이를 기반으로 웹 서비스를 구축하여 한국어 작문을 지원하고자 한다.

2. 캡스톤디자인 목표 및 비전

본 프로젝트는 한국어에 특화된 문법 및 맞춤법을 교정하고 글쓰기를 보조하는 서비스 개발을 목표로 한다. 현재 시장에는 영어 사용자를 위한 글쓰기 보조 도구가 널리 이용되고 있으나, 한국어 사용자를 위한 서비스는 부족한 실정이다. 이에 본 프로젝트는 한국어에 최적화된 GEC를 수행하는 LLM 모델을 개발하고, 이를 서비스하기 위해 웹 어플리케이션을 구축할 것이다.

LLM 모델은 LLaMa 기반의 모델을 파인튜닝하고 성능을 평가할 것이다. 글쓰기 평가 데이터셋을 활용하여 문법 오류 교정뿐만 아니라 글쓰기 평가 및 개선하는 한국어 특화 모델을 구축한다. 사용자는 해당 모델을 사용하여 문법적으로 정확한 글쓰기를 할 수 있을 뿐만 아니라 문체, 적절한 단어 선택 등에 있어 도움을 받을 수 있다.

또한, 개발한 모델을 기반으로 서비스를 구축하고 운영하기 위해 MLOps (ML + Operations)를 도입한다. 이는 모델 저장, 모니터링, 지속적 학습, ci/cd 등을 포함하며, 머신러닝 모델을 개발하고 배포하는데 필수적이다. 웹 어플리케이션을 구축하는데 있어 MLOps 원칙을 준수하여 지속적인 서비스 성능 개선을 달성하는 것을 목표로 한다.

3. 캡스톤디자인 내용

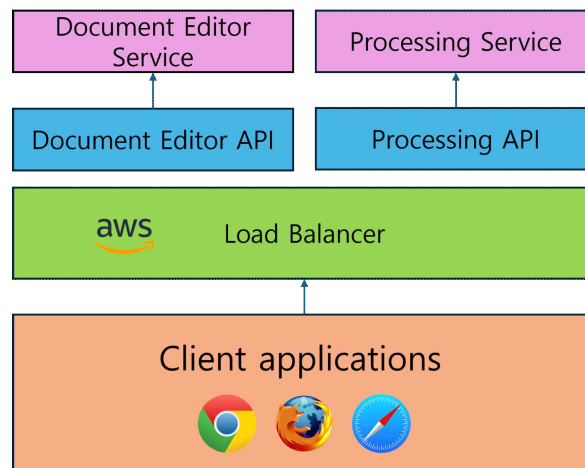
주요 기능

기능	내용
문장 교정	- 한국어 문장의 맞춤법 및 문법적 오류를 식별하고 수정
표현 개선	- 문맥상 더 적절한 단어 선택이나 문장 구조를 제안하여 글의 품질을 향상
문제 검사	- 글의 목적과 대상 독자에 맞는 적절한 문체를 유지하도록 지원
실시간 서비스	- 웹 어플리케이션을 구축하여 웹 상에서 작문과 동시에 실시간으로 GEC 서비스 제공

비기능적 요구사항

	내용
성능	- 기존의 한국어 맞춤법 검사기 이상의 성능 달성
MLOps	- MLOps를 도입하여 모델 학습 및 배포 과정을 자동화하고, 지속적인 서비스 성능 개선
결과	- 연구한 모델을 바탕으로 논문 작성 - 개발한 모델 기반의 웹 어플리케이션을 구축하여 사용자에게 기능 제공
API 공개	- 개발한 모델의 기능을 사용할 수 있는 api를 공개하여 이를 기반으로 한 어플리케이션 개발 지원

웹 어플리케이션 예상 기본 아키텍처



- Load Balancer: 사용자의 요청을 처리하고 Grammarly의 다양한 서비스(예: 문서 편집 및 문법 검사)에 접근할 수 있게 연결하는 인터페이스를 제공
- APIs: 사용자의 요청을 처리하고 다양한 서비스(예: 문서 편집 및 문법 검사)에 접근할 수 있게 연결하는 인터페이스를 제공
- Document Editor Service: 사용자가 문서를 쉽게 생성하고, 편집하며, 저장할 수 있도록 하는 기능 제공
- Processing Service: 클라이언트 애플리케이션으로부터의 요청을 처리하고, 핵심 기능인 문법 오류 교정 및 글쓰기 보조를 위한 분석과 제안을 제공

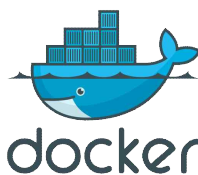
4. 캡스톤디자인 추진전략 및 방법

본 프로젝트의 모델 학습을 위해 준비된 데이터 셋은 맞춤법 교정 데이터 셋과, 글에 대한 평가 데이터 셋 두 가지이다. 두 데이터 셋 모두 연구를 위한 비공개 데이터 셋이며, 해당 데이터 셋의 학습을 위해 한국어로 튜닝된 LLaMA 기반의 모델을 instruction tuning의 형태로 추가 학습시킨다. 이러한 모델의 학습을 위해 서울과학기술대학교에서 제공하는 A6000, A100, H100 gpu 서버들을 활용할 예정이다. 서울과학기술대학교의 mlp lab을 이끌고 있는 임경태 교수를 멘토로 섭외하였다. llm 기반 자연어 처리 분야의 전문가인 임경태 교수의 멘토링을 받아 모델을 개발한다.

모델의 개발과 함께, MLOps(ML + Operations)를 도입하여 웹 어플리케이션의 개발과 운영을 효율적으로 수행한다. MLOps의 도입은 모델의 지속적인 학습, 개선, 배포를 자동화하여, 서비스의 안정성과 신뢰성을 보장하고, 신속한 업데이트를 가능하게 한다. 자동화된 테스트, 모델 버전 관리, 데이터 파이프라인 구축 및 모니터링 시스템을 구성하여 서비스 개발의 복잡성을 줄이고, 모델 개발을 효율화한다.

	팀 구성	성명	역할
1	팀장	이용빈	MLOps 구축, 웹 어플리케이션 개발
2	팀원	육정훈	모델 개발, 논문 작성

사용 프레임워크 및 기술



5. 참고문헌

- [1] Touvron, Hugo, et al. "Llama: Open and efficient foundation language models." arXiv preprint arXiv:2302.13971 (2023).
- [2] Touvron, Hugo, et al. "Llama 2: Open foundation and fine-tuned chat models." arXiv preprint arXiv:2307.09288 (2023).
- [3] Yoon, Soyoung, et al. "Towards standardizing Korean grammatical error correction: Datasets and annotation." arXiv preprint arXiv:2210.14389 (2022).
- [4] System and Organization Controls (SOC 3) Report, grammarly, (2021)