종합설계 프로젝트 제안서

팀 코드	A6	제출일	2026년 3월 17일
팀 이름	O(1)		
과제 명	"ECHO" English Correction & Hearing Optimization		
팀 구성원	김형준, 김지용, 백지영, 이찬영, 신영환, 최현지		
과제 개요	기존 발음 학습 앱의 사용자 불만은 두 가지로 수렴된다: **① 발음을 제대로 인식하지 못하는 것, ② 왜 틀렸는지 알려주지 않는 것.** 전자는 MDD(Mispronunciation Detection & Diagnosis) 시스템 자체의 한계에서 비롯된다. L2-ARCTIC 데이터셋(6 명 화자, 29,087 음소)에 Wav2Vec2-Base/Large, HuBERT-Large 세 모델로 베이스라인 실험을 수행한 결과: MDD F1 최대 0.6078, FRR 5.38~7.48%, FAR 36.85~42.77%로, FAR-FRR 사이에 trade-off 가 존재했다. Duolingo(고 FAR·관대)와 말해보카(고 FRR·엄격)는 같은 문제의 양 끝단에 있다. 한국어 화자(YKWK)에서는 MDD F1 이 0.4974~0.5166, FAR 은 최대 54.52%로 더 두드러졌다. /r/·/l/혼동, /θ/→/s/대치, 자음군 모음 삽입 등 L1 간섭 오류 패턴을 범용 모델이 충분히 포착하지 못했기 때문이다. 결과적으로 탐지 수준의 한계로 교정 피드백이 부재하며, 잘못된 쉐도잉 반복은 오히려 발음 오류를 고착시킨다.		
해결 방안	ECHO는 한국인 대상 영어 발음 교정 서비스로, 두 문제를 다음과 같이 해결한다. **① 탐지 성능 개선** - FRR 감소 우선. HuBERT 기반 음소 인식기에 Canonical Sequence Injection 을 적용해 정발음 인식 오류율(PER_C)을 낮추고, 데이터 불균형(정발음 25,710 vs 오발음 4,318)은 Focal CTC Loss 로 보완한다. 한국어 화자 성능 저하에는 AI-Hub 한국인 영어 발화 데이터셋을 추가 학습에 활용한다. **② 피드백 개선** - 탐지된 오류 음소와 유형(대치·삭제·삽입)을 LLM 에 전달해 한국어 음소 체계 기반 교정 가이드를 생성한다. [조선시대 영어교재 아학편] 방식의 프롬프트 엔지니어링으로 IPA 없이 즉시 따라 할 수 있는 안내를 제공하며, 구강 단면 이미지 시각화와 최소대립쌍 훈련 콘텐츠도 함께 제공한다. **③ 범용성 개선** - 학습 목적별 스크립트 선택 또는 직접 입력 기능을 구현한다		
달성 목표	L2-ARCTIC 기준 성능 평가와 소프트웨어 파이프라인 정상 동작 시, 성공 기준: ① FRR을 4%대로 감소시키되 MDD F1(현재 0.6078) 하락 없이 유지. ② LLM 피드백의 음성학적 타당성을 자체 벤치마크로 평가. ③ 단일 발화(~5 초) 기준 end-to-end 추론 - CPU 환경에서 실시간 응답 속도 **소프트웨어 테스트**: 모듈 단위 테스트 / 통합 테스트(오디오 입력→피드백 출력 전체 파이프라인) / API 응답코드·속도·동시처리 안정성 검증.		
기타	**역할분담** - 음소 탐지 모델 학습 및 실험: 신영환 / 백엔드 API 서버와 오디오 처리: 김지용 / 프론트엔드 인터페이스와 피드백 시각화: 이찬영, 최현지 / QA와 데이터 처리: 김형준, 백지영		