



이력서

## 기채운(Chaewoon Ki)

전화번호 010-6817-0826

이메일 [befilledwith@kaist.ac.kr](mailto:befilledwith@kaist.ac.kr)

GitHub <https://github.com/71c1nw00n>

학력사항 EDUCATION	2022.03 – 현재	한국과학기술원(KAIST) 전산학부 학사 GPA: 3.74/4.3 (94.4) 주요 수강 과목: 기계학습, 파이썬을 통한 자연어처리, 인공지능개론, 3 차원 데이터를 위한 기계학습
	2019.03 – 2022.02	한성과학고등학교
경력사항 EXPERIENCES	2025.11 – 현재	KAIST AI 대학원 <b>DAVIAN Lab (지도교수: 주재걸)</b> LLM 기반 Personalized Web Agent 연구 (공저자 논문 작성)
	2024.07 – 2025.01	KAIST 바이오및뇌공학과 <b>SynBi Lab (지도교수: 이관수)</b> 신약 후보 물질 발굴 및 단백질 활성 예측 (연구 참여)
	2023.12–2024.07	<b>Pebblous 데이터사이언스팀 연구인턴</b> 배터리 SOH 예측, ML/DL 비교 실험 및 리서치 페이퍼 작성
프로젝트 SIDE PROJECTS	개인 프로젝트 (3 인)	<b>POPO: 대학생 개발자를 위한 포트폴리오 생성 플랫폼</b> 백엔드 개발, LLM, RAG, Fine-tuning
	LG 전자 산학협력	<b>FossLight Dataset Scanner</b> LLM, 재귀적 의존성 탐색, 평가 데이터셋 구축
	카카오임팩트 산학협력	<b>FitFor: 휠체어 이용자를 위한 운동 추천 플랫폼</b> LLM, 사용자 인터뷰를 통한 문제 재정의
	개인 프로젝트 (6 인)	<b>Asymmetric Encoder-Decoder 기반 음성 분리</b> PyTorch, 데이터 증강, 베이스라인 대비 지표 개선
기술사항 SKILLS	프로그래밍 언어	Python
	AI/ML & Agent	LLM Web Agent, RAG, Fine-tuning, Agent memory systems, Context Engineering
	프레임워크	PyTorch, FastAPI, Playwright, Vector DB
	언어	한국어, 영어 (TOEIC 960, OPIc AL)
수상/장학 AWARDS	수상	공과대학 Dean's List
	장학금	KAIST 성적 우수 장학금 (전액 장학) 임미숙 장학금 KIAT 한미 첨단분야 장학금

# 경력기술서

LLM 을 비롯한 AI 가 사람처럼 언어를 이해하고, 맥락을 기억하며, 적절한 시점에 정보를 활용하도록 설계하는 것에 관심이 있습니다. Web Agent, AI 신약개발, 배터리 SOH 예측 등 다양한 도메인에서 AI 를 활용해 문제를 해결한 경험이 있습니다. AI 가 단순히 언어를 처리하는 것을 넘어 맥락을 이해하고 기억하는 시스템을 만드는 엔지니어가 되고 싶습니다.

## 연구 및 인턴

KAIST 김재철 AI대학원  
DAVIAN Lab

학부연구생  
(지도교수: 주재걸)  
2025.11 - 현재

### Personalized Web Agent 연구 (공동저자, 논문 준비 중)

#### Description

사용자의 웹사이트 도메인 선호와 세부 조건이 명시되지 않은 모호한 쿼리가 주어졌을 때 사용자의 Implicit Preference 를 추론해 개인화된 웹 태스크를 수행하는 Web Agent 설계 및 벤치마크 구축

#### What I Did

- 메모리 시스템 설계: Web 탐색 히스토리를 task 단위로 정제하는 Factual Memory 와 반복된 행동 패턴에서 선호를 추출하는 Preference Memory 로 분리, 단기 행동 로그와 장기적 선호 패턴을 계층적으로 관리
- Reranking 기반 Context 최적화: Preference Memory 전체 주입시 context 초과하는 문제를 해결하기 위해 Embedding 기반 Top-N 검색 후 LLM 으로 관련성을 재판단하는 Reranking 기능 구현
- 벤치마크 설계: top-1M website 수집 및 기능 기준 분류, Personalized Web Task 정의, Implicit Preference 추론 능력을 정량 평가하는 지표 설계

#### Tech Stack

Python, Playwright, Pydantic, Anthropic Claude API, Embedding 기반 Retrieval

KAIST 바이오및뇌공학과  
Synthetic Bio Lab

학부연구생  
(지도교수: 이관수)  
2024.07 - 2025.02

### AI 기반 신약 후보 물질 발굴 연구 (연구 보조)

#### Description

단백질-약물 결합 예측 정확도 향상을 목표로 AI 기반 신약 후보 물질 발굴 연구에 보조로 참여. Fragment-based Drug Generation 모델의 분자 표현 품질 개선을 위한 GNN 기반 사전 학습 구현

#### What I Did

- GNN 기반 단백질 사전학습 구현: 단백질 구조의 원자-결합 관계를 그래프로 모델링하여 fragment-based drug generation 모델에 적용. 기존 sequence 기반 표현 대비 구조적 정보를 직접 반영하고자 GNN 구조 선택
- GAE 를 활용한 단백질 데이터 압축: 고차원 단백질 그래프 데이터를 차원 축소 및 정보 압축해서 학습 데이터로 사용하고 Graph Autoencoder 로 인코딩하여 핵심 구조 정보만 추출
- GPCR 기반 활성 예측 모델 탐색: Hierarchical Graph Transformer 기반 단백질 활성 예측 모델을 검토하고, orphan GPCR 예측으로의 확장 가능성을 일부 구현하며 탐색

#### Tech Stack

Python, PyTorch Geometric, GNN, GAE, Graph Transformer, DiffDock

## 리튬이온 배터리 SOH 예측 모델 구현 및 리서치 페이퍼 작성

### Description

AI 데이터 스타트업 Pebblous 에서 연구 인턴으로 근무하며 리튬이온 배터리의 잔여 수명(SOH) 예측을 위한 ML/DL 알고리즘을 비교 실험하고 리서치 페이퍼 작성. 데이터 기반 수명 예측 분야에서 유효한 알고리즘 검증

### What I Did

- 공개 데이터셋 8 종 수집 및 전처리: NASA PCoE, CALCE CS2 등 주요 벤치마크 데이터셋을 수집하고 EDA 를 통해 각 데이터셋의 특성과 변수 분포를 분석. 여러 데이터셋 환경에서 일관되게 동작하는 통합 전처리 파이프라인 구현
- ML/DL 알고리즘 비교 실험 및 인사이트 도출: LSTM, GRU, RNN 등 신경망 모델과 LightGBM, XGBoost 등 트리 기반 모델을 비교 실험. PyCaret AutoML 로 탐색한 결과 LightGBM 이 가장 낮은 RMSE 를 기록하였으며, DL 모델(LSTM/GRU)에서 shifting 문제(직전 값을 그대로 예측하는 허위 정확도)를 확인 — DL 대비 ML 알고리즘이 배터리 SOH 예측 태스크에서 더 안정적임을 분석, 추가로 모델 성능이 알고리즘 종류보다 추출하는 feature 에 따라 더 크게 변동함을 확인.
- 리서치 페이퍼 작성: 위 분석 내용을 바탕으로 리서치 페이퍼 작성

### Tech Stack

Python, PyCaret, LightGBM, XGBoost, LSTM, GRU, RNN

## 프로젝트

### POPO

개인 프로젝트  
(팀 3 인, 백엔드 및 고도화)  
2025 봄학기  
2026 봄학기



Front



Back

### 대학생 개발자를 위한 포트폴리오 생성 플랫폼

#### Description

대학생 개발자가 프로젝트 정보를 입력하면 LLM 이 포트폴리오를 자동 생성해주는 웹 플랫폼. 사용자 입력이 짧거나 모호한 상황에서도 일관된 품질의 결과물을 생성하도록 백엔드 구현. 현재 Fine-tuning 이용한 고도화 진행 중

#### What I Did

##### [ 백엔드 구현 | 2025 봄학기 ]

- LLM 포트폴리오 생성 파이프라인 구현: 사용자 입력을 구조화된 컨텍스트로 변환한 뒤 OpenAI API 에 전달하는 파이프라인 설계. 입력이 짧거나 모호할 때 발생하는 hallucination 을 줄이기 위해 입력 정보를 정형화해 프롬프트 주입하는 방식으로 일관성 확보
- FastAPI 백엔드 및 DB 설계: API 서버 아키텍처 설계 및 DB 스키마 구성, OpenAI API 연동 등 전반적인 백엔드 구현

##### [ 고도화 | 2026 봄학기 (진행 중) ]

- Llama 3.1 LoRA Fine-tuning: 범용 LLM 의 포트폴리오 특화 표현이 부족한 문제를 해결하기 위해 포트폴리오 데이터를 수집하여 Llama 3.1 기반 LoRA Fine-tuning 진행. 도메인 특화 표현 품질 개선이 목표
- RAG 파이프라인 구축: 우수 포트폴리오 사례를 Vector DB 에 저장하고 유사 사례를 검색해 생성 시 참조하는 RAG 파이프라인 구축. 프롬프트만으로는 커버하기 어려운 구체적인 표현 품질을 레퍼런스 기반으로 보완하는 것을 목표로 구현 중

#### Tech Stack

FastAPI, Next.js, OpenAI API, Llama 3.1, LoRA, RAG, Vector DB

## FOSSLight Scanner

LG 전자 산학협력 프로젝트  
(팀 2 인)

2025 가을학기



Code



PDF

## 오픈소스 모델/데이터셋의 라이선스 및 의존성을 자동 분석하는 도구 개발

### Description

AI 개발 시 사용하는 오픈소스 모델·데이터셋의 라이선스를 수동으로 검증하는 비효율을 해소하기 위해, Hugging Face/GitHub URL 입력만으로 의존성을 재귀적으로 탐색하고 라이선스를 자동 분석하는 CLI 도구 개발. README 분석기 구현 및 평가 체계 구축 담당.

### What I Did

- LLM 기반 README 분석 모듈 구현: 비정형·다형식 README 에서 신뢰성 있는 의존성 추출을 위해 정규식 대신 LLM 기반 분석 모듈을 채택, 참조 데이터셋·모델을 식별하여 라이선스·의존성 정보를 종합 제공
- 재귀적 의존성 탐색 파이프라인 설계: URL 입력 시 README 분석과 링크 추적으로 전체 의존성을 재귀적으로 탐색하는 파이프라인 설계. 복잡하게 중첩된 오픈소스 의존성을 한 번에 추적할 수 있도록 구조화
- 평가 체계 구축: JSON 테스트케이스 50 개를 직접 구축하여 성능 검증 (Accuracy 95.24%, Precision 100%, Recall 95.24%, F1 Score 97.56%)

### Tech Stack

Python, LLM, HuggingFace API, GitHubAPI

## Fitfor

카카오임팩트 주관 프로젝트  
(팀 7 인, 기획 및 백엔드 개발)

2024 가을학기



Front



Back



PDF



Demo

## 휠체어 이용자를 위한 운동 추천 플랫폼

🏆 카카오임팩트 주관 <전산학특강: 테크포임팩트> 공감인기상 수상

### Description

운동을 처음 시작하는 휠체어 이용자가 자신에게 맞는 운동 종목을 파악하기 어렵다는 문제를 위해, 신체 능력 평가 기반으로 맞춤 운동을 추천하는 웹 플랫폼 개발, 카카오임팩트 주관 <테크포임팩트> 수업에서 장애인 스타트업 캠퍼스즈의 멘토링 하에 진행

### What I Did

- 사용자 인터뷰 기반 문제 재정의: 기획 초기안(운동 동호회와 장소 탐색)을 장애인 사용자 직접 인터뷰로 검증한 결과, 실제 니즈는 동호회 탐색보다 “자신이 가능한 운동 종목 파악”임을 확인, 동일 병명이라도 개인별 운동 가능 범위가 상이함을 근거로 LLM 기반 적응형 운동 추천으로 방향 전환
- LLM 적응형 추천 파이프라인 구현: 실제 장애인 운동 대회에서 사용하는 동작 능력 평가 데이터 수집, 컨텍스트로 구조화해 LLM 에 제공하는 적응형 추천 파이프라인 설계 및 구현해 일반 LLM 의 추천 한계 보완
- 동호회 데이터 크롤링: 대한장애인체육회 생활체육정보센터에 등록된 동호회 정보를 크롤링하여 추천 결과와 연계
- 백엔드 개발: FastAPI 백엔드, DB 설계, API 연동을 맡아 OpenAI API 와 연동

### Tech Stack

FastAPI, Next.js, OpenAI API, Python

## 그 외 프로젝트

---

### Separate and Reconstruct

(팀 4 인, 모델 구현)

2024 가을학기



Code



Demo



PDF

### Asymmetric Encoder-Decoder for Speech Separation

베이스라인 대비 SI-SNRi, SDRi 성능 향상, 인공지능개론 팀프로젝트 2 위

- 인간 음성에만 최적화된 기존 음성 분리 모델의 한계를 극복하고자, 인간-동물 혼합 오디오 환경까지 커버하는 Asymmetric Encoder-Decoder 기반 음성 분리 모델 개발
- 연산 효율을 유지하면서 모든 베이스라인 모델을 상회하는 분리 정확도 달성 (SI-SNRi +2.3dB, SDRi +1.96dB)
- 기술 스택: PyTorch, Speech Processing, Deep Learning

### AI 생성 음성 판별 모델

(팀 6 인)

2024 가을학기

실제 인간의 음성과 생성된 가짜 음성을 구별하는 AI 모델

- Res2Net 아키텍처 기반으로 실제 음성과 AI 생성 음성을 판별하는 모델 개발
- KAIST 대표로 선발되어 SW 중심대학 디지털 경진대회 출전
- 기술 스택: PyTorch, Res2Net, Deep Learning

---

## 학업 외 활동

---

- KAIST 새내기지도그룹 프락터 (2023 봄학기-2024 가을학기)
- KAIST 생활관자치회 아름관 동장 (2024 가을학기-2025 봄학기)
- ICISTS-KAIST (2022 봄학기-2023 가을학기, 과학기술과 사회를 연결하는 컨퍼런스 개최 동아리)