PORTFOLIO

예측을 넘어 **신뢰**할 수 있는 답을 설계하는
데이터 사이언티스트 김준희입니다.

김준희
kim8881472@gmail.com
+82 10-8543-0390

Kim Junhee

김준희



Birth. 2001.02.09

E-mail. kim8881472@gmail.com

Tel. 010.8543.0390

Web. https://github.com/joon0390

불확실한 데이터 속에서 명확한 의사결정의 근거를 설계하는 분석가

1. 통계적 엄밀성과 신뢰도 확보
단순 집계를 넘어 데이터의 노이즈와 불확실성을 검증하고, 왜곡을 최소화한 인사이트를 도출합니다.

2. 현업 중심의 가설 설계 및 문제 정의
비즈니스 맥락을 파악해 풀고자 하는 문제를 명확히 정의하고, 검증 가능한 가설로 구체화합니다.

3. 실행으로 이어지는 지표 전환
복잡한 분석 과정 및 결과를 이해관계자가 즉시 활용할 수 있는 직관적인 언어와 계획으로 전달합니다.

수상 및 성과

2026.09	2026 인하 인공지능 챌린지 우수상
2025.08	2025 인하 인공지능 챌린지 최우수상
2025.04	LG Aimers 6기 최종 오프라인 해커톤 7위
2023.12	가을 VIP 발표대회 동상
2023.08	한국 ITS 학회 추계학술대회 우수논문상
2019~2025	총 6개 학기 우등생 선정

보유 기술

Language	Python, R, SQL
ML/DL Framework	Pytorch, Tensorflow, Keras, PyMC
Etc	Git, Notion, Docker, Pandas, Scikit-Learn, Airflow

CAREER SUMMARY

원천 데이터 전-처리 및 통계적 모델링까지, 불확실한 데이터를 신뢰도 높은 지표로 전환한 End-to-End 분석 경험

학사 과정 (엔지니어링 & 응용)

산업용 회전 설비 이상 진동 탐지

Arduino 및 Edge Impulse를 활용한 엣지 컴퓨팅 기반 실시간 예지보전 시스템 End-To-End 구현

대조학습 기반 주행 패턴 이상 탐지

레이블이 없는 비정형 스마트폰 센서 시계열 가속도 데이터를 활용한 자기지도학습 기반의 이상 탐지 모델 제안

강화학습을 이용한 조난자 이동 경로 예측

복잡한 공간 데이터를 강화학습 환경으로 구축하고 이에 따른 맞춤형 보상 함수를 설계한 Pointer Network DQN 에이전트 개발(특허 출원)

석사 과정 (고급 통계 & 최적화)

One-Class Classification using Bayesian Optimization

정상 데이터만 존재하는 불균형 환경을 타겟으로, 베이지안 최적화 기반의 자동화된 이상 탐지 구조 설계(SCI급 투고)

베이지안 딥 커널 머신을 이용한 성장 예측

DNN의 비선형 특징 추출과 가우시안 프로세스 회귀를 결합하여 예측의 오차를 최소화 및 신뢰 구간 정량화 달성

GAT 임베딩 기반 베이지안 커널 릿지를 활용한 분자 예측

그래프 어텐션(GAT)으로 추출한 구조 임베딩과 베이지안 커널 릿지를 결합하여 예측 성능 극대화 및 불확실성 정량화 달성(국제학회 발표)

도메인 지식 결합 복합 커널 가우시안 프로세스 기반 분자 독성 예측

분자 도메인 특성을 반영한 복합 커널과 가우시안 프로세스 분류기를 결합하여, 독성 예측 성능 향상 및 원칙적인 불확실성 정량화 달성(SCI급 투고)

핵심 역량

정교한 머신러닝 모델링(ML)과 비즈니스 지표 설계(DA)로 공간 데이터의 가치를 완성합니다.

CV Link



시공간 궤적 모델링 및 행동 패턴 분석

[역량]
- 시공간 이동 경로 모델링 경험
- 행동 패턴 및 변화점 탐지 경험

[Evidence]
- 이동 경로 예측 특허 및 학회 포스터 발표
- 시계열 패턴 탐지 논문 투고 및 우수논문상 수상



원천 센서 노이즈 제어 및 확률론적 불확실성 정량화

[역량]
- 센서 로그 노이즈 보정 경험
- 베이지안 불확실성 정량화

[Evidence]
- Bayesian ML 논문 게재 및 투고
- 엣지 센서 파이프라인 구축 프로젝트 경험



도메인 특화 KPI 설계 및 데이터 시각화·제품화

[역량]
- 비즈니스 가설 기반 지표 정의
- 직관적인 대시보드 템플릿 구축

[Evidence]
- LG Aimers 7위(Feature Engineering)
- 데이터 시각화 / 회귀분석실습 TA 수행
- Streamlit을 사용한 이동경로 모델링 대시보드 구축

PROJECTS

산업용 회전 설비 이상 진동 탐지 프로젝트 01

회전 설비의 미세 진동 데이터를 엣지 디바이스를 통해 실시간으로 수집, 분석하는 시스템을 구축

One-Class Classification Using Bayesian Optimization 02

정상 데이터만을 가지고 비정상을 분류해내는 OCC 모델과 베이지안 최적화에 대한 연구

베이지안 딥 커널 머신을 이용한 양식 넙치 성장 예측 03

딥러닝 비선형 표현학습과 베이지안 커널 머신의 불확실성 정량화를 결합한 BDKMR 모델을 제안 및 검증

대조학습을 이용한 주행 패턴 이상 탐지 04

라벨이 부족한 운전 패턴 데이터를 활용하여 대조학습 기반의 이상탐지 시스템을 설계

강화학습 기반 공간 이동 경로 예측 시스템 05

산악 조난자의 시간에 따른 이동 경로를 강화학습을 이용해 모델링하여 예측

01. 산업용 회전 설비 이상 진동 탐지 프로젝트

Edge Computing 기말 프로젝트 | 2023.09~2023.12

Edge AI | Anomaly Detection | Arduino | End-To-End Pipeline | Classification

1. 프로젝트 목표 (GOAL)

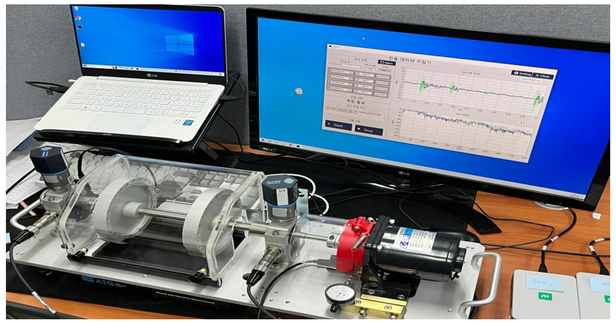
회전 설비에서 발생하는 미세 진동 데이터를 엣지 디바이스(Edge Device) 단에서 실시간으로 수집 및 분석하여, 고장을 예측하고 분류하는 시스템을 구축

2. 문제 상황(Problem)

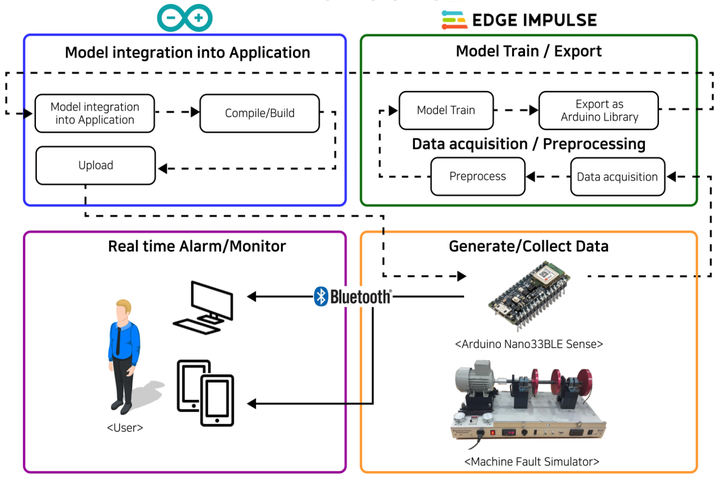
1. 기존 클라우드 전송 방식은 통신 지연으로 인해 실시간 이상 대응이 어렵고 보안의 우려가 존재함. 또한, 설비마다 진동의 패턴이 다르기에 일관된 규칙 기반의 탐지가 불가능하여 현장 맞춤형 경량화 AI 모델이 필수적임

2. 모사 설비에 센서를 부착하는 데에 물리적 한계가 존재하여 전용 케이스를 제작하는 것이 필요함

회전 모사 설비에서 데이터를 수집하는 모습



프로젝트 파이프라인



3. 접근 방식 및 해결(Solution)

1. 하드웨어 한계 돌파(Agile): Arduino Nano 33 BLE 도입 후 설비 맞춤형 케이스가 부재하자, 단기간에 3D 모델링을 독학하여 케이스를 직접 설계 및 제작해내는 실행력 발휘.

2. 엣지 AI 파이프라인 구축: Edge Impulse 플랫폼을 활용하여 진동 시계열 데이터를 전처리하고, 마이크로 컨트롤러 환경에 최적화된 경량화 딥러닝 모델 학습.

3. 실시간 검증 완료: 구축된 모델을 C++ 라이브러리로 변환하여 Arduino에 배포(Upload), 정확도 96.8%로 3가지 설비 상태(정상/불균형/축틀어짐) 실시간 분류 성공.

4. 경량화 모델링 및 파이프라인 상세

1. Spectral Analysis(FFT) 전처리
   - 3축 가속도 시계열 데이터를 고속 푸리에 변환하여 주파수 도메인으로 치환, RMS, Skewness, Kurtosis, 주요 주파수 Peak 에너지를 추출해 물리적 진동 특성을 반영하여 데이터 차원 대폭 축소
2. 1D-CNN 기반 경량화 신경망 설계 및 양자화
   - 주파수 특징맵 기반의 경량화 1D-CNN 아키텍쳐 설계. MCU 리소스 한계(RAM 256KB) 극복을 위해 INT 8 양자화를 수행하여 분류 성능 유지 및 모델 용량 압축
3. C++ 라이브러리 포팅 및 On-Device 추론(정확도 96.8%)
   - 학습모델을 C++ 라이브러리로 Export하여 Arduino IDE에 통합. 클라우드 통신 없이 디바이스 자체 연산 만으로 높은 정확도의 실시간 이상 탐지 달성

5. 기여할 수 있는 역량

1. 대용량 데이터 환경을 고려한 경량 분석 파이프라인 구축
   - 결측과 노이즈가 존재하는 원천 데이터에 전처리 및 모델 경량화 기법을 적용하여, 대규모 데이터가 지속적으로 유입되는 환경에서도 안정적으로 작동하는 저지연 분석·모니터링 체계 구축
2. 신규 과제의 빠른 End-to-End PoC 수행
   - 데이터 수집과 전처리부터 모델링, 검증 및 배포까지 전 과정을 수행한 경험을 바탕으로, 새로운 과제의 기술적 가능성과 실효성을 신속하게 검증하고 실제 적용까지의 시간을 단축

02. One-Class Classification Using Bayesian Optimization

논문 연구 | Computational Statistics (Under Revision)

One-Class Classification | Bayesian Optimization | Anomaly Detection | Classification

1. 프로젝트 목표 (GOAL)

불량(비정상)데이터가 극도로 부족한 환경에서, 정상 데이터만을 활용하여 이상 패턴을 탐지하는 One-Class Classification(OCC) 모델의 성능과 탐색 효율성 극대화

2. 문제 상황(Problem)

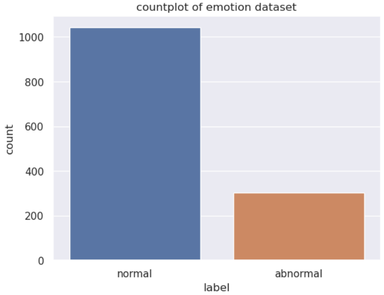
1. 제조 현장 등에서는 불량(Abnormal) 데이터가 희귀하여 일반적인 이진 분류(Binary Classification) 적용이 불가함

2. 딥러닝(Deep-SVD) 및 복잡한 커널 모델(OC-SVM)은 한 번 학습하고 평가하는 데 큰 컴퓨팅 자원과 시간이 소요됨. 기존의 Grid Search나 Random Search를 무작정 적용할 경우 탐색 비용이 기하급수적으로 증가함

3. 접근 방식 및 해결(Solution)

1. Bayesian Optimization(BO) 도입
   - Gaussian Process(GP) 기반의 Surrogate 모델을 도입하여, 불필요한 모델 학습을 최소화하고 '가장 정답일 확률이 높은 구간(UCB)'만 똑똑하게 탐색하는 최적화 프레임워크 적용

2. 검증 및 성능 달성
   - 가상 데이터 및 극도 불균형 데이터(24:1 산불 데이터) 등을 대상으로 Grid Search, Random Search, Hyperband 등과 벤치마킹 수행
   - Random Search나 Grid Search 대비 가장 적은 횟수의 모델 평가(평가 시간 단축) 만으로도 최고 수준의 F1-Score 및 AUC를 달성하여 비용 대비 압도적인 효율성을 입증함

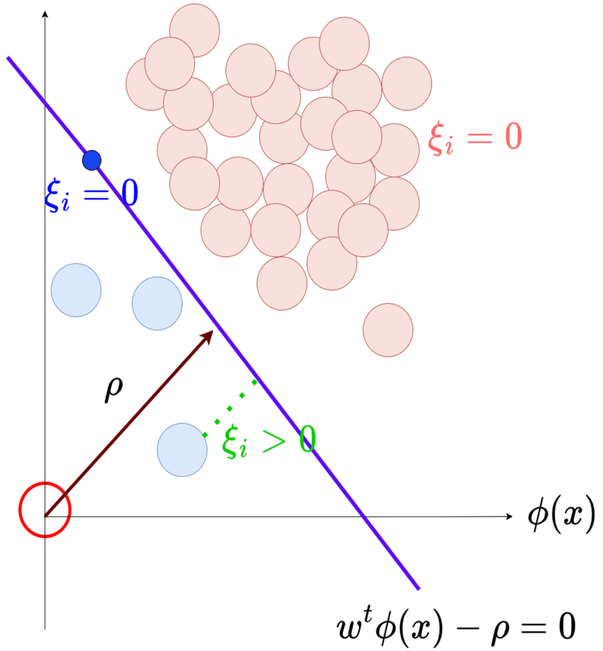


불균형 데이터

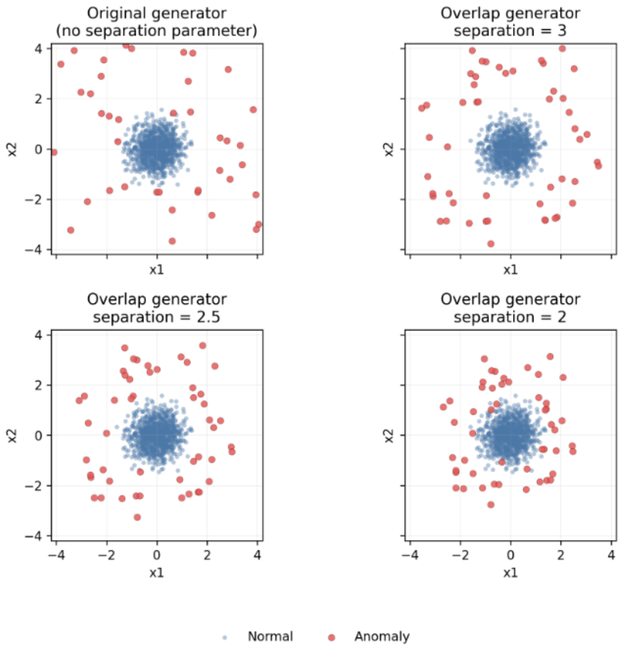
Metric	GC	RC	BO	HB	BOHB
Recall	0.4800	0.4350	0.4700	0.6200	0.6650
F-1	0.6142	0.5592	0.6569	0.4563	0.4155
AUC	0.6929	0.6407	0.7050	0.6407	0.6224

BO를 적용한 방법의 F1-Score와 AUC가 우수한 모습

OC-SVM



클래스 불균형 시뮬레이션 데이터 예시



4. 기여할 수 있는 역량

1. 정상 데이터 기반의 이상 패턴 탐지
   - 비정상 사례와 레이블이 부족한 환경에서 정상 데이터의 공통 패턴을 학습하고, 정상 범위를 벗어나는 이상 징후를 안정적으로 식별하는 단일 클래스 분류(OCC) 파이프라인 구
2. 베이지안 최적화를 통한 모델 튜닝 및 인프라 리소스 절감
   - 평가 비용이 높은 하이퍼파라미터를 제한된 실험 횟수로 최적화하여 모델 개발에 필요한 연산 자원과 탐색 시간을 절감하고, 새로운 데이터와 과제에 대한 모델 적용 주기를 단축

03. 베이지안 딥 커널 머신을 이용한 성장 예측

Growth Prediction | Deep Kernel Machine | Gaussian Process | Uncertainty Quantification

논문 연구 | Predicting Flatfish Growth in Aquaculture Using Bayesian Deep Kernel Machines, Applied Sciences

1. 프로젝트 목표 (GOAL)

1. 온도, 용존 산소량 등 통제하기 어려운 다수의 외부 환경 센서 변수가 타겟에 미치는 영향을 분석하는 하이브리드(DNN + 통계) 예측 모델 구축
2. 단순 점 예측을 넘어 예측 결과의 불확실성(Uncertainty)을 수학적으로 정량화하여, 리스크가 큰 상황에서도 선제적 대응이 가능한 의사결정 프레임워크 제안

2. 문제 상황(Problem)

1. 다중 센서에서 수집되는 환경 변수들은 복잡하고 비선형적인 상호작용을 보이므로, 기존 선형 모델이나 기존 커널 머신(KRR, BKMR)으로는 데이터의 특징(Feature)을 온전히 추출하기 어려움
2. 반면 딥러닝(DNN)을 단독으로 사용할 경우 비선형 패턴은 잘 찾지만, 모델이 내놓은 결과값을 얼마나 믿을 수 있는 지 알 수 없는 블랙박스(Black-Box) 특성 탓에 보수적인 현장 도입이 까다로움

3. 접근 방식 및 해결(Solution)

1. 하이브리드 BDKMR 모델 제안
   - 딥러닝의 뛰어난 "특징 추출(Feature Extraction)" 역량과 가우시안 프로세스(Gaussian Process) 기반의 "불확실성 정량화 능력"을 결합한 BDKMR 모델 아키텍처 설계
2. 복합 센서 데이터 기반 성능 검증
   - 장시간 누적된 대규모 환경 센서 및 생장 데이터를 활용하여 모델 검증 수행
3. 통계적 교차 검증(LOOCV) 결과, 기존 베이스라인 모델 대비 예측 오차(MAE 0.1895)를 대폭 개선하였으며 강건한 신뢰구간 산출에 성공함

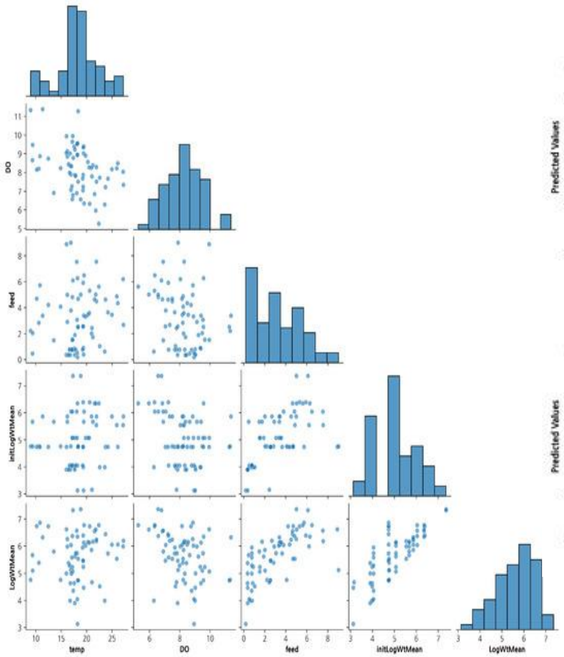
Model	MAE	MSE
KRR	1.1141	3.5665
BKMR	0.6977	0.9447
BDKMR (Equal)	0.2006	0.0721
BDKMR	**0.1895**	**0.0629**

벤치마킹 모델 간 성능 비교(MAE, MSE)

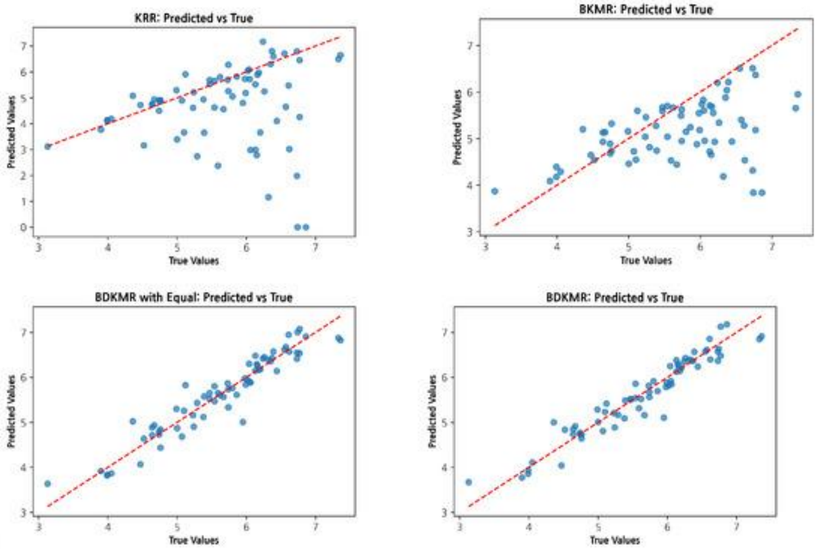
Model	Interpretability	Interaction Modeling	Uncertainty	Nonlinearity	Comp. Cost	Scalability
BKMR	Moderate	Strong	Yes	Strong	High	Moderate
BDKMR (ours)	Moderate	Strong	Yes	Very Strong	High	Moderate
Random Forest	Moderate	Limited	No	Moderate	Low	High
SVM (RBF kernel)	Low	Limited	No	Strong	Moderate	High

타 비선형 머신러닝 모델들과의 비교

모델에서 사용한 주요 변수들 간의 산점도



벤치마킹 모델 별 예측값 대비 실제값 그림



4. 기여할 수 있는 역량

1. 신뢰도 높은 의사결정을 위한 예측 불확실성 정량화
   - 단순한 점 예측을 넘어 결과의 불확실성을 함께 제시하여, 신뢰도가 낮거나 위험이 큰 구간을 사전에 식별하고 안정적인 의사결정을 지원
2. 다변량 데이터셋에서의 비선형 관계 모델링
   - 딥러닝과 베이지안 커널 기법을 활용하여 변수 간 비선형 관계와 상호작용을 분석하고, 높은 예측 성능과 결과에 대한 해석 가능성을 함께 확보

04. 대조학습 기반 주행 패턴 이상 탐지

논문 연구 | 2023 한국 ITS 학회 추계 학술대회 - 우수 논문상

Contrastive Learning | Self-Supervised Learning | Anomaly Detection | Time-Series

1. 프로젝트 목표 (GOAL)

운전자의 스마트폰 3축 가속도 센서를 분석하여, 급가속/급정거/급회전 등 이상 주행 패턴과 노면 상태의 변화 등을 정확하게 탐지하는 모델 구축

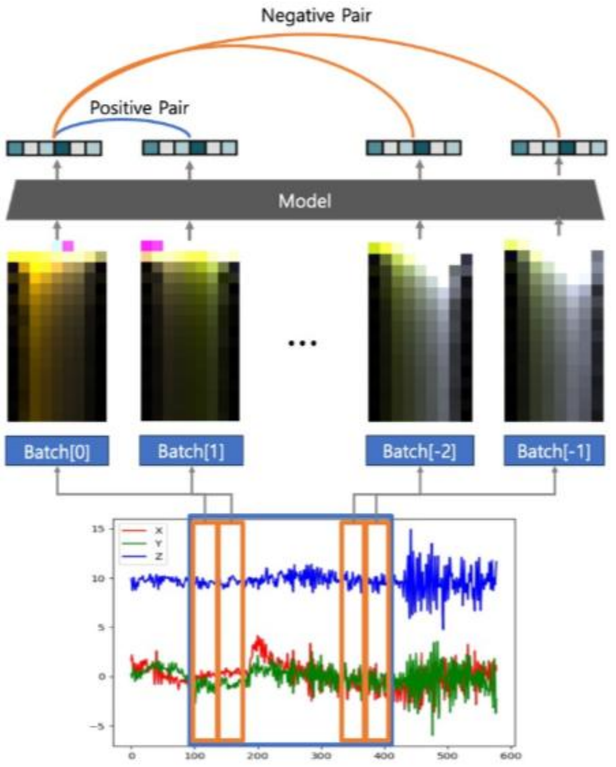
2. 문제 상황(Problem)

1. 기존의 지도학습 기반 딥러닝 방식은 대량의 레이블 데이터가 요구됨
2. 데이터에 운전 패턴 레이블을 만드는 어노테이션 작업은 시간적, 인적 비용이 크게 요구되어 실제로 레이블링된 데이터는 소규모로 존재함
3. 학습된 차량과 실제 적용 차량이 달라지는 도메인 변화 상황에서 모델의 일반적인 성능을 유지하기 어려움

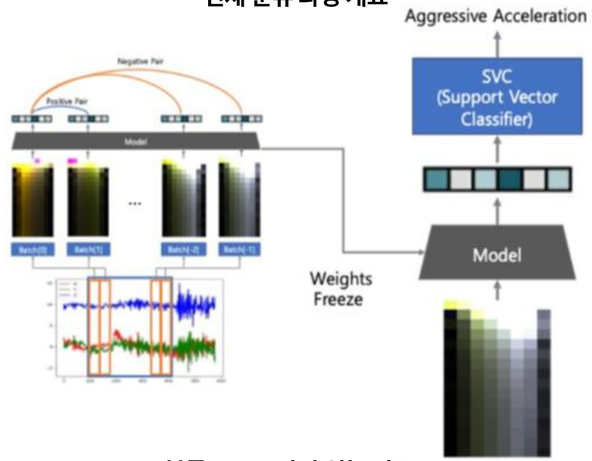
3. 접근 방식 및 해결(Solution)

1. 시계열 데이터 전-처리
   - 스마트폰 가속도 데이터를 작은 시간 단위로 나누고 단시간 푸리에 변환(STFT)을 적용하여 신호의 시간-주파수 해상도를 확보함
2. 시간 일관성 가정 기반 대조학습
   - 인접한 시간의 데이터는 특징의 변화가 크지 않다는 시간 일관성(Time Consistency)을 가정하여, 긍정 쌍(Positive Pair)과 부정 쌍(Negative Pair)을 구성하고 대조학습(Contrastive Learning)을 수행함
3. 적은 레이블의 한계 극복
   - 대조학습으로 사전 학습된 모델을 특징 추출기로 사용하고, SVC(Support Vector Classifier)를 결합하여 클래스별 30개 내외의 아주 적은 레이블 데이터만으로도 분류 학습을 완료함
4. 성능 달성 및 일반화 검증
   - 추후 차량 탑재 가능성을 고려하여 SqueezeNet, RegNet 등 경량화 딥러닝 모델을 적용하여 정밀도 1.0, 재현율 1.0의 높은 성능을 재현하였으며, 학습과 검증에 서로 다른 차량 데이터를 사용하여 도메인 변화에 대한 강건성을 입증함

Time Consistency 가정 하에서의 학습 개요



전체 분류 과정 개요



분류 Head 간의 성능 비교표

Model Name	Change Point Detection Result		Classification Result	
	Threshold	Precision	Precision	Recall
SqueezeNet	0.0	0	1.0	1.0
	0.4	0.75		
	0.6	0.69		
ShuffleNet	0.0	0	0.886	0.869
	0.4	0.73		
	0.6	0.61		
RegNet	0.0	1	0.956	0.951
	0.4	0.92		
	0.6	0.72		
MobileNet	0.0	0	0.039	0.197
	0.4	0.69		
	0.6	0.60		
EfficientNet	0.0	0	0.26	0.34
	0.4	0.67		
	0.6	0.76		
MnasNet	0.0	0	0.17	0.25
	0.4	0		
	0.6	0		

4. 기여할 수 있는 역량

1. 대규모 비정형 로그 기반 이상 징후 및 변화점 탐지
   - 레이블이 없는 대규모 시계열·행동 로그에 자기지도학습을 적용하여, 정상 패턴에서 벗어나는 이상 징후와 급격한 상태 변화를 조기에 탐지하는 파이프라인 구현
2. 환경 및 도메인 변화에 강건한 데이터 표현 구축
   - 서비스 운영 환경에 따라 데이터 분포가 달라지는 상황에서도 일반화 성능을 유지할 수 있도록, 공통 패턴과 핵심 특성을 반영한 강건한 표현(Representation) 모델링

05. 강화학습 기반 공간 이동 경로 예측 시스템

특허 출원 (10-2024-0185328) | The 12th IASC-ARS Conference 포스터 발표 (대만, 2024)

Reinforcement Learning | Spatial-Temporal Modeling | Trajectory Prediction | Agent Simulation | Time-Series

1. 프로젝트 목표 (GOAL)

지형 및 환경적 제약 조건이 존재하는 복잡한 시공간 환경에서, 강화학습(RL) 기반 에이전트 모델링을 통해 대상자의 최적 이동 경로와 확산 궤적을 정밀하게 예측하는 시스템 구축

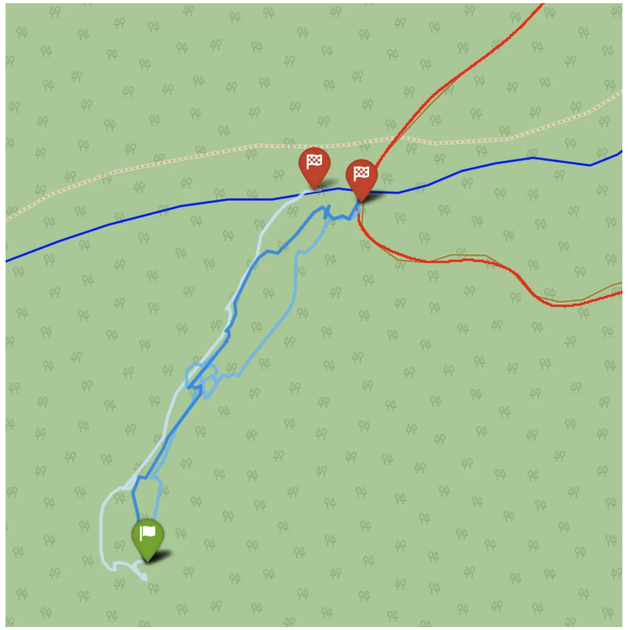
2. 문제 상황(Problem)

1. 극심한 시공간 불확실성: 실제 이동 데이터는 결측이 많고 관측이 불연속적 이어서 전통적인 통계적 보간법(Interpolation)만으로는 비선형적 이동 패턴을 예측하기 어려움
2. 환경적 제약 요인의 복합성: 고도, 경사도, 이동 장애물 등 공간의 물리적 구조가 에이전트의 이동 의사결정에 미치는 영향을 정량적으로 반영할 모델 부재
3. 순차적 의사결정 모델링의 한계: 단순 룰 기반(Heuristic) 접근법은 환경 변화에 따른 대상자의 목적 지향적 행동 메커니즘을 동적으로 시뮬레이션 하지 못함

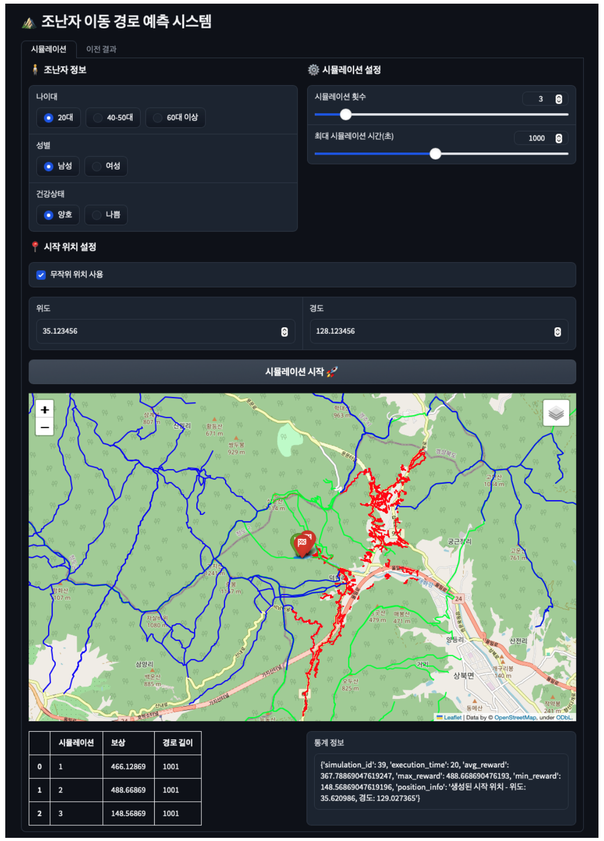
3. 접근 방식 및 해결(Solution)

1. 시공간 환경의 MDP(Markov Decision Process) 정식화
   - 지형 공간 좌표 및 이동 제약 요소를 상태(State) 공간으로 정의하고, 이동 가능한 방향과 속도를 행동(Action) 공간으로 구조화
2. 목적 지향적 보상 함수(Reward Function) 설계
   - 공간 내 장애물 회피, 경로 효율성, 이동 관성을 종합적으로 반영한 다중 목적 보상 체계를 구축하여 현실적인 이동 궤적 유도
3. 강화학습 기반 최적 이동 정책(Policy) 학습
   - 누적 보상을 극대화하는 에이전트 정책을 학습시켜 미지의 환경에서도 물리적 제약 조건을 준수하며 이동할 확률이 가장 높은 경로를 순차적으로 생성

조난자 에이전트의 이동경로



시뮬레이션 대시보드



4. 기여할 수 있는 역량

1. 시공간 데이터 기반 이동 경로 및 다음 행동 예측
   - 시공간 좌표와 행동 이력 데이터를 바탕으로 객체의 이동 궤적을 모델링하고, 향후 위치나 행동을 예측하는 경로 분석 알고리즘 개발
2. 공간 구조 및 자원 배치 최적화 시뮬레이션
   - 공간 구조와 자원 배치 변화에 따른 이동 흐름과 운영 효율을 사전에 시뮬레이션하여, 데이터 기반의 공간 설계 및 배치 의사결정 지원

OTHER PROJECTS & ACHIEVEMENTS

다양한 도메인과 방법론(생성 모델, 강화학습 및 그래프 모델)을 어우르는 기술 스펙트럼

1. 다채널 어레이 센서 기반 실시간 화학 작용제 분류 프로젝트(국방과학연구소(ADD) 협업)
- 센서 드리프트 및 스파이크 같은 노이즈를 보정한 전-처리 파이프라인과 독자적인 아키텍처를 설계하여, 다중 채널 환경에서 실시간 추론이 가능한 시제품을 구현함

2. BayesComp 2025 국제학회 포스터 발표(싱가포르 2025.06) – link
- 주제 : Bayesian Kernel Ridge on GAT Embeddings for Molecular Prediction
- 그래프 신경망(GAT)를 통해 분자 구조의 복잡한 네트워크 특징을 추출하고, 베이지안 회귀를 적용하여 물성 예측 및 불확실성 정량화를 동시에 달성

3. 언어정보 기반 흑백 이미지 색채화(2025 인하 인공지능 챌린지, 최우수상) - link
- 자연어 지시문을 바탕으로 흑백 이미지를 정밀하게 채색해내는 주제의 대회에 출전하여 좋은 성적을 거둠
- 컴퓨팅 자원을 고려하여 사전 학습된 Diffusion Model을 파인튜닝하고 출력 결과물을 사후 보정하는 방식의 전략을 세움

4. 도메인 지식 기반 복합 커널 GP 분자 독성 예측(Journal of Cheminformatics 투고 완료) - link
- 분자가 갖는 고유의 물리/화학적 특성을 가장 잘 설명하기 위한 복합 커널(Tanimoto Similarity기반의 Binary Kernel을 제안)을 설계하고 이를 가우시안 프로세스 회귀에 적용

5. LG Aimers 6기 : 의학 도메인 기반 난임 환자 임신 성공 확률 예측 프로젝트
- 난임 환자 시술 데이터를 기반으로 임신 성공 확률을 예측하여 의료진의 의사결정을 보조하고 복잡한 metric 평가 지표에 최적화된 머신러닝 파이프라인 구축 - **link**

6. 강화학습 기반 공간 이동 경로 예측 시스템
- 지형 및 환경적 제약 조건이 존재하는 복잡한 시공간 환경에서, 강화학습(RL) 기반 에이전트 모델링을 통해 대상자의 최적 이동 경로와 확산 궤적을 정밀하게 예측하는 시스템 구축 – **link**

7. 다수의 데이터 사이언스 경진대회 및 해커톤 참가 및 입상 - link

NOW, I AM DOING …

현재 집중하고 있는 연구(Current Research & Focus)

1. 구조 변화 예측을 위한 그래프 생성 모형(Graph Generative Models):
   - 분자에 방사선이 조사되었을 때 일어나는 동적인 구조 변화를 그래프 생성 모형으로 모델링하는 방법을 연구하고 있습니다.
   - ➢ 이러한 복잡한 네트워크 구조의 상태 변화 예측 경험은 특정한 변화를 AI를 통해 시뮬레이션 하는 데 강력하게 응용될 것이라고 생각합니다.

2. 강건한 추론을 위한 Predictive Score Posterior:
   - 모형 오지정 하에서의 사후 예측분포가 갖는 문제점을 해결하기 위해 사후 예측 분포를 직접 타겟으로 하여 임의의 Scoring Rule을 최적화하는 샘플링 기법을 연구하고 있습니다.

3. 자율적 의사결정을 위한 Agent AI (Self-Study):
   - 연구 이외에도 최근 급속도로 발전하고 있는 Agentic AI를 꾸준하게 공부를 진행하고 있습니다. 단순 예측을 넘어, 주어진 환경(World)를 스스로 인지하고 시뮬레이션을 통해 최적의 선택을 계획 및 수행하는 에이전트 기반 기술을 향후 어떻게 접목할 수 있을지 고민하고 있습니다.

AI ENGINEER로서의 자세 4가지

끊임없는 탐구

익숙한 방법에 머무르지 않고 새로운 지식과 가능성을 탐색합니다.

- GitHub와 기술 블로그를 통한 학습 내용 기록
- 최신 AI·통계 논문을 읽고 방법론 직접 구현
- 대조학습, Gaussian Process, Bayesian Optimization 등 다양한 방법론 탐구

논리적인 사고

문제를 구조화하고, 데이터와 평가 기준에 근거해 해결책을 설계합니다.

- 평가지표를 분석한 후 학습·검증 전략 설계
- Time-Consistency 가설을 기반으로 변화점 탐지 방법 개발
- LG Aimer 7기 프로젝트에서 도메인 지식을 근거로 파생변수와 커널 구조 설계
- 반복 실험과 Ablation Study를 통해 모델 개선 효과 검증

끝까지 구현하는 실행력

아이디어를 모델에 머물게 하지 않고, 실제로 작동하는 시스템까지 구현합니다.

- 센서 수집부터 모델 학습·경량화·BLE 통신까지 End-to-End 구현
- 논문 속 방법론을 코드로 재현하고 실제 데이터에 적용
- 제한된 시간 안에 모델 개선과 제출을 반복해 해커톤 본선 7위 달성
- 팀장으로서 일정, 역할, 발표자료와 최종 결과물 관리

신뢰할 수 있는 AI

높은 정확도뿐 아니라, 예측의 근거와 불확실성까지 함께 고민합니다.

- Gaussian Process를 활용한 예측 불확실성 정량화
- Conformal Prediction을 공부하고 독성 예측 연구에 적용
- 확률 예측 문제에서 Brier Score와 Calibration을 중심으로 모델 평가
- 새로운 데이터와 환경에서도 성능이 유지되는지 반복적으로 검증