

# 2025학년도 소중한 JUMP-UP Labs 결과보고서(공개용)

과제 현황					
과제명	AI(인공지능)를 활용한 산업체 전력 사용량 이상치 탐지				
팀명	ECO AI LAB				
과제 수행기간	2025년 4월 ~ 2025년 11월				
Github주소 (해당시)	<a href="https://github.com/HBNU-SWUNIV/jum-up-labs25-ecoai-2025">https://github.com/HBNU-SWUNIV/jum-up-labs25-ecoai-2025</a>				
지도 교수					
소속	국립한밭대학교 컴퓨터공학과				
성명	이상금				
전화					
E-mail					
참가 인원(팀원)					
No.	성명	소속학과	학번	전화번호	E-mail
1	이지상	컴퓨터공학과			
2	김준연	컴퓨터공학과			
3	박범도	컴퓨터공학과			
4	장현석	컴퓨터공학과			
5	박준성	컴퓨터공학과			
6	권우현	컴퓨터공학과			
7	정민성	컴퓨터공학과			
8	유동건	컴퓨터공학과			
9	이정섭	컴퓨터공학과			

## 과제 세부 설명

1. 개발 동기 및 목적, 필요성	<p>전 세계적으로 탄소 중립이 주요 과제로 대두됨에 따라, 대한민국 역시 2050년까지 탄소 중립 실현을 목표로 하고 있다. 이산화탄소 배출의 상당 부분은 화석 연료와 산업 활동에서 비롯되며, 이에 대한 해결책으로 산업체의 전력 데이터를 기반으로 최적의 솔루션을 제공하는 '그린버튼(Green Button)' 플랫폼이 주목받고 있다. 하지만 산업 공정 데이터는 다양한 환경 요인과 비정상적 운영으로 인해 신뢰성을 저해하는 이상치(anomaly)를 포함할 가능성이 높다. 본 연구는 이 산업체 전력 소비 데이터를 분석하고 이상치를 효율적으로 탐지하여 그린버튼 플랫폼의 신뢰성과 정확성을 높이는 방안을 제시하는 것을 목적으로 한다.</p>  <pre> graph LR     subgraph Generation [생성]         direction TB         G1[에너지 공급사]         G2[금속 산업]         G3[식품 산업]         G4[에너지 산업]         G5[요업]         G6[제지 목재]         G7[화학공학]         G8[산업 기타]     end          subgraph DataCollection [데이터 수집]         direction TB         DC1[Open - API]         DC2[DBMS 수집]         DC3[웹 수집]         DC4[서비스 연계]     end          subgraph DataStorage [데이터 저장]         direction TB         DS1[데이터 전처리]         DS2[데이터 매칭]         DS3[NOSQL]         DS4[Hadoop]     end          subgraph DataAnalysis [데이터 분석]         direction TB         DA1[예측 분석]         DA2[통계 분석]         DA3[다차원 분석]         DA4[실시간 분석]     end          subgraph DataSharing [데이터 공개·공유 및 서비스]         direction TB         DS1[시각화]         DS2[일반 시각화]         DS3[특성화 시각화]         DS4[서비스]         DS5[모니터링]         DS6[소셜분석]         DS7[API]         DS8[Open API]         DS9[데이터 프리즌]         DS10[데이터 거래마켓]     end          subgraph Services [서비스 대상]         direction TB         S1[전문]         S2[정부 지자체]         S3[학교]         S4[연구기관]         S5[에너지 업체]     end          G1 -- 전달 --&gt; DC1     G2 -- 전달 --&gt; DC1     G3 -- 전달 --&gt; DC1     G4 -- 전달 --&gt; DC1     G5 -- 전달 --&gt; DC1     G6 -- 전달 --&gt; DC1     G7 -- 전달 --&gt; DC1     G8 -- 전달 --&gt; DC1          DC1 -- 전달 --&gt; DS1     DC2 -- 전달 --&gt; DS1     DC3 -- 전달 --&gt; DS1     DC4 -- 전달 --&gt; DS1          DS1 -- 전달 --&gt; DA1     DS2 -- 전달 --&gt; DA1     DS3 -- 전달 --&gt; DA1     DS4 -- 전달 --&gt; DA1          DA1 -- 전달 --&gt; DS1     DA2 -- 전달 --&gt; DS1     DA3 -- 전달 --&gt; DS1     DA4 -- 전달 --&gt; DS1          DS1 -- 전달 --&gt; S1     DS2 -- 전달 --&gt; S1     DS3 -- 전달 --&gt; S1     DS4 -- 전달 --&gt; S1     DS5 -- 전달 --&gt; S1     DS6 -- 전달 --&gt; S1     DS7 -- 전달 --&gt; S1     DS8 -- 전달 --&gt; S1     DS9 -- 전달 --&gt; S1     DS10 -- 전달 --&gt; S1          S1 -- 전달 --&gt; S2     S1 -- 전달 --&gt; S3     S1 -- 전달 --&gt; S4     S1 -- 전달 --&gt; S5 </pre>
2. 과제 수행 내용	<p>본 과제는 산업 데이터의 비정상성(Non-stationarity)과 클래스 불균형 문제를 해결하기 위해 다음과 같은 다각도의 접근 방식을 수행했다.</p> <ul style="list-style-type: none"> <li>○ GELU CNN-GRU AE 기반 이상치 탐지 모델 개발: 시계열의 지역적 특징(CNN)과 시간적 흐름(GRU)을 동시에 학습하는 오토인코더(Autoencoder) 모델을 설계했다. 활성화 함수로는 GELU를 사용하여 비선형성을 강화했으며, 재구성 오차 측정 시 시간 왜곡에 강인한 DTW(Dynamic Time Warping)를 적용하고 Isolation Forest로 최종 이상치를 판별하는 하이브리드 방법론을 제안했다.</li> <li>○ 데이터 해상도에 따른 모델 최적화 연구: 데이터의 수집 주기(시간 해상도)가 모델 성능에 미치는 영향을 분석했다. 고해상도 데이터에서는 LSTM Autoencoder가, 중·저해상도 데이터에서는 적대적 학습(Adversarial Training) 기반의 USAD(Unsupervised Anomaly Detection) 모델이 더 우수한 성능을 보임을 규명했다.</li> <li>○ MI-FGSM 기반 데이터 증강 (Data Augmentation): 이상치 데이터가 극히 부족한 화학 산업 데이터의 클래스 불균형 문제를 해결하기 위해, 정상 데이터에 미세한 노이즈를 주입하여 적대적 샘플을 생성하는 MI-FGSM(Momentum Iterative-Fast Gradient Sign Method) 기법을 적용했다. 이를 통해 탐지 모델의 F1 Score를 비약적으로 향상시키는 임계값을 도출했다.</li> </ul>

	<p>본 연구를 통해 도출된 정량적 및 정성적 성과는 다음과 같다.</p> <ul style="list-style-type: none"> <li>○ 산업 특화 이상치 탐지 성능 확보: 화학 산업: MI-FGSM 증강 기법을 LSTM-AE에 적용하여, 최적 노이즈 임계값에서 F1 Score 0.9778을 달성하며 클래스 불균형 문제를 효과적으로 해결했다. 산업 전반: GELU CNN-GRU AE 모델을 통해 급격한 피크 부하와 미세한 에너지 누수 패턴을 모두 탐지하는 데 성공했으며, DTW 적용으로 단순 시간 밀림에 의한 오탐지를 방지했다.</li> </ul>
3. 과제 수행 결과	<ul style="list-style-type: none"> <li>○ 해상도별 모델 선정 가이드라인 제시: 실제 산업 현장의 중·저해상도 데이터 환경에서는 USAD 모델이 미세한 이상 변동을 더 민감하게 포착함을 입증하여, 현장 데이터 특성에 맞는 모델 선택 기준을 수립했다.</li> <li>○ 학술적 성과: 본 과제 수행 결과를 바탕으로 총 3편의 학술대회 논문을 발표하였으며, 이를 통해 제안 방법론의 독창성과 우수성을 검증받았다. FFT 기반 주기 추출 및 GELU CNN-GRU AE 모델 (박범도 등) LSTM AE와 USAD의 해상도별 성능 분석 (박준성 등) MI-FGSM 데이터 증강 기법 (장현석 등)</li> </ul>
4. 팀원 역할 분담	<ul style="list-style-type: none"> <li>• 박범도(모델 개발) FFT 기반 전처리 및 GELU CNN-GRU AE 모델 설계 및 구현</li> <li>• 권우현(모델 개발) CNN-GRU 하이퍼파라미터 튜닝 및 Isolation Forest 연동 구현</li> <li>• 장현석(데이터 증강) MI-FGSM 알고리즘 구현 및 클래스 불균형 해소 실험 수행</li> <li>• 정민성(데이터 분석) 적대적 샘플 생성 실험 및 <math>\epsilon</math> 임계값에 따른 성능 변화 분석</li> <li>• 박준성(성능 비교 연구) LSTM-AE 및 USAD 모델 구현, 데이터 해상도별 비교 실험</li> <li>• 이정섭(데이터 전처리) LEAD1.0, KPI 등 공개 데이터셋 전처리 및 정규화 수행</li> <li>• 유동건 (실험 지원) 모델 성능 지표(F1, Precision, Recall) 측정 및 결과 시각화</li> <li>• 이지상, 김준연(연구 자문) 최신 이상치 탐지 동향 조사, 실험 설계 검토 및 논문 작성 지도</li> </ul>

5. 결과물 사진  
또는 동작 화면

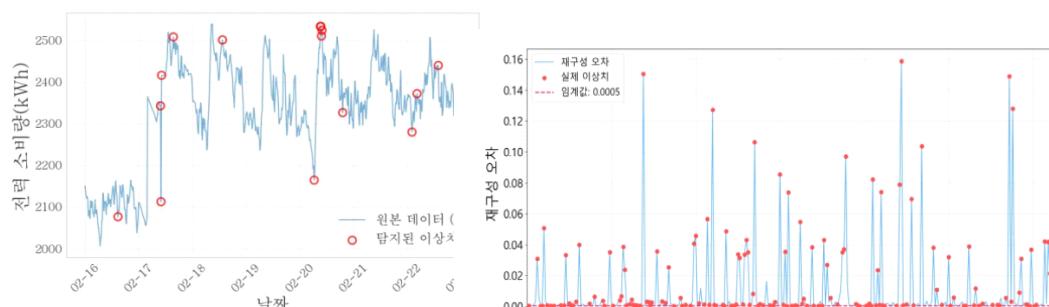


그림 4. (2020년 2월 16일 ~ 2020년 2월 22일)  
이상치 검출 결과

그림 2 이상치 탐지 재구성 오차 분포( $\epsilon = 0.01$ )

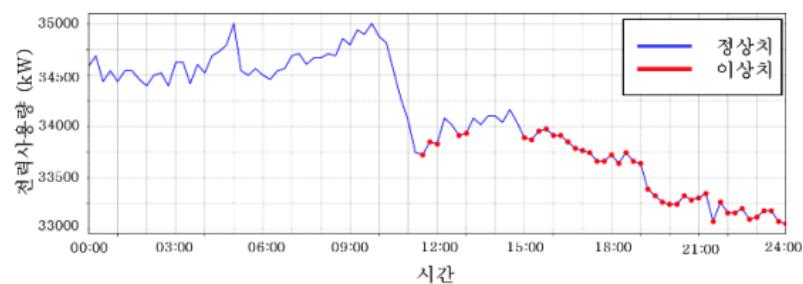


그림 3. 실제 산업 전력 데이터셋의 전력사용량

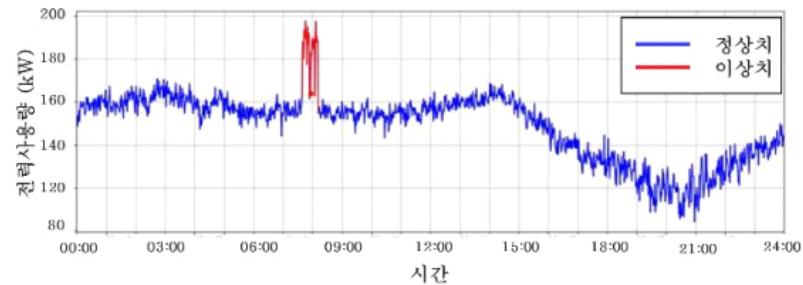


그림 4. KPI 데이터셋의 전력사용량

6. 기타	<ul style="list-style-type: none"> <li>학술대회 발표 성과: 본 과제 수행 결과를 바탕으로 '대한전자공학회 하계종합학술대회' 등 주요 학술대회에서 총 3건의 논문을 발표하였음.</li> <li>후속 연구 계획: 본 연구에서 검증된 모델을 기반으로, 금융(사기 탐지) 및 보안(침입 탐지) 분야로의 도메인 확장 가능성을 확인하였으며, 향후 실시간 스트리밍 데이터 처리를 위한 파이프라인 최적화 연구를 진행할 예정임.</li> </ul>
7. 작품사진	<p>The image shows a full-page research paper. At the top left is the logo of the Korea National University of Maritime Science and Technology. The title ' 이상치 탐지 성능 향상을 위한 MI-FGSM 데이터 증강 기법 ' is prominently displayed. Below the title is the author information: 장현석1, 정민성1, 이충호2, 허태욱2 *이상금1. The paper is divided into several sections: 'Introduction &amp; Research Background', 'Abstract', 'Experiment &amp; Result', 'Conclusion', and 'Future Research'. The 'Experiment &amp; Result' section contains numerous figures, including bar charts, scatter plots, and line graphs, which show the performance of the proposed MI-FGSM method compared to baseline models like Isolation Forest. A QR code is located in the top right corner of the paper.</p>



# LSTM Autoencoder와 Unsupervised Anomaly Detection 모델의 시간 해상도별 이상치 탐지 성능 분석

박준성<sup>1</sup>, 이정섭<sup>1</sup>, 유통건<sup>1</sup>, 이충호<sup>2</sup>, 허태욱<sup>2</sup>, \*이상금<sup>1</sup>

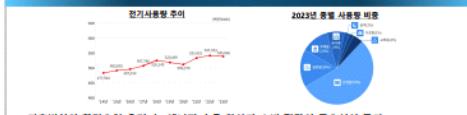
<sup>1</sup>국립한밭대학교<sup>1</sup>, 한국전자통신연구원<sup>2</sup>

EcoAI Lab

## Abstract

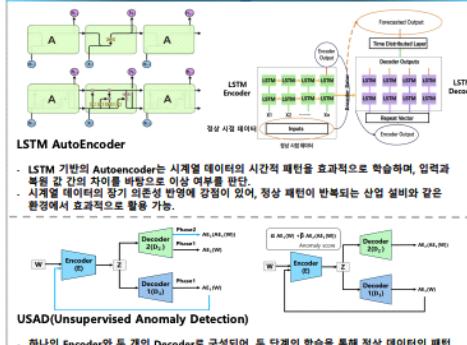
- 본 연구는 산업 전력 데이터 내 이상치를 탐지하여 에너지 낭비를 줄이고 전력 효율을 향상시키는 것을 목표로 하며, 이를 위해 시계열 기반 비지도 학습 모델의 성능을 분석.
- LSTM(Long Short-Term Memory) Autoencoder와 USAD(Unsupervised Anomaly Detection) 모델을 적용하여, 시간 해상도에 따른 이상 탐지 성능 차이를 비교.
- 실험에는 저해상도의 LEAD1.0, 고해상도의 KPI, 실제 산업 전력 데이터셋을 활용하고, F1-score, Precision, Recall 지표를 통해 정량적 성능 평가를 수행.
- USAD는 종자해상도 데이터에서 LSTM Autoencoder보다 높은 성능을 보이며, 이는 실제 산업 환경에서의 선서 제작이나 저장 관리를 고려했을 때 실용적임을 시사.
- 시계열 데이터의 해상도에 따라 모델의 이상 탐지 성능이 달라질 수 있으며, 데이터 특성에 맞는 모델 선택 기준의 필요성을 강조.

## 개요



- 기후변화와 환경오염 유려 속, 에너지 효율 향상을 소비 절감의 중요성이 증가.
- 산업 부문은 2023년 기준 전체 전력 사용량의 53% 가지며, 에너지 관리의 핵심 대상으로 부각.
- 최근 10년간 전력 사용량이 지속적으로 증가했으며, 산업을 전력 소비의 과반 이상을 차지.
- 산업 설비 운영 중 발생하는 이상치는 에너지 낭비와 시스템 불안정성을 초래할 수 있음.
- 이상치 탐지를 통해 효율적인 전력 관리가 가능하므로, 이에 대한 필요성이 강조됨.

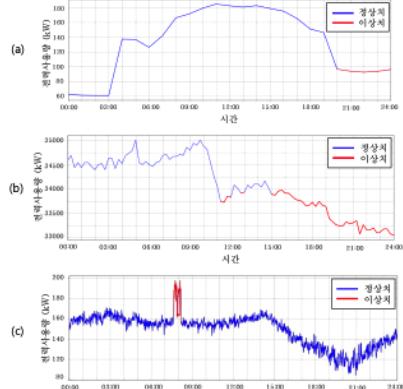
## 이상치 탐지 모델



- LSTM 기반의 Autoencoder는 시계열 데이터의 시간적 패턴을 효과적으로 학습하여, 입력과 복원 간의 차이를 바탕으로 이상치를 판단.
- 시계열 데이터의 장기 의존성 반영에 강점이 있어, 정상 패턴이 반복되는 산업 설비와 같은 환경에서 효과적으로 활용 가능.

- 하나의 Encoder와 두 개의 Decoder로 구성되어, 두 단계의 학습을 통해 정상 데이터의 패턴을 효과적으로 학습하고 입력과 복원 간의 차이를 이용해 이상치를 탐지.
- 정상과 비정상 사이의 경계에 위치한 예제는 이상치 구간까지 효과적으로 구분할 수 있음.

## 데이터 구성 및 전처리



데이터셋의 시간별 전력 사용량 그래프 (a) LEAD1.0 데이터셋 (b) 실제 산업 전력 데이터셋 (c) KPI 데이터셋

- 이상치 탐지 성능 평가를 위해, 두 공개 데이터셋 LEAD1.0 저해상도, 1시간 간격과 KPI-Anomaly-Detection(고해상도, 1분 간격)과, 실제 산업 환경에서 15분 간격으로 수집된 중해상도 산업 전력 데이터를 활용.
- LEAD1.0과 KPI 데이터셋에는 수작업 또는 전문가에 의해 라벨링 된 이상치 정보가 포함되어 있으며, 실제 산업 데이터는 라벨 없이 모델 성능을 간접 평가하는 기준으로 사용.

## 데이터 전처리

- LEAD1.0 : 산업 전력과 사용한 건물을 선택.
- KPI-Anomaly-Detection : 산업 전력의 정상적인 분포를 보이는 항목을 선택.
- 세 데이터셋 모두 : 전력 사용량은 Min-Max 정규화를 적용해 일자 폭정 간 스케일 차이 조정.

## 실험 방법 및 결과 분석

	LEAD1.0	Industrial Power	KPI
Train	6541	10054	65049
Validation	1747	3481	21588
Test	1747	3482	21584

표1 세 데이터셋을 통한 검증/검증/테스트 6:2:2의 비율로 분할한 표.

	LEAD1.0	Industrial Power	KPI
Model	LSTM Autoencoder	USAD	LSTM Autoencoder
Recall	0.8510	0.8179	0.7777
Precision	0.8665	0.8665	0.8676
F1 Score	0.7015	0.8690	0.8195
Recall	0.8065	0.8147	0.8285
Precision	0.8626	0.8764	0.8511
AUC	0.8686	0.8685	0.8190
Recall	0.8686	0.8685	0.8686
Precision	0.8686	0.8685	0.8686

표2. LEAD1.0, KPI와 실제 산업 전력 데이터셋에서 LSTM Autoencoder와 USAD 이상치 탐지 성능 비교

(표1) 이상치 탐지 성능을 평가하기 위해 산업 전력 데이터 구조를 변경 한 세 가지 시계열 데이터셋을 훈련/검증/테스트 6:2:2 비율로 활용.

(표2) USAD 모델이 저해상도 데이터에서는 더 높은 탐지 성능을 보이지만, KPI와 같은 고해상도 데이터에서는 상대적으로 낮은 성능을 보임. 반면 폭이 작고 이상치가 드물게 발생하는 환경에서 USAD의 적대적 학습 구조가 미세한 이상치에도 민감하게 반응.

실제 산업 환경에서 자주 발생하는 저해상도 데이터 수집 환경을 고려할 때, USAD의 Decoder 구조를 통해 미세한 이상치 변화에도 민감한 이상치 탐지에 강점을 지닌 모델로 평가.

## 결론 및 향후 연구

- LSTM Autoencoder와 USAD를 활용하여 서로 다른 시간 해상도를 가진 산업 전력 데이터셋에서의 이상치 탐지 성능을 비교.
- 실험 결과, 데이터셋의 해상도에 따라 두 모델의 성능 차이가 나타나며, 저해상도 환경에서는 USAD가 높은 탐지 성능을 나타내고, LSTM Autoencoder는 상대적으로 낮은 탐지 성능을 보임.
- USAD는 복잡한 LSTM 구조 없이 이상치에 민감하게 반응할 수 있어 저해상도 데이터를 주로 이루는 산업 환경에 적합한 대안.
- 향후 두 모델의 강점을 결합한 하이브리드 이상치 탐지 모델을 설계하여 해상도 변화나 실시간 환경에서도 강한 이상치 탐지 성능을 유지하는 시스템을 구축할 예정.
- 다양한 해상도의 산업 전력 데이터의 신뢰성을 향상시키고, 에너지 효율 최적화 및 탄소 배출 저감 목표 달성을 기여하고자 함.

**고속 푸리에 변환(FFT) 기반 주기 추출 및  
윈도우 구성을 활용한 GELU CNN-GRU AE 모델의  
산업 전력 시계열 이상치 탐지**

박범도<sup>1</sup>, 권우현<sup>1</sup>, 이충호<sup>1</sup>, 허태욱<sup>2</sup>, \*이상금<sup>1</sup>  
 \*국립한발대학교<sup>1</sup>, 한국전자통신연구원<sup>2</sup>

**EcoAI Lab**

**Abstract**

(연구 배경 및 목표) 산업 현장의 복잡한 전력 데이터로부터 성비, 교정 예방 및 에너지 효율 최적화를 목표로, FFT 기반의 동적 윈도우 구조와 GELU CNN-GRU 오토인코더를 결합한 이상치 탐지 방법론을 제안 (데이터 정제화 및 주기 추출 1차 차분으로 데이터의 변화 패턴을 강조하고, FFT 분석으로 주요 주파수 성분을 식별하여 삼위 주기를 효과적으로 추출, 이를 동적 분석 윈도우로 실행). (다중 윈도우 기반 이상 탐지 모형) 주제별 주기들로 다중 스케일 윈도우를 구성, CNN으로 지역적 패턴을 예측하고, GRU 시간적 의존성을 학습하는 오토인코더 모델을 설계, GELU 활성화 함수로 학습 안정성을 확보 (DTW 및 Isolation Forest 기반 이상치 판별) 시간 축 외곡에 강인한 오차 조정을 위해 DTW를 활용하고, 계산된 오차 평균에 비지도 학습인 Isolation Forest을 적용하여 이상치 윈도우를 최종 판별 (연구 의의) 제안된 단계별 방식은 실제 산업 현장의 다양한 이상 상황에 정교하게 대응 가능하며, 스마트 맥로리의 데이터 기반 예지 보전 시스템 구축을 위한 핵심 기술로 활용될 수 있음

**Introduction & Research Background**

산업 설비의 예지보전과 에너지 관리 효율화, 시스템 안정성 확보를 위해서는 전력 데이터의 이상 탐지가 필수적이다. 기존의 통계적 방법은 복잡한 이상 패턴을 포착하는 데 한계가 있으며, 딥러닝 기반 Autoencoder 구조가 이상 탐지의 효과적인 대안으로 주목받고 있다. 본 연구는 FFT 기반 다중 주기 분석, CNN-GRU Autoencoder, DTW, Isolation Forest 등을 활용하여, 실제 산업 환경에서 다양한 시계열의 이상 상황을 신뢰성 있게 탐지하는 방안을 제시한다.

**Objective**

**Materials & Methods**

**데이터 전처리**

데이터는 (2020~2022년) 3년간 15분 단위로 수집된 특정 산업계의 전력 사용량 데이터로, 결속자 처리, sklearn.KNN Imputer를 이용, 시계열 데이터 평균 고려 및 연속성 확보, 1차 차분 (Differencing): 비정상성을 원화하고 급격한 변화를 강조하여 동적 윈도우 구성에 활용

**Results & Discussion**

• (정밀 이상 탐지) 본 연구는 1차 차분, FFT 기반 주기 분석, GELU-CNN-GRU Autoencoder, DTW, Isolation Forest의 단계별 결합으로 산업 전력 데이터의 다양한 이상 패턴을 효과적으로 탐지할 수 있음을 실험함  
 • (확장성) 향후 모델의 지속적인 발전을 통해 보다 다양한 산업 데이터셋으로 확장 및 이상 원인 분석, 실시간 예측 모델과의 연동을 통하여 자동화·고도화 추진 예정  
 • (스마트 맥로리 적용) 제안된 방법론은 스마트 맥로리 실시간 모니터링, 에너지 효율 관리 및 설비 예보전을 통한 산업 현장에서의 실질적 기여를 기대함

**Future Works**

**FFT 기반 주기 추출 & 동적 윈도우 구성**

그림 2. 원본 데이터의 FFT 분석 및 폭스 시험 결과  
 그림 3. 1차 차분 데이터의 FFT 분석 및 폭스 시험 결과

**GELU CNN-GRU Autoencoder 모델**

• CNN 계층으로 윈도우 내 지역적 패턴을 추출  
 • GRU 계층이 시계열의 시간적 의존성을 학습하여 저자율 장애 벡터로 감지 후 다시 원본 차원으로 복원하는 구조  
 → GELU 활성화 함수를 적용하여 학습 안정성 및 성능을 향상

**DTW 재구성 오차 기반 IsolationForest 이상치 탐지**

• 모델이 재구성한 윈도우와 원본 윈도우 간의 형태적 유사성을 시간 축 외곡에 강인한 DTW(Dynamic Time Warping)로 계산하여 재구성 오차를 산출  
 • 이 오차 평균에 비지도 학습인 Isolation Forest을 적용하여 이상지 윈도우를 최종 판별

그림 4. 원본 폭스 결과와 DTW 재구성 오차 평균

**(정밀 개선 방안)**

CNN 레이어를 GCN으로 구성하여 모델 훈련  
  
 - 이 경작한 시계열 데이터 활용  
 - 데이터 간 관계에 더 활용  
 - 미래 예측에서 더 활용  
 - RNN 대체로 고려해 활용 가능  
 - 다른 GCN 활용 노드