

# Anomaly Detection Method of LSTM-AE through SAC-based Dynamic Threshold Adjustment

(SAC 기반 동적 임계값 조정을 통한 LSTM-AE 이상 탐지 기법)

2025.09.19

발표자: 정민성

팀원: 권우현, 정민성, 이정섭

---

# 목 차

1. Abstract
2. 배경지식 설명
3. 방법론 설명
4. 실험 및 결과
5. 연구 계획

# **Abstract**

---

# Abstract

This study proposes a novel framework combining a Long Short-Term Memory Autoencoder (LSTM-AE) and the Soft Actor-Critic (SAC) reinforcement learning algorithm to overcome the limitations of traditional anomaly detection methods in handling highly volatile industrial power data. The LSTM-AE model learns normal patterns from time-series data, utilizing the reconstruction error as an anomaly score. Subsequently, the SAC agent receives the statistical properties of the reconstruction error distribution as its state and dynamically searches for an optimal anomaly detection threshold that maximizes the F1-Score. Experimental results on a real-world industrial power dataset show that the proposed model achieves higher performance in terms of F1-Score and Precision compared to a standalone LSTM-AE model. This demonstrates that a reinforcement learning-based approach for dynamic threshold setting can effectively enhance the performance of anomaly detection in complex time-series data.

# **배경 지식**

# 소프트 액터-크리틱(SAC) 알고리즘

SAC

최대 엔트로피 강화학습 프레임워크 기반

Off-policy 액터-크리틱 알고리즘

기대보상을 최대화하는 동시에 엔트로피 최대화

연속적인 공간에서의 강화학습에 활용

목표 함수

$$\pi^*(\cdot | s_t) = \operatorname{argmax}_\pi E_\pi \left[ \sum_t R_t + \beta H(\pi(\cdot | s_{-t})) \right]$$

표현식 설명

$\pi^*(\cdot | s_t)$ : 상태  $s_t$ 에서의 엔트로피와 그에 따른 기대보상을 최대화하는 정책의 확률 분포

$R_t$ : 시점 t에서의 보상

$\beta$ : 최적 정책에 영향을 미치는 상관계수로 0이면 결정적, 아니라면 확률적

$H(\pi(\cdot | s_t))$ : 정책  $\pi$ 의 통제 하에 있는 방문했던 상태의 정책에 대한 엔트로피

$\operatorname{argmax}_\pi E_\pi [\dots]$ : 정책이  $\pi$ 일 때, 괄호 안 표현식의 값이 가질 수 있는 평균이 최대값이 되는 정책 탐색

# 방법론 설명



# SAC 기반 임계값 조정

- 임계값 지정 수식

$$\tau(k) = \mu + k\sigma$$

Sac를 활용해 신뢰 구간( $k$ )의 값을 조정하면서 F1-Score를 최대화

- 보상함수

$$r(k) = F1(k) - \lambda I[Recall(k) = 0], \quad \lambda \gg 0$$

지도 학습으로 진행해 보상은 F1-score

Recall이 0이 되는 것을 막기 위해 미탐 시 -1.0 패널티 부여

# **실험 및 결과**



# 실험 및 결과

Method	F1	Precision	Recall
LSTM-AE (grid)	0.6718	0.6369	<b>0.7108</b>
SAC (ours)	<b>0.7191</b>	<b>0.7408</b>	0.6986

LSTM 단일 모델과 SAC 적용 모델 비교

데이터셋: 화공 데이터셋(MI-FGSM 논문에서 활용)



테스트 데이터의 재구성 오차 분포  
(파란색-정상, 빨간색-이상)

## 결과

- 단일 모델 대비 F1-score 0.4 상승, Precision 0.11 상승
- 다만, 전체적으로는 낮은 점수
- 테스트 데이터의 재구성 오차 분포가 정상, 이상이 일부 겹침
- 강화학습의 대상이나 보상함수의 변경이 필요

# 연구 계획



# 연구 계획

## 통신학회

- 더 좋은 데이터를 통한 실험
- 보상 함수 변경 고민
- 시각화 추가

## SCIE 10장 계획

- 빅데이터 분석에서 작성 계획 중인 라벨링 없는 데이터에 라벨링 주제 논문과 병합 가능성 고민
- 현재 산업 전력 데이터 라벨링 후 이번보다 개선된 이상탐지 강화학습 적용
- 전체적인 논문 흐름은 라벨링 없는 데이터에 라벨링 적용 강화학습 후 비지도 혹은 준지도 강화학습 적용을 통한 이상탐지
- 라벨링을 위한 강화학습 알고리즘과 탐지를 위한 알고리즘 선택이 필요

# Anomaly Detection via DTW Clustering and Reinforcement Learning with Rolling IQR

(DTW 군집화 기반 롤링 IQR 규칙 강화학습을 통한 이상치 탐지)

2025.09.19

발표자: 이정섭

팀원: 권우현, 정민성, 이정섭

---

# 목 차

1. Abstract
2. 배경지식 설명
3. Rolling IQR + RL
4. 논문 계획

# **Abstract**

---

## Abstract

We present a time-series anomaly detection framework that couples DTW-based clustering with a Rolling Interquartile Range (IQR) detector whose sensitivity parameter is adapted by reinforcement learning (RL). By first isolating pattern-homogeneous clusters with DTW- $k$ means and then learning per-cluster decision sensitivity, our method handles temporal misalignment, nonstationarity, and shifting noise regimes, improving precision–recall trade-offs over fixed-threshold baselines. Experiments on industrial power-consumption traces show consistent F1 gains, with inference throughput suitable for near real-time deployment.

# 방법론

# DTW Clustering

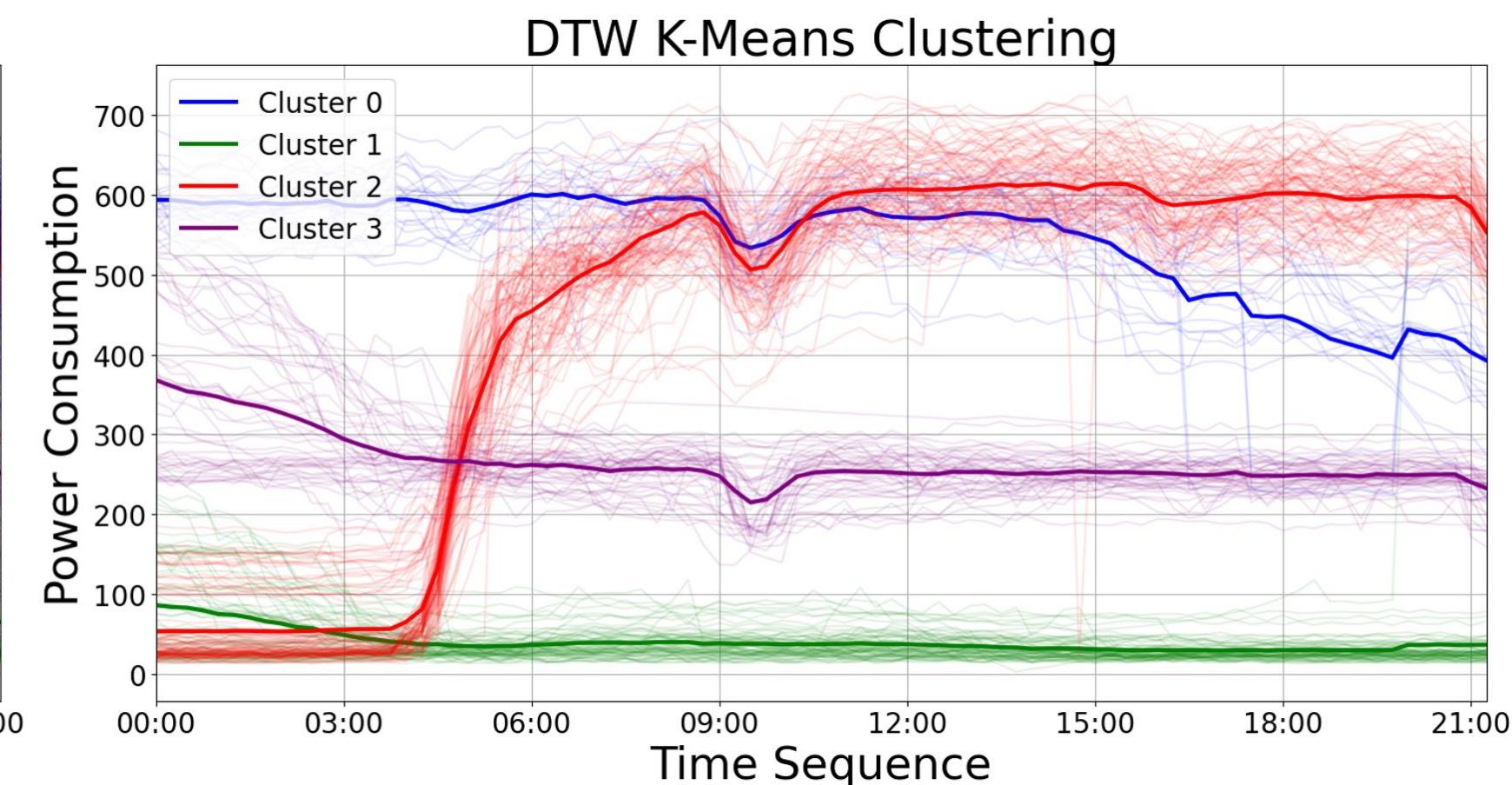
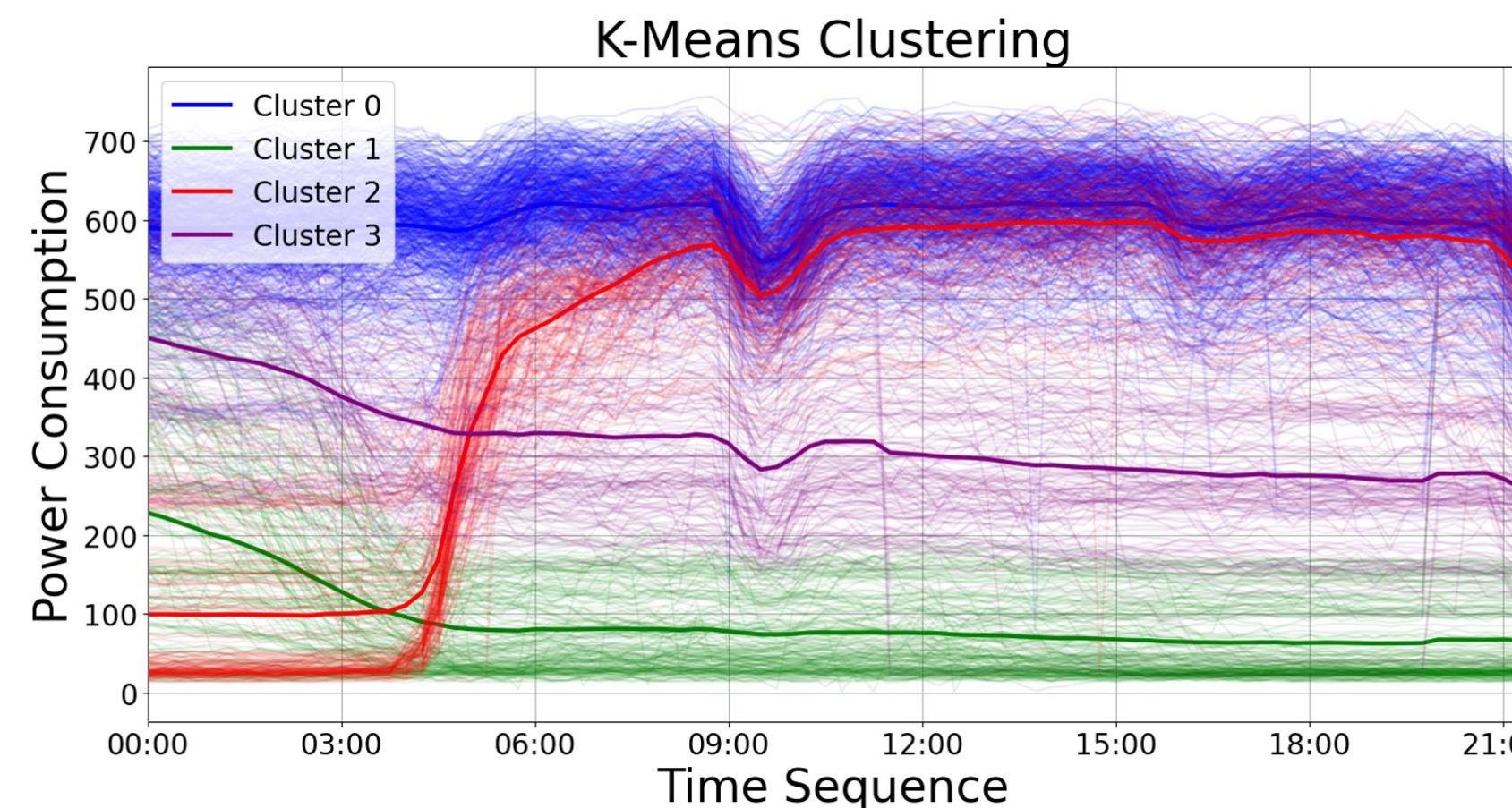
## K-means Clustering

- 주변의  $K$ 개의 이웃과 비교하여 비슷한 분포(= 크기)를 지닌 데이터끼리 군집화
- 유clidean 거리를 기반으로, 시계열 패턴이 비슷해도 다른 클러스터로 분류될 수 있음  
(= 값의 절대적인 크기에 민감)
- 정규화 및 표준화 필수

## DTW K-means Clustering

- 패턴을 중심으로 군집화
- 시간축이 늘어나거나 줄어들어도 패턴의 유사성을 비교함
- 시계열 데이터의 길이가 달라도 일부 매핑하여 계산이 가능

# DTW Clustering



# Rolling IQR

IQR

Rolling IQR

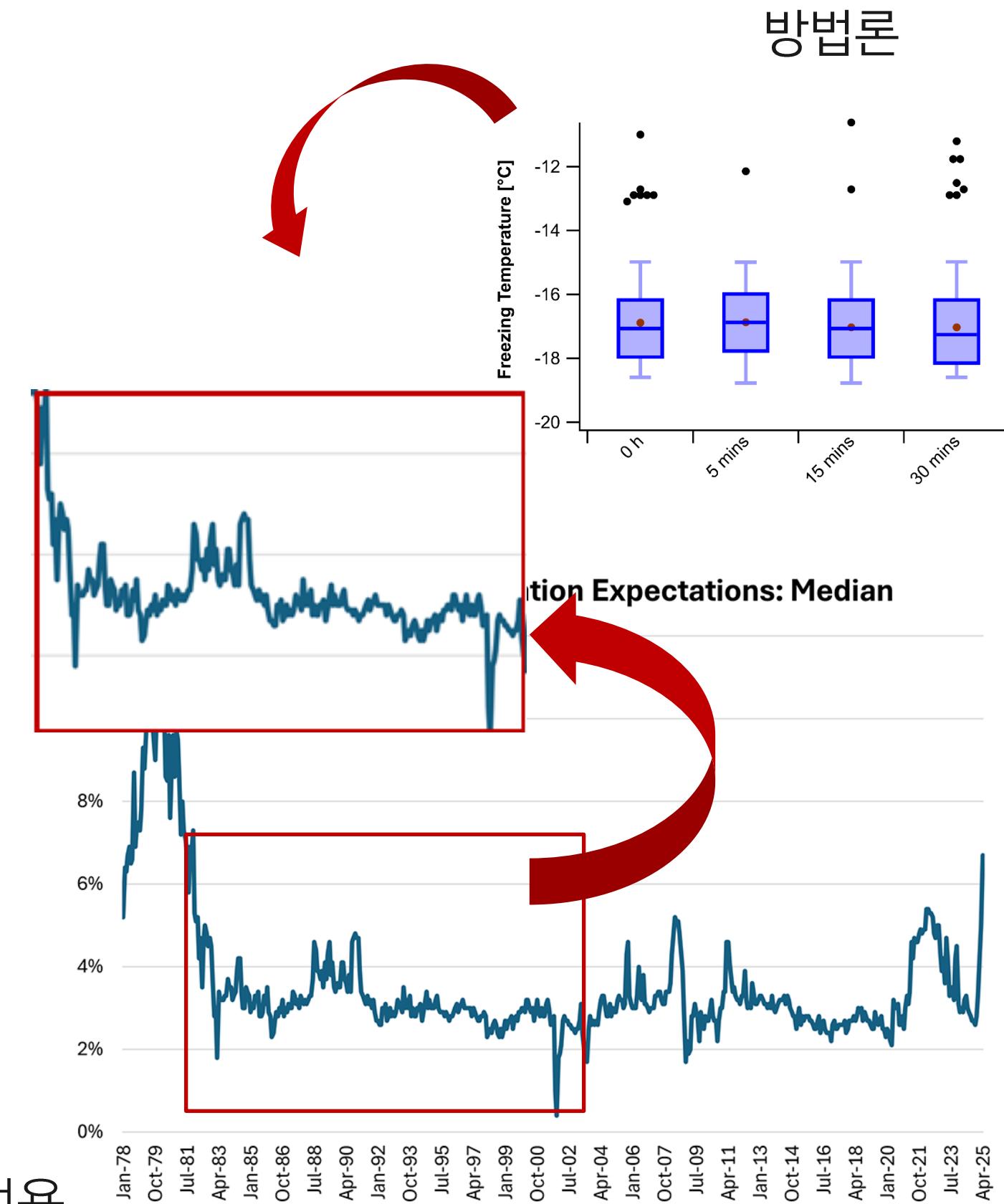
- t는 특정 시간(= 윈도우)에서의 IQR값
- Rolling Window 개념을 넣어, 윈도우 구간을 설정, 윈도우마다 IQR 적용
- 시계열 데이터나 변동성이 큰 지표에서 많이 사용

# Rolling IQR

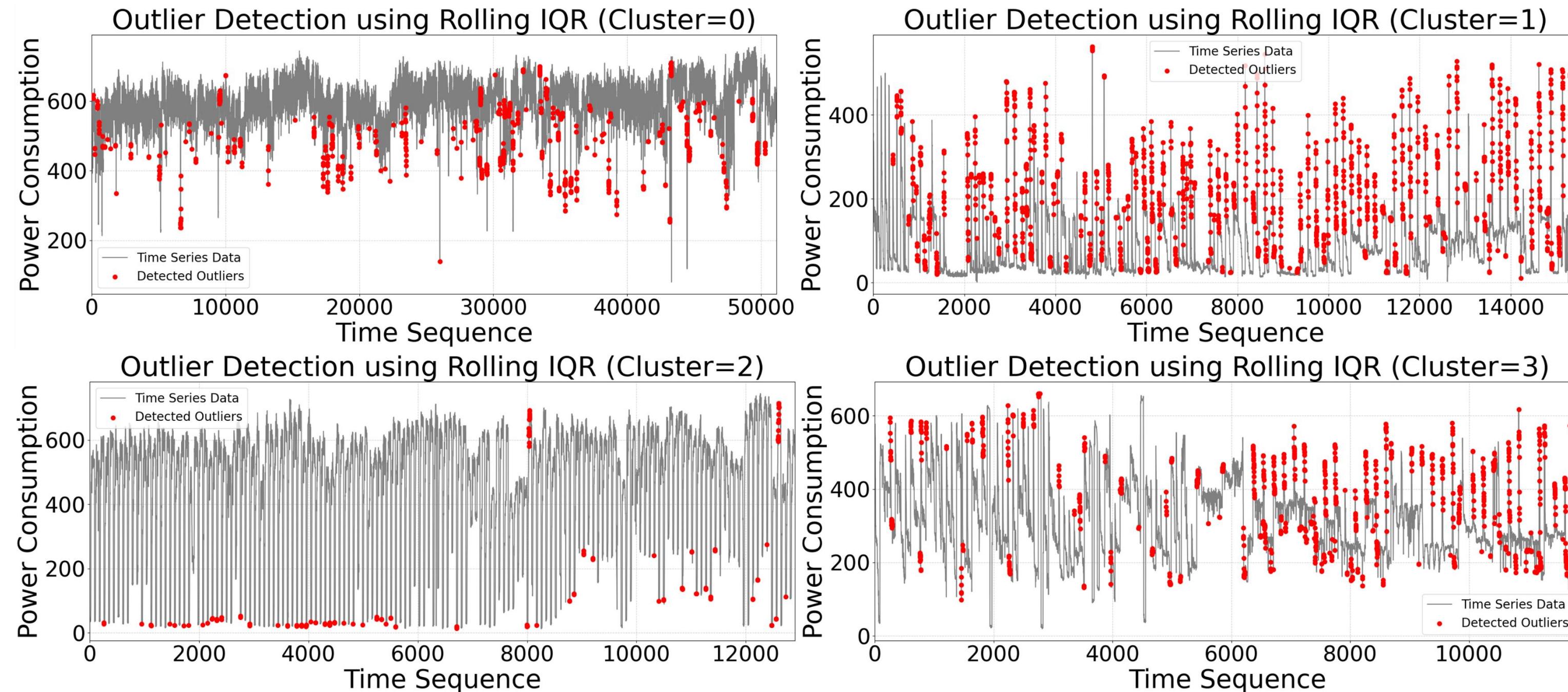
IQR

Rolling IQR

- $t$ 는 특정 시간(= 윈도우)에서의 IQR값
- Rolling Window 개념을 넣어, 윈도우 구간을 설정, 윈도우마다 IQR 적용
- 파라미터 K값을 조정하며, 데이터에서 이상치를 검출



# Rolling IQR



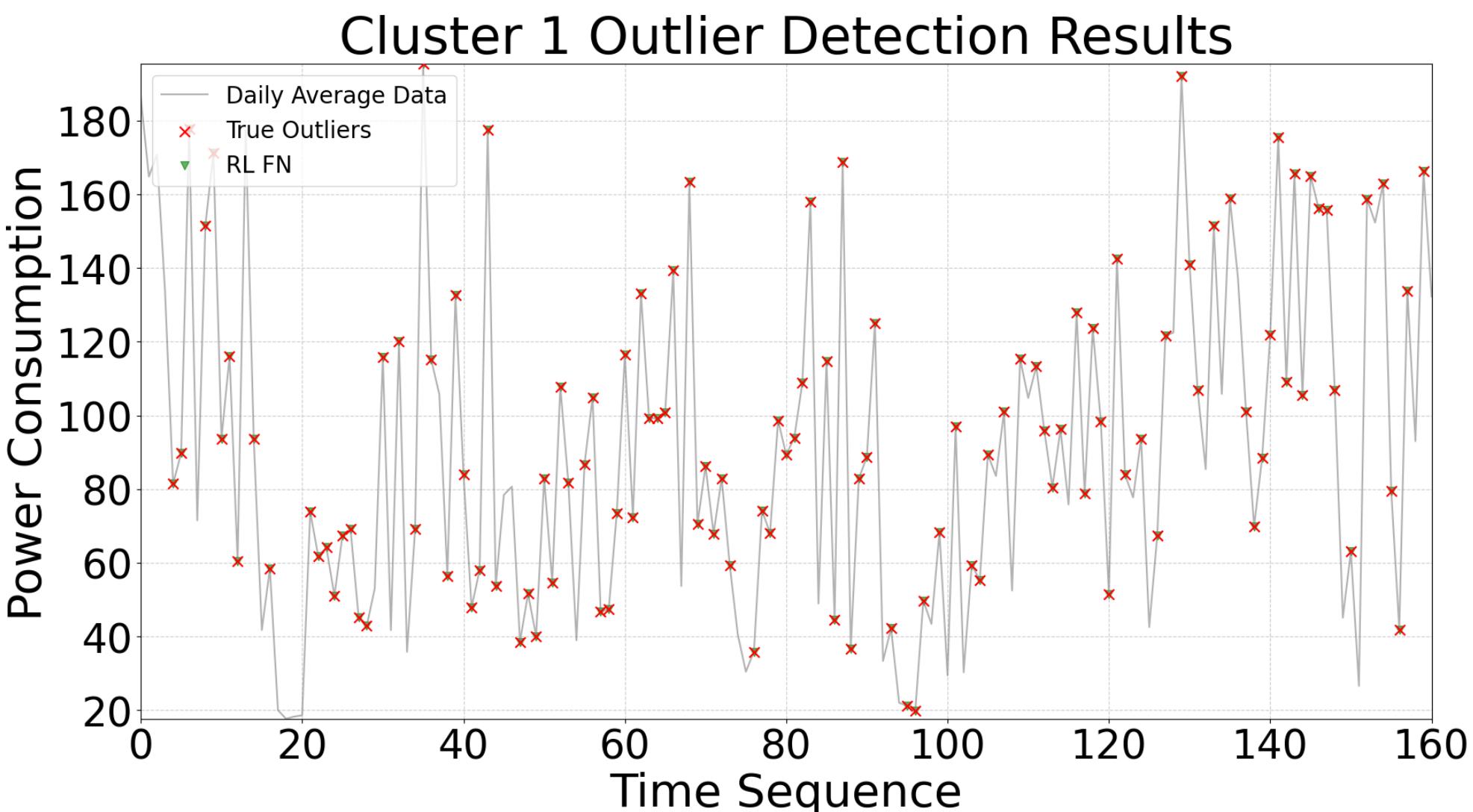
# **Rolling IQR + RL**

# Rolling IQR + RL

## Rolling IQR + RL

- Rolling IQR에서 파라미터 K값의 동적인 조정을 위해, 강화학습을 적용
- Q-Learning  
환경(Environment):  
 $\mathcal{E} = \{C_1, C_2, \dots, C_N\}, \text{where } C_k$   
상태(State):  $S_t = [v_t, Q1, Q3_t, IQR_t]$   
행동(Action):  $A = \{a_{normal}, a_{anomaly}\}$   
보상(Reward):  
 $R(s_t, a_t) = +1 \cdot \mathbb{I}(a_t = a_{anomaly} \wedge IQR \text{ Rule holds}) - 1 \cdot \mathbb{I}(a_t \neq a_{anomaly} \vee IQR \text{ Rule fails})$
- 모델은 데이터의 변동성에 따라, 누적 보상을 최대화시키고 동적으로 파라미터를 조정함

Rolling IQR + RL



# **실험 및 결과**



# 실험 및 결과

실험 및 결과

Model	F1	Precision	Recall	PR-AUC
IQR ( $k=1.5$ )	0.78	0.75	0.82	0.80
Isolation Forest	0.72	0.69	0.76	0.74
Autoencoder	0.76	0.78	0.74	0.79
<b>Ours(adapt-<math>k</math>)</b>	<b>0.85</b>	<b>0.83</b>	<b>0.87</b>	<b>0.88</b>

Table1. 모델 별 비교지표

데이터셋: Greebutton 2020~2022년 간 약 30만개의 데이터셋

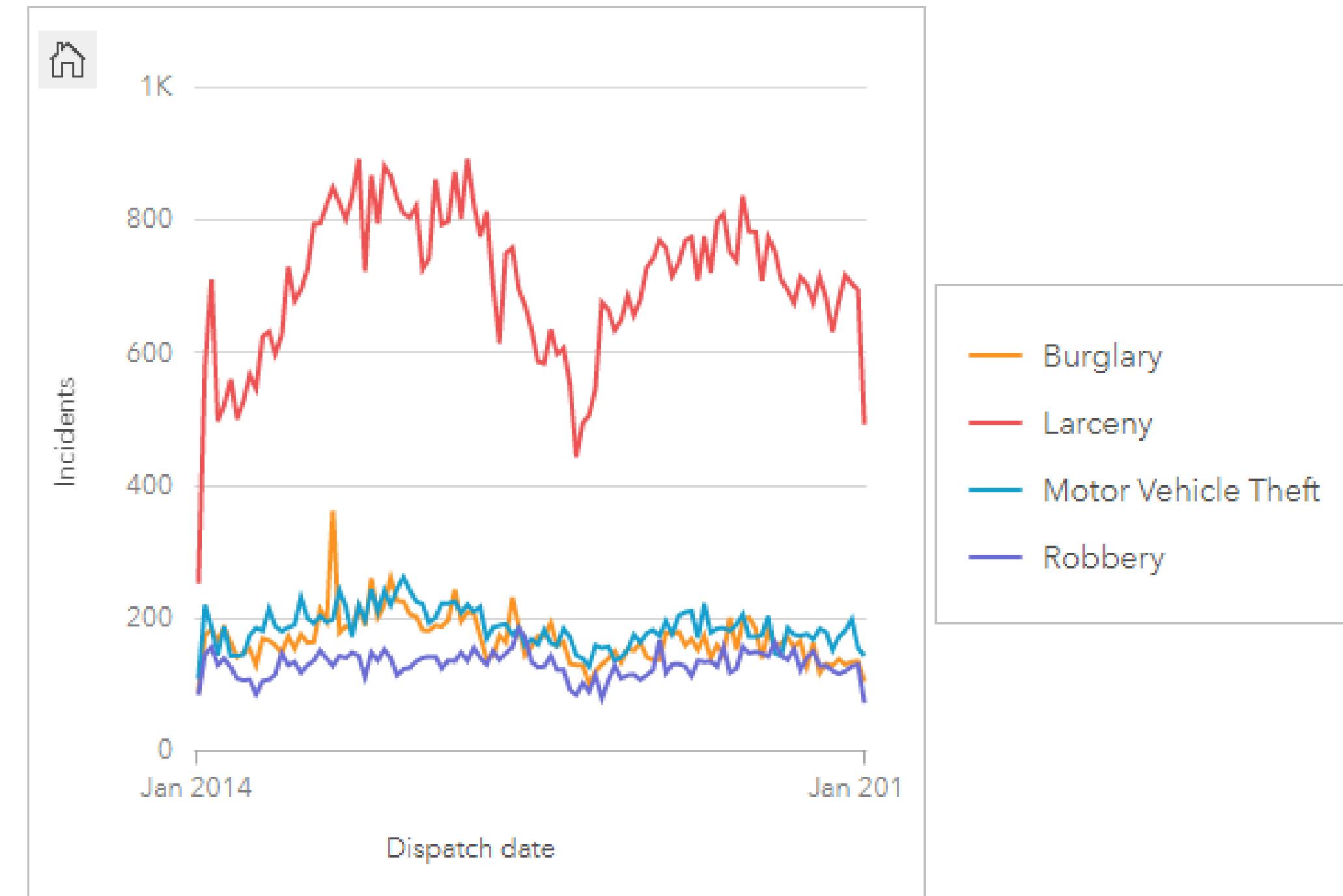
결과

- Rolling IQR + RL모델이 IQR 단독 모델과 Isolation forest, Autoencoder대비, 12%가량 우수한 성능을 보임

# 연구 계획



# 연구 계획



**감사합니다**