

국책 Rule-Based 성능평가 시나리오

☰ 태그

Rule Base 이상징후 감지정확도(F-score) 시나리오 및 분류 가능한 이상 패턴 수 확인

[배경 및 목표](#)

[진행 과정](#)

[개발 환경](#)

[시나리오](#)

[Data Generation](#)

[#1 시료 \(Nosise anomaly type\)](#)

[#2 시료 \(Drift anomaly type\)](#)

[#3 시료 \(Shift anomaly type\)](#)

[#4 시료 \(Spike anomaly type\)](#)

[#5 시료 \(Quantize anomaly type\)](#)

[Results](#)

Rule Base 이상징후 감지정확도(F-score) 시나리오 및 분류 가능한 이상 패턴 수 확인

배경 및 목표

본 프로젝트는 과제 "모델 학습 과정 없이 바로 적용 가능한 AI기반 생산 설비 이상감지 및 예지보전 관리 일정 추천 Edge Computer System 개발" 의 개발 목표 중 "Rule Base 이상징후 감지정확도(F-score)" 를 검증하는 성능 평가 시나리오와 "분류 가능한 이상 패턴 수"를 확인하는 프로젝트임

- 목표 1: 총 시료 5개에 대해서 Rule Base 이상감지 모델의 성능을 **F1 score** 로 평가하여 각 시료 별 0.95 이상($0 \leq f1 \text{ score} \leq 1$)의 성능을 가지는 것을 개발 목표로 함
- 목표 2: 감지한 이상이 학습한 5개의 패턴 중 어느 패턴으로 분류되는지 예측하는 예측 모델 개발을 목표로 함

진행 과정

시료는 반도체 도메인의 cmp 공정의 데이터로, 하나의 설비에서 5개의 'header' 압력 센서 데이터를 이용함

각 압력 센서 데이터의 랜덤한 구간에 각 header 별 서로 다른 타입(

'noise', 'drift', 'shift', 'spike', 'quantize')의 이상 패턴 데이터를 주입하여 5개의 시료를 생성하여, Rule Base 이상감지 모델의 성능을 분류 성능 지표 **F1 score**로 평가하고, 이상 패턴을 분류 예측함

시료	이상 패턴 종류	모델	데이터 타입	이상 패턴 주입 구간
#1	noise	adt model	time series	1_1_Act. value \geq 5.0 & SLOT in [7,8] & STEP_ID in [5]
#2	drift	adt model	time series	2_2_Act.value \geq 2.0 & SLOT in [8,9] & STEP_ID in [5]
#3	shift	adt model	time series	3_3_Act.value \geq 1.5 & SLOT in [6,9] & STEP_ID in [5]
#4	spike	adt model	time series	4_4_Act.value \geq 2.0 & SLOT in [6,7] & STEP_ID in [5]
#5	quantize	adt model	time series	2_2_Act.value \geq 5.0 & SLOT in [7,8] & STEP_ID in [5]

개발 환경

터미널에서 다음의 명령어를 실행 (PowerShell, git Bash, macOS/Linux 등)

1. git clone

```
git clone git@gitlab.***
```

2. 디렉토리 이동

```
cd brique_product/adt_model_assessment
```

3. 필요 라이브러리 설치

```
pip install -r requirements.txt
```

4. `scenario.ipynb` 파일 실행 및 **시료 별 성능 및 분류한 이상 패턴** 확인

시나리오

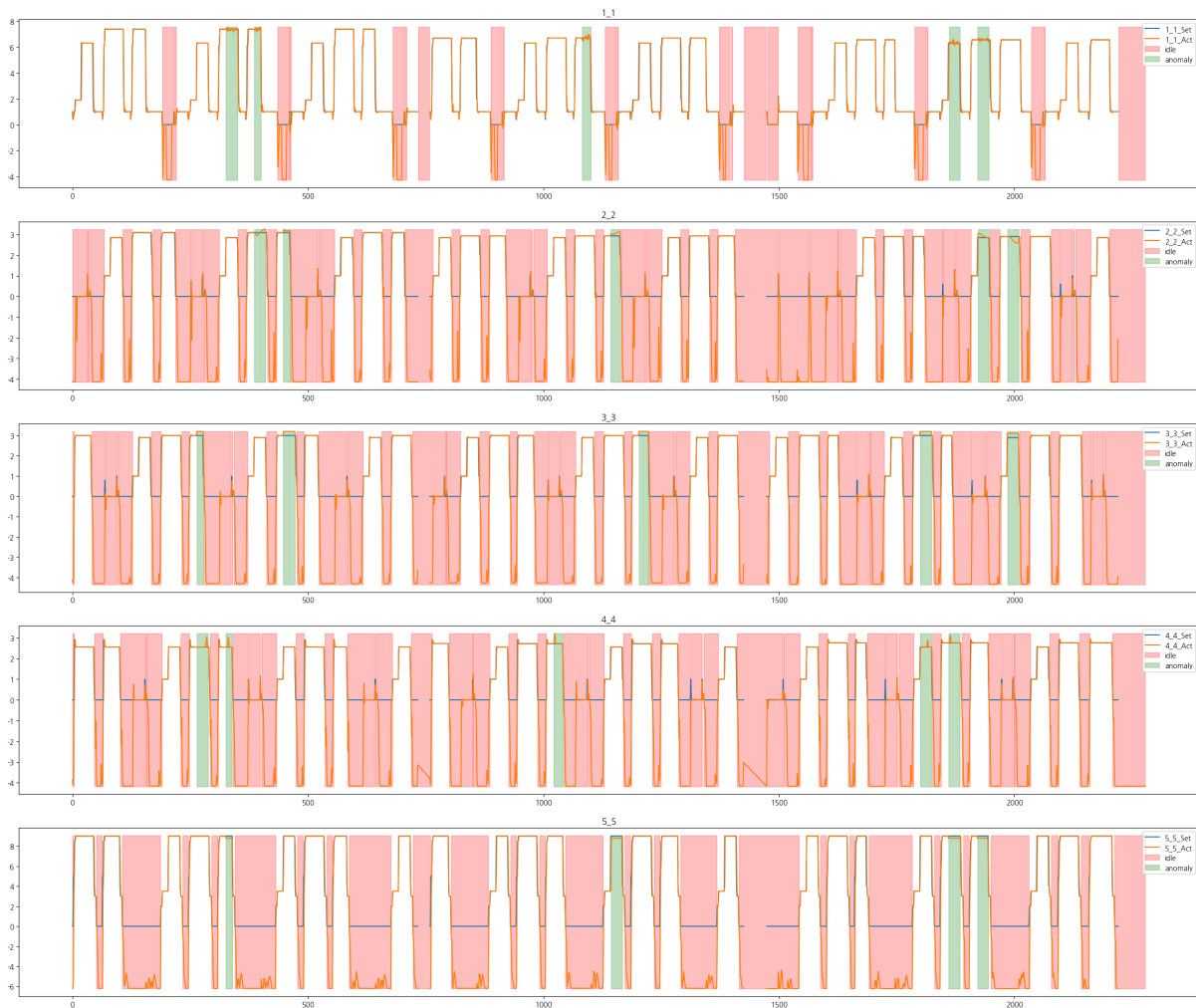
실행 파일: `scenario.ipynb`

- 반도체 cmp 공정의 정상 데이터('`x_bee_good_3.csv`')에 특정 구간에 **이상 데이터를 주입** 하고 정상 데이터와 이상 데이터 라벨을 기입하여 **5개의 시료 생성**
- Rule Base 이상감지 모델**을 통해 동일한 크기의 **구간 별 이상을 감지** 및 이상 감지 지표인 CMK 산출
- 이상 감지 지표의 Threshold 를 탐색 및 적용하여 **이상 구간을 예측**하고 실제 이상 데이터가 포함된 구간과 비교하여 성능 지표인 **F1 Score** 를 통한 **모델 성능을 평가**
- 이상 패턴 학습 (`adt_model_pattern_generation.ipynb` 참고)
- 예측 한 **이상 감지 구간에 대한 이상 패턴 분류 예측**

Data Generation

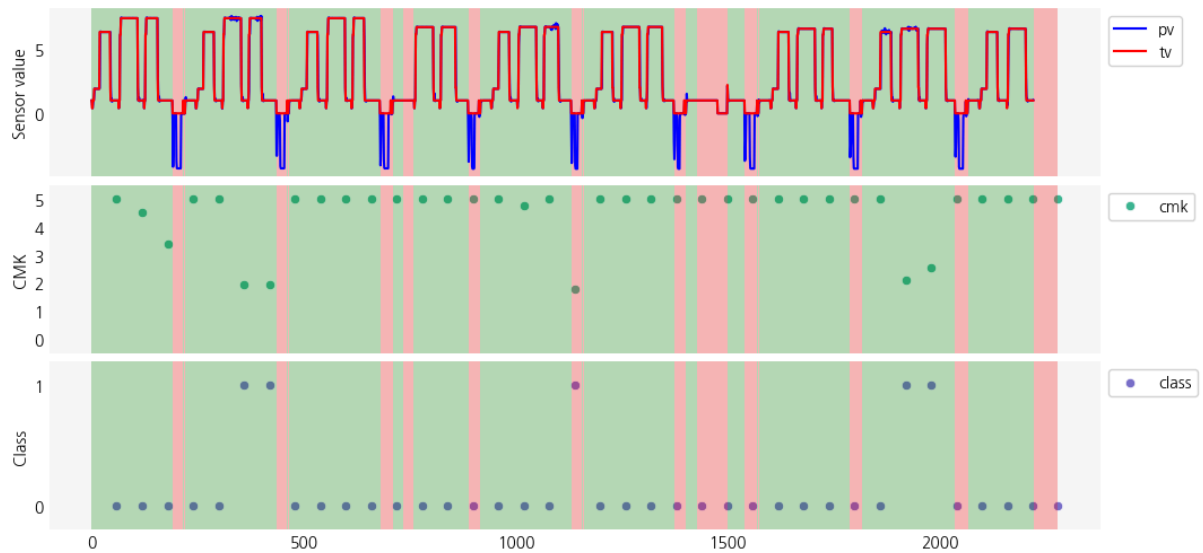
정상 데이터에 특정 구간(초록 영역) 별 이상 데이터 증강

- #1 시료
 - noise: cmp 공정의 '1_1 헤더 압력' 데이터
- #2 시료
 - drift: cmp 공정의 '2_2 헤더 압력' 데이터
- #3 시료
 - shift: cmp 공정의 '3_3 헤더 압력' 데이터
- #4 시료
 - spike: cmp 공정의 '4_4 헤더 압력' 데이터
- #5 시료
 - quantize: cmp 공정의 '5_5 헤더 압력' 데이터

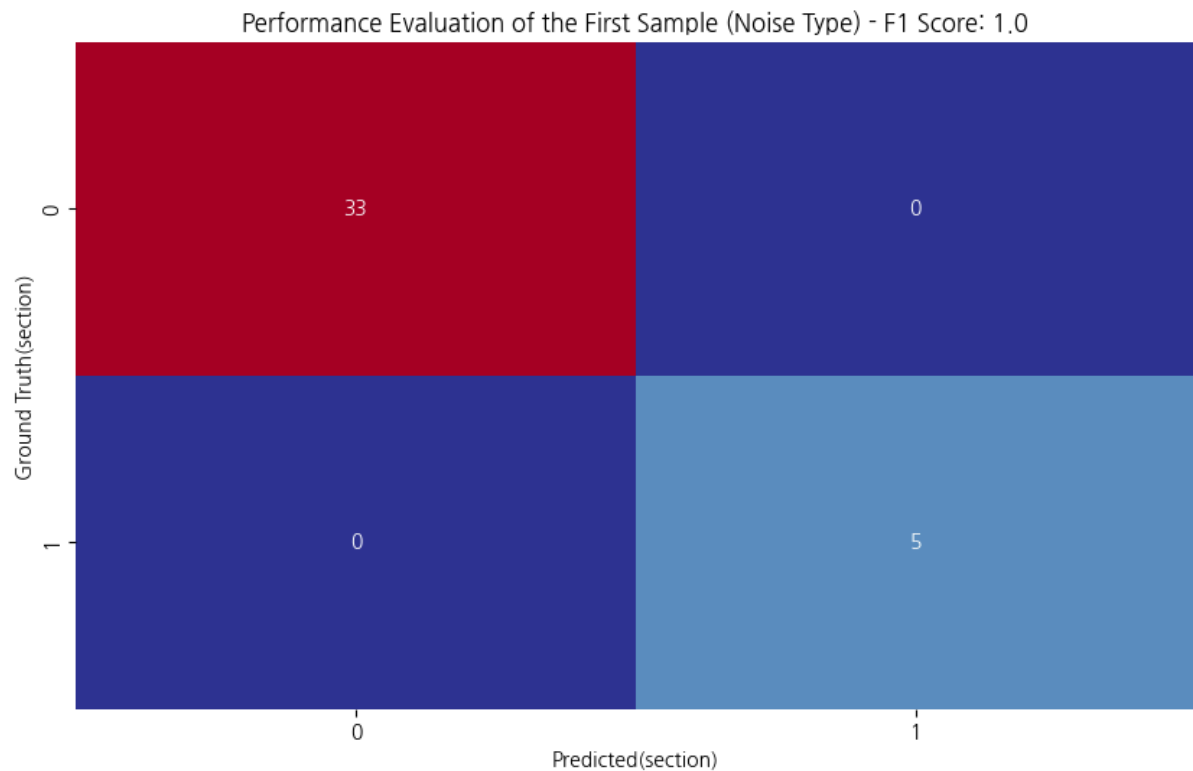


#1 시료 (Nosisse anomaly type)

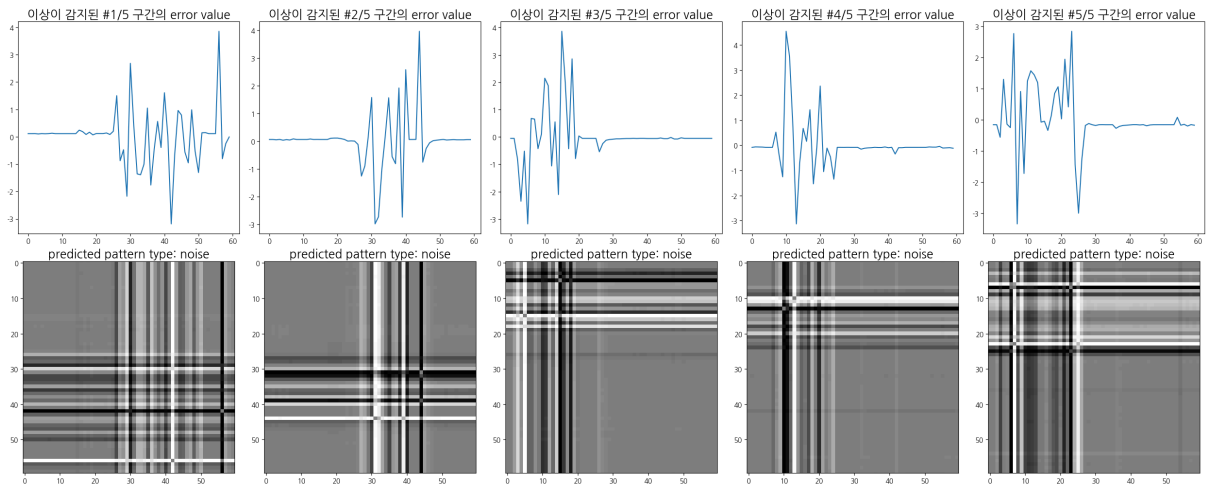
- Anomaly Detect



- Assessment (F1 score)

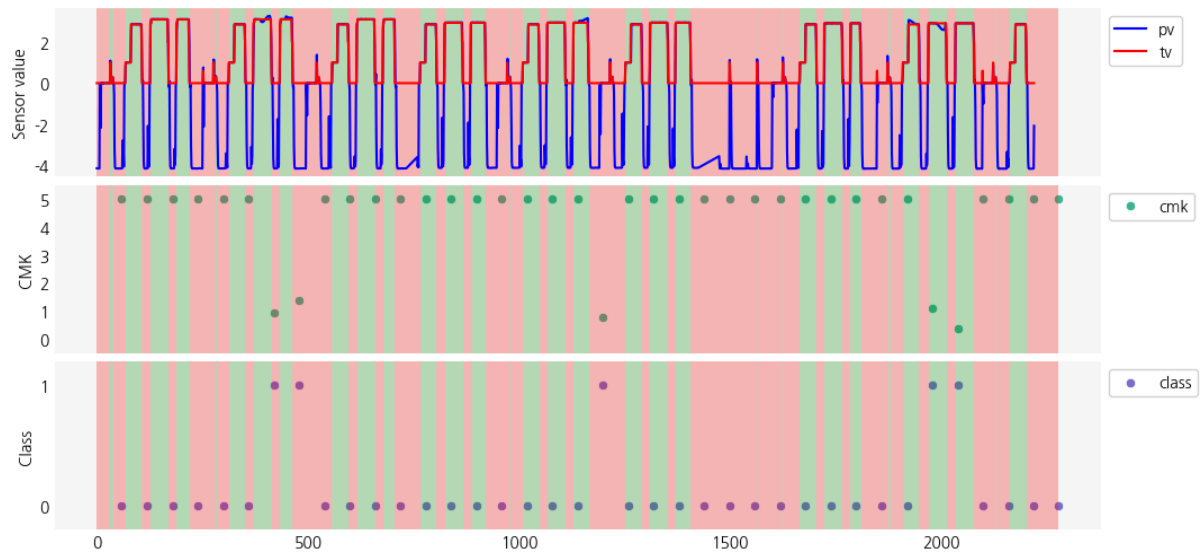


- Classification of anomaly patterns

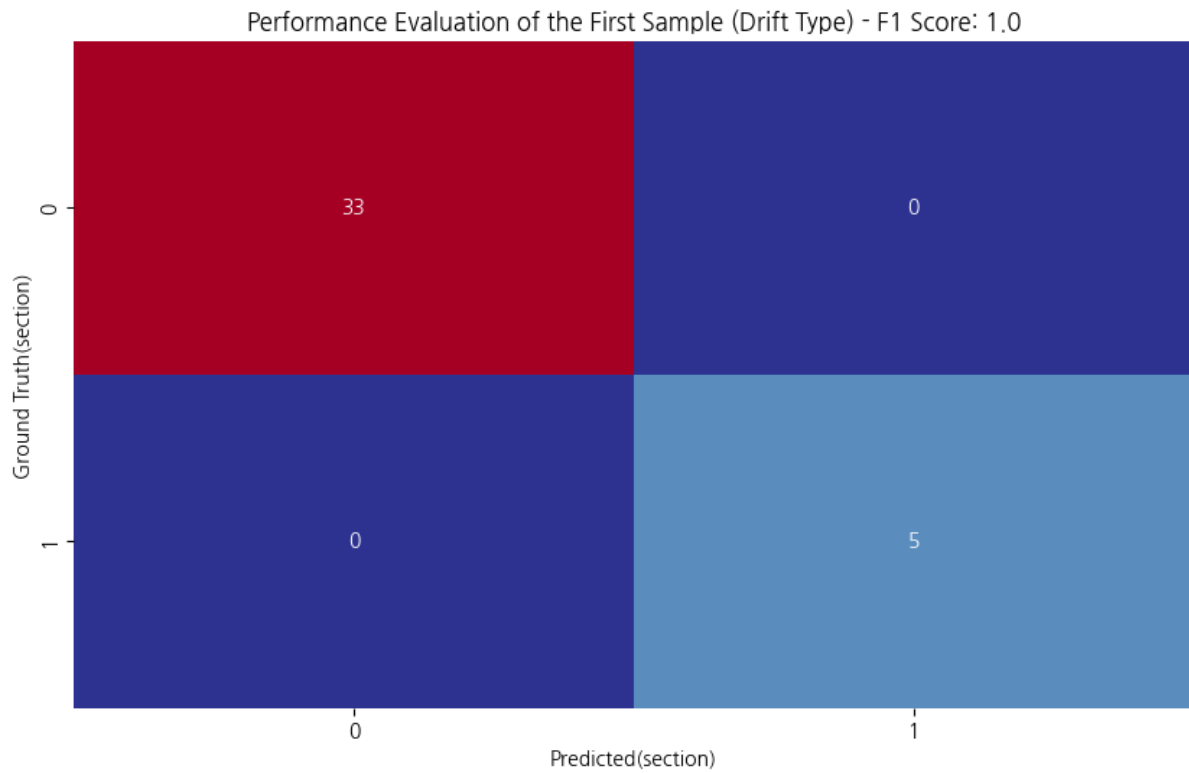


#2 시료 (Drift anomaly type)

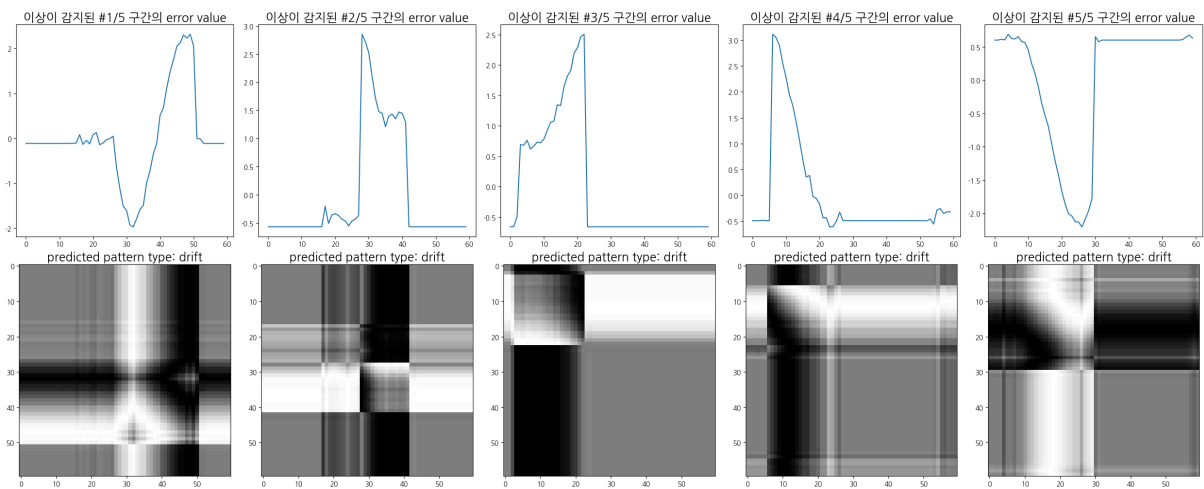
- Anomaly Detect



- Assessment (F1 score)

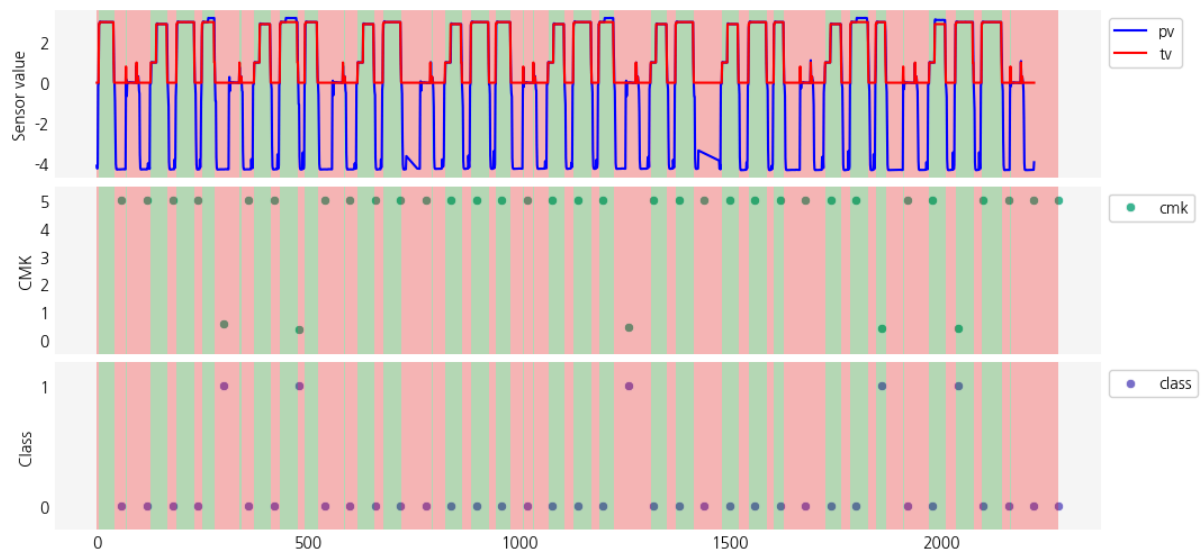


- Classification of anomaly patterns

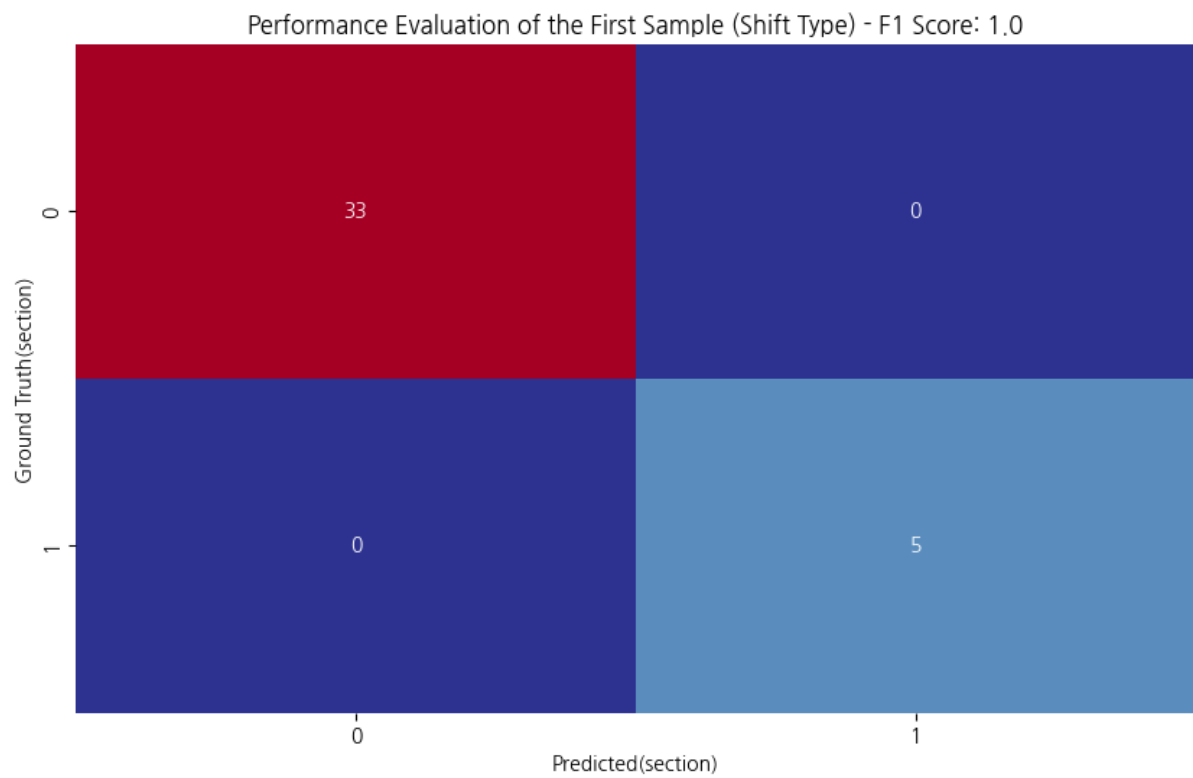


#3 시료 (Shift anomaly type)

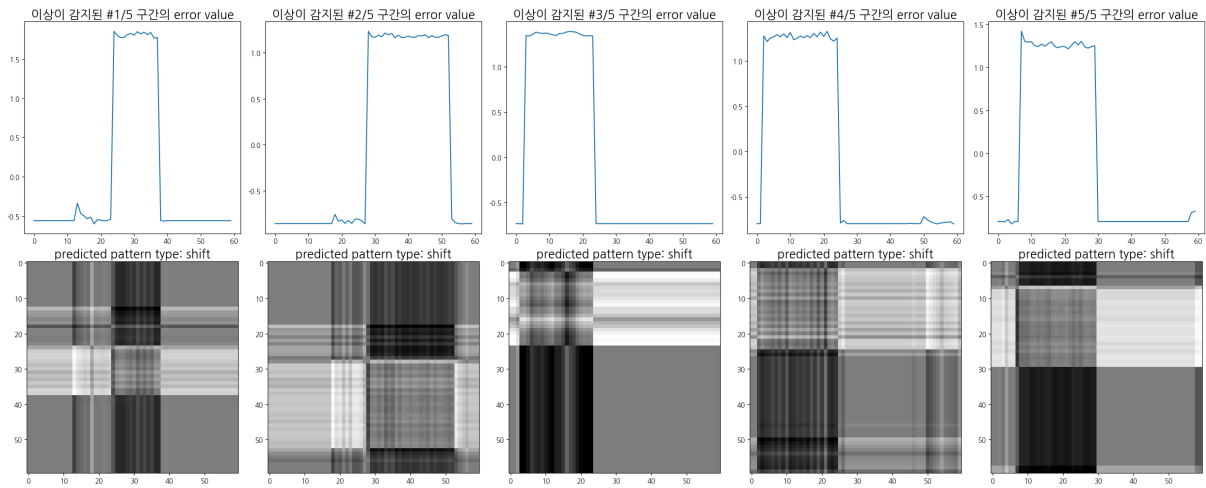
- Anomaly Detect



- Assessment (F1 score)

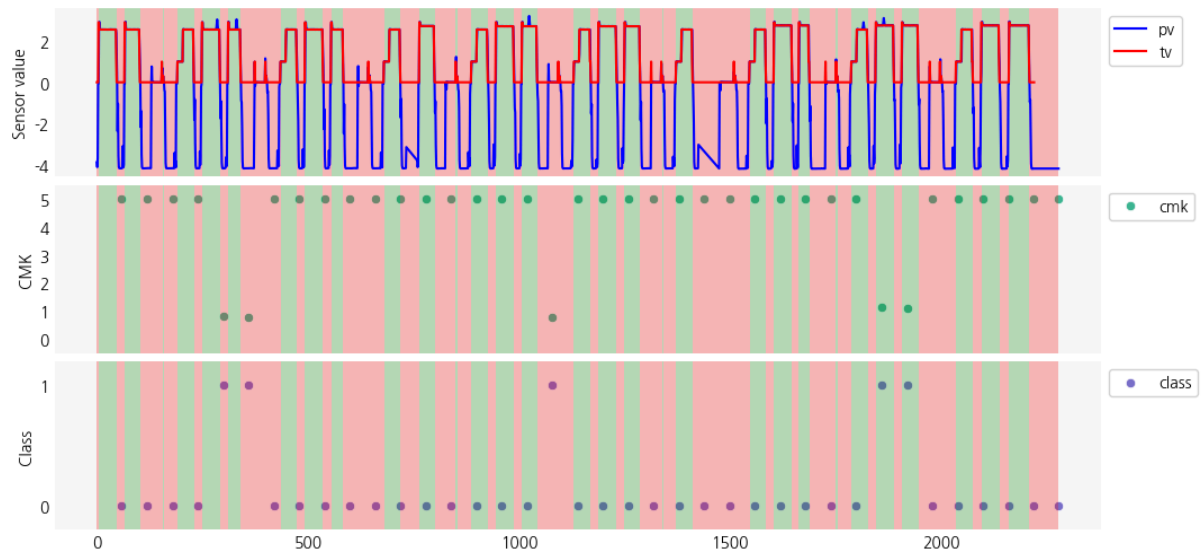


- Classification of anomaly patterns

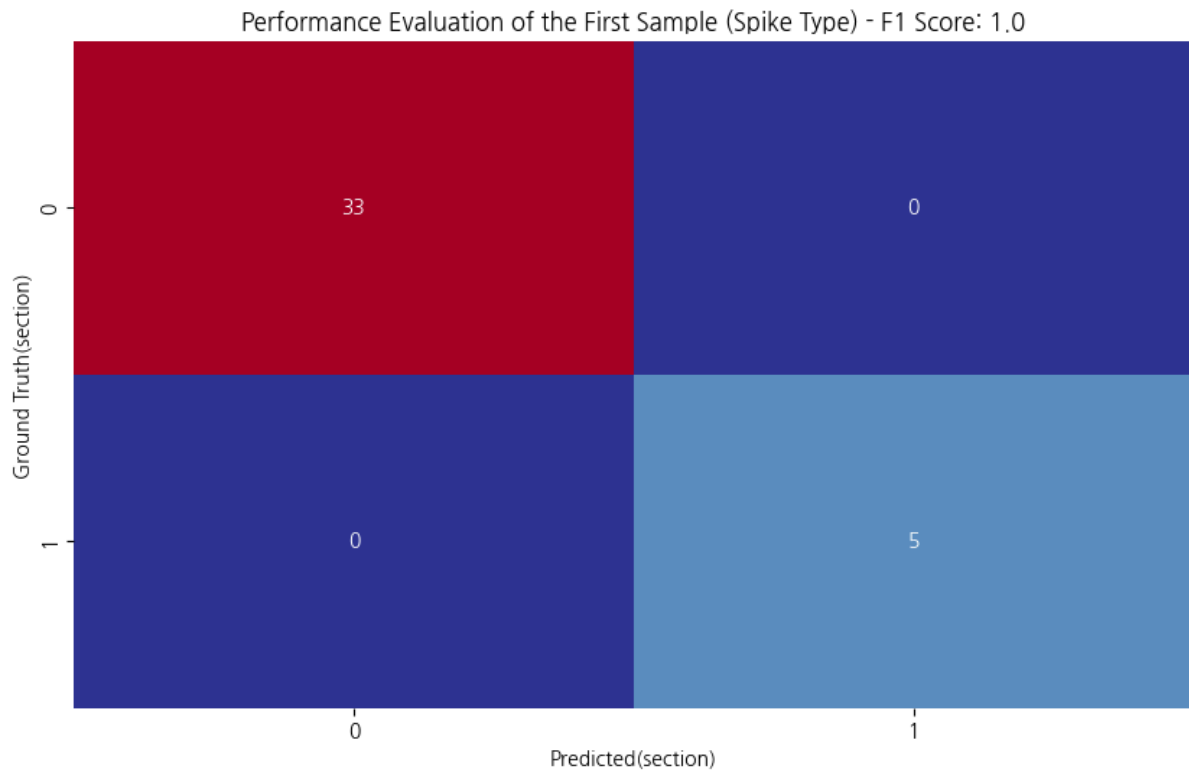


#4 시료 (Spike anomaly type)

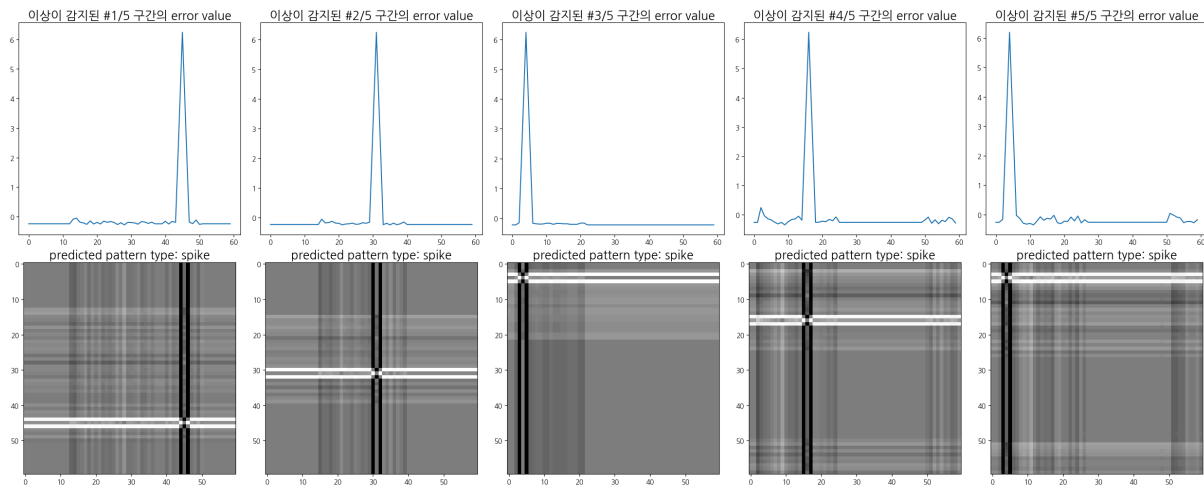
- Anomaly Detect



- Assessment (F1 score)

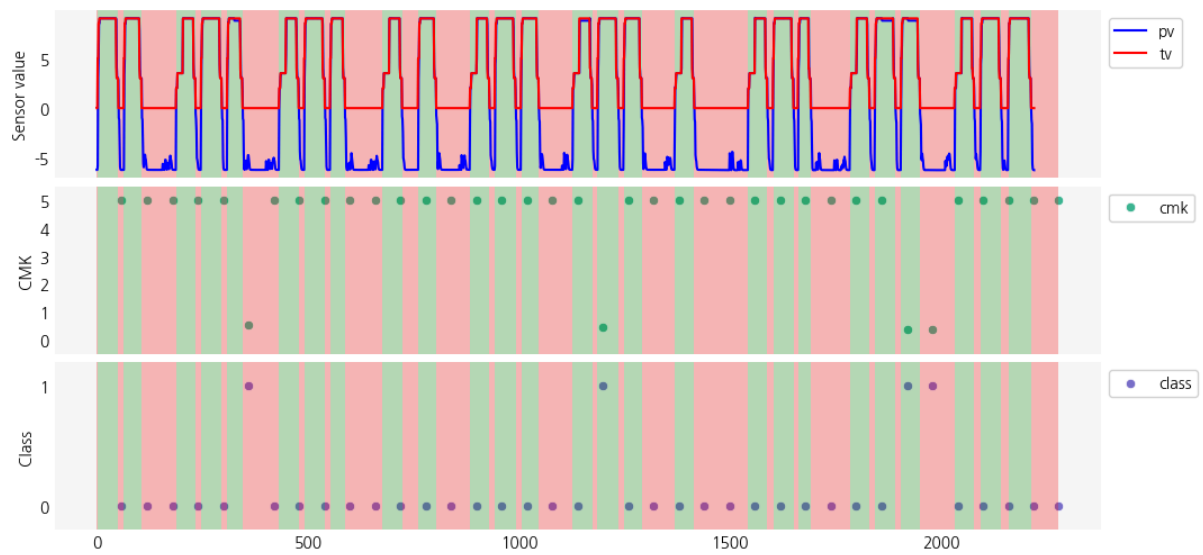


- Classification of anomaly patterns

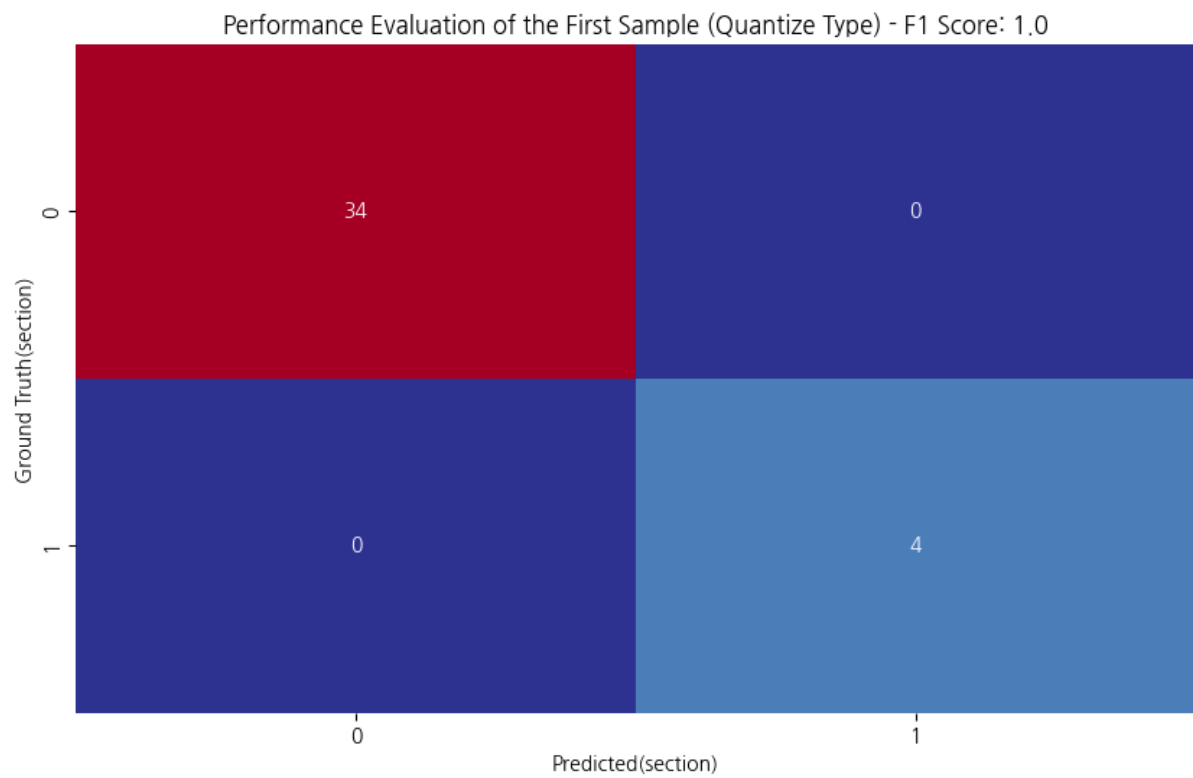


#5 시료 (Quantize anomaly type)

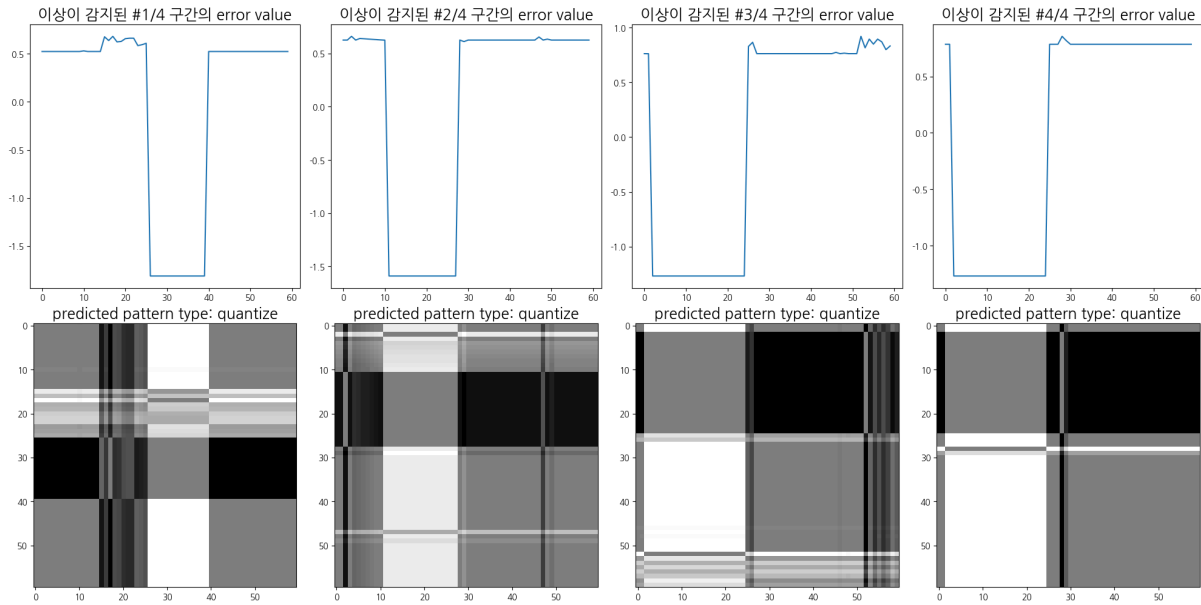
- Anomaly Detect



- Assessment (F1 score)



- Classification of anomaly patterns



Results

반도체 CMP(Chemical Mechanical Polishing) 공정의 정상 데이터를 기반으로, 5가지의 대표적인 이상 패턴을 서로 다른 센서 데이터에 적용하여 증강(Augmentation)된 5개의 샘플 데이터를 생성하였다. 이후, Rule-Based 이상 감지 모델을 활용하여 각 샘플 내에서 이상 발생 구간을 탐지하였으며, 모델의 성능을 평가하기 위해 각 샘플별 F1 Score를 성능 지표로 사용하였다.

본 실험을 통해, 각 이상 패턴이 적용된 샘플에서 Rule-Based 모델의 이상 탐지 성능을 정량적으로 분석하였으며, F1 Score를 기준으로 패턴별 탐지 성능의 차이를 비교하였다. 또한, 감지된 이상 구간에서 5가지 이상 패턴을 정확하게 분류할 수 있음을 확인하였다.

