

PyCon Korea 2019

Python을 활용한 데이터 분석 사례

- 지하철역 환기실 공조기 고장 감지 모델 -

- 국가수리과학연구소 김민중
mjgim@nims.re.kr

김민중

- **국가수리과학연구소 산업수학혁신센터**
 - 산업체 및 공공기관의 산업 문제를 발굴하고 수학적 해결 지원 업무 수행('16년 설립)
- **이학박사(확률론)**

Why Pycon & Why Python?

1. 문제 소개

- 기계설비 자동제어 빅데이터 분석시스템 고도화
- 공조기 및 전류 데이터 소개

2. 데이터 분석 사례(실적)

- 공조기 V벨트 탈락 및 슬립 감지 알고리즘
- 공조기 모터 주요 부품의 이상(상태) 감지

SAMBA



지하철 기계설비

1. 환기: 지하역사 신선한 공기 공급
2. 냉방: 시원한 공기 공급
3. 배수: 터널 내 지하수 외부 배출
4. 소방: 화재 감시 및 예방
5. 위생: 화장실, 급수
6. 승강: 승객 이동 편의시설
7. 자동제어: 기계설비 감시 및 제어

- 기계설비 자동제어 빅데이터 분석시스템(SAMBA) 고도화
 - 지하철 역사 내 기계설비의 고장을 감지하거나 사전 예측 모델
 - 공조기(송풍기) 이상 상황 감지 알고리즘 개발
 - 장비 노후로 인한 최근 3년간 가장 많은 고장
 - 고장 보수 시간 및 유지보수 비용 증가

공조기(Air handling unit)

- **승강장과 대합실 환기 설비**
- 급기 공조기: 외부의 공기를 내부로 공급
- 배기 공조기: 내부의 공기를 외부로 배출



공조기(Air handling unit)

- V벨트는 V자 모양의 벨트로, **모터의 동력을 팬(fan)에 전달**하는 구성품
- 벨트와 풀리 사이에 작용하는 마찰력을 높이기 위해 단면적이 넓은 V 모양



데이터

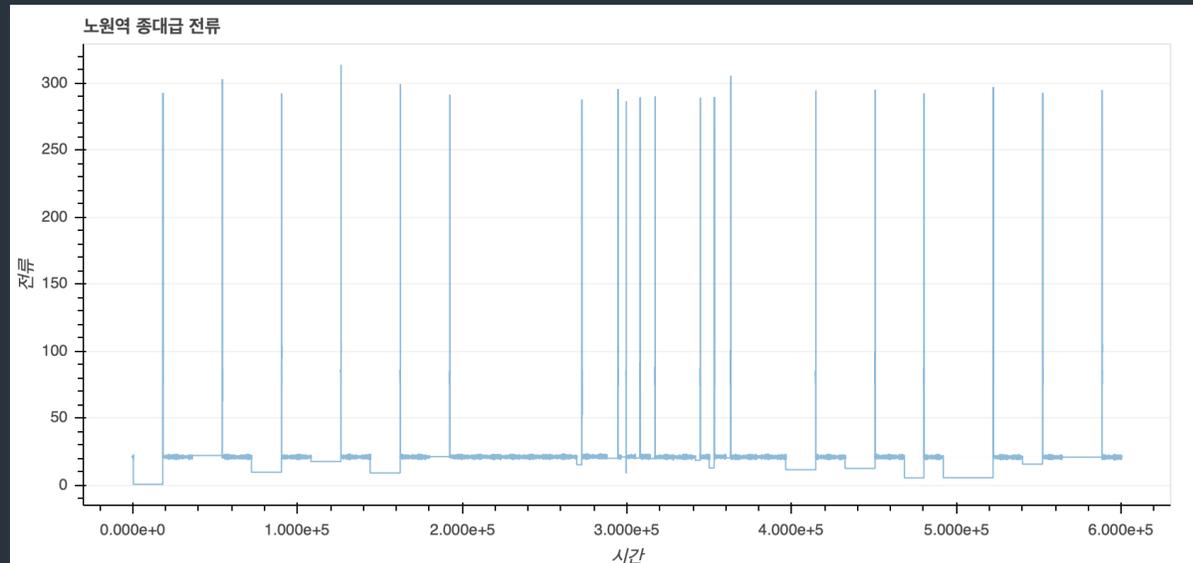
- 공조기 모터에서 수집된 작동 운전 전류(current) 데이터
- 회전력(torque): 물체를 회전시키는데 필요한 힘(kg-cm)

| | DateTime | Value |
|---|------------------------|-----------|
| 0 | 2019-01-29 오전 12:00:00 | 23.540001 |
| 1 | 2019-01-29 오전 12:00:01 | 23.469999 |
| 2 | 2019-01-29 오전 12:00:01 | 24.100000 |
| 3 | 2019-01-29 오전 12:00:02 | 23.290001 |
| 4 | 2019-01-29 오전 12:00:02 | 24.250000 |

- 0.5s 단위로 수집된 전류

노원역 종대급 전류 데이터 예시

- 스케줄에 의해서 자동으로 on/off 되는 방식
- 정지 상태에서 구동시 전류값이 크게 뛴



문제 정의

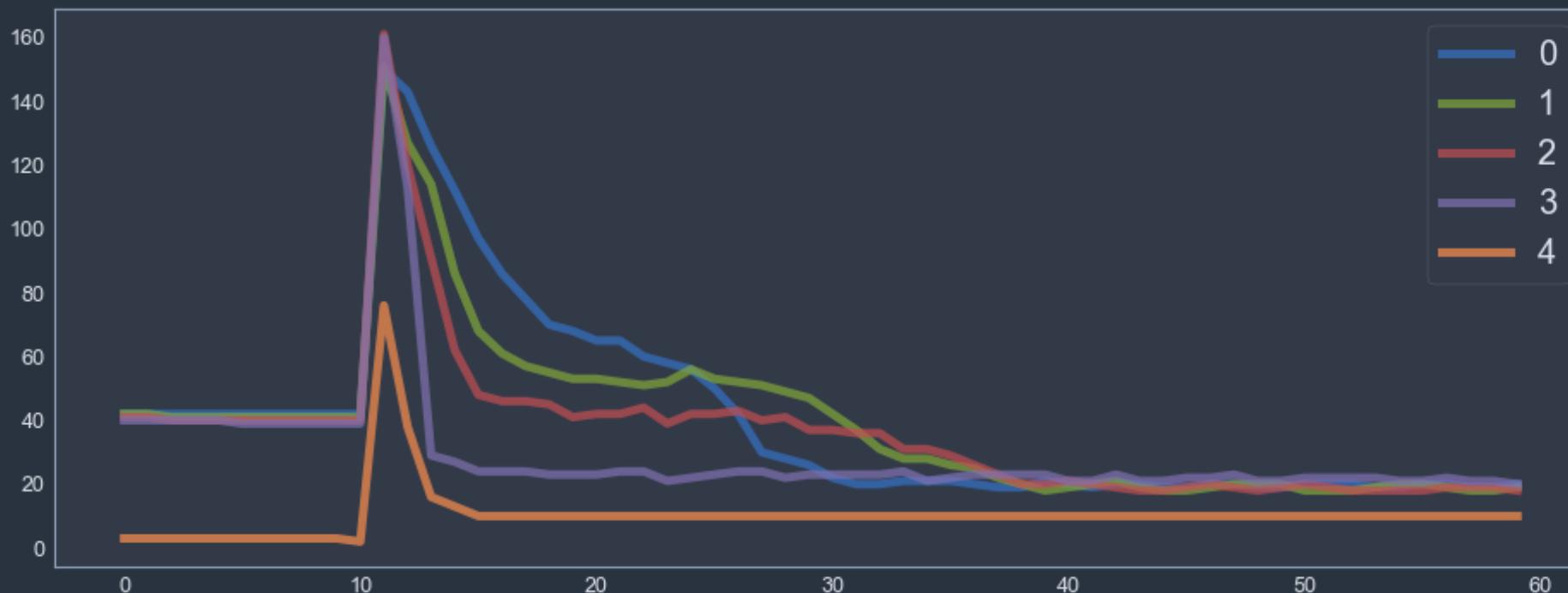
- **공조기 V벨트 탈락 및 슬립 감지 알고리즘**
 - 벨트 장력 부족 또는 저부하 운전으로 풍량 부족(효율 저하)
 - 정상 / 비정상 운전 시의 전류값 변화에 따른 고장 감지
 - V벨트를 한 줄씩 제거해서 만든 탈락(불량) 데이터를 활용
- **공조기 모터 주요 부품의 이상(상태) 감지(교체 시기 감지)**
 - V벨트, 베어링 교체 이전/이후 자동 분류 모델 개발
(샘플 장비 데이터 탐색 통해서 가능성 확인)

추진 실적

- **공조기 V벨트 탈락 및 슬립 감지 알고리즘**
 - 벨트 장력 부족 또는 저부하 운전으로 풍량 부족(효율 저하)
 - 정상 / 비정상 운전 시의 전류값 변화에 따른 고장 감지
 - V벨트를 한 줄씩 제거해서 만든 탈락(불량) 데이터를 생성

문제 해결 과정

- 공조기 모터 전류 데이터(작동 시점 기준)



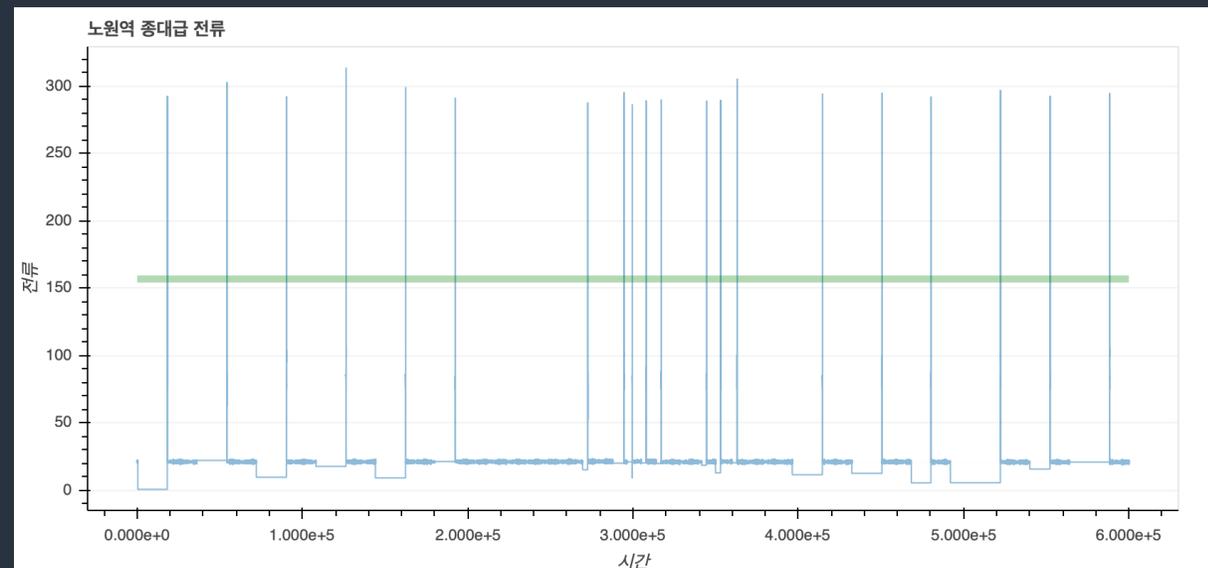
데이터 직접 수집

- 노원역 시점 환기실
 - 정상 가동 10회
 - V벨트 제거 후 10회



제시된 해결 방법

- 시계열 데이터 동기화 기법을 활용한 공조기 작동 시점을 기준으로 1차원 데이터 분할(segmentation)
 - 전류 최대값의 1/2 넘는 구간 임시 저장
 - 위 구간에서 시간 index 차이가 1보다 큰 index 기록(켜지는 시점 찾기)
 - 켜지는 시점 기준 \pm 일정 시간 기록



제시된 해결 방법

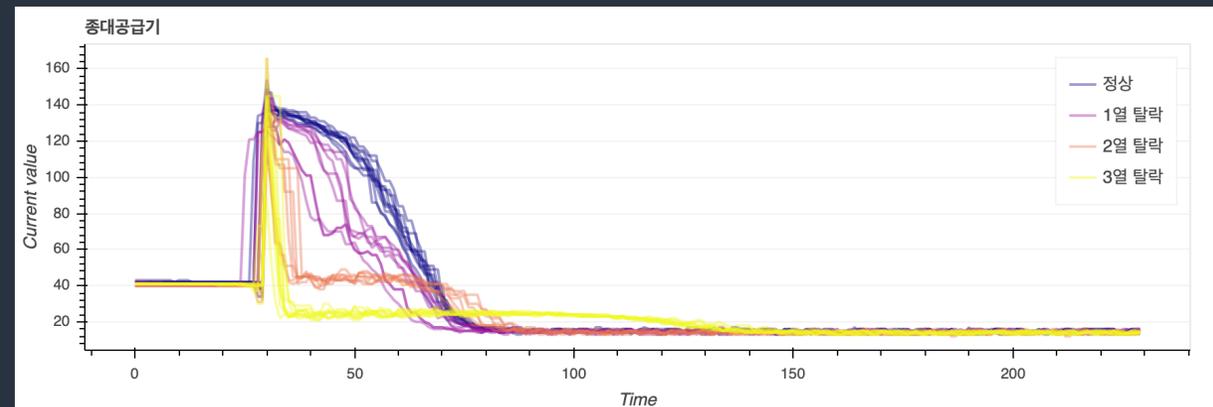
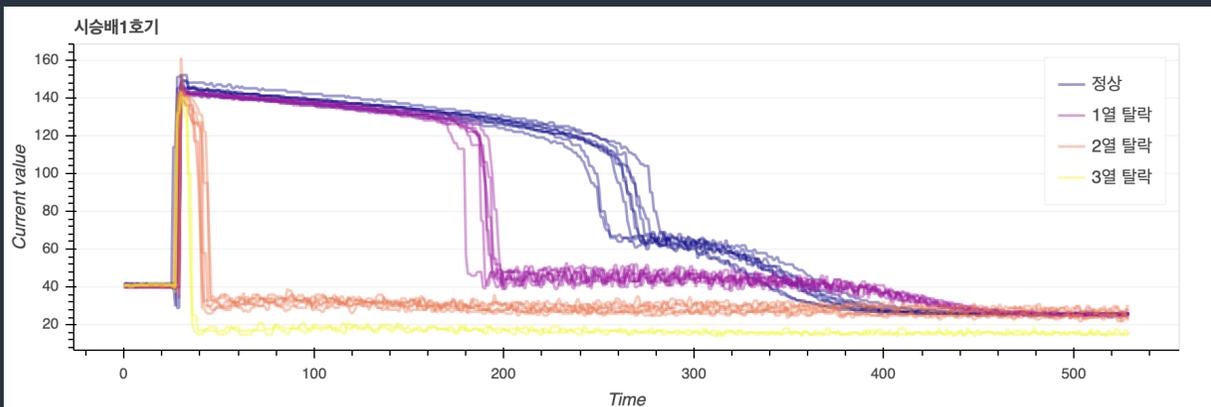
- 시계열 데이터 동기화 기법을 활용한 공조기 작동 시점을 기준으로 1차원 데이터 분할(segmentation)

- 전류 최댓값의 1/2 넘는 구간 임시 저장
- 위 구간에서 시간 index 차이가 1보다 큰 index 기록(켜지는 시점 찾기)
- 켜지는 시점 기준 \pm 일정 시간 기록

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | ... |
|---------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----|
| 2019-01-29 01:02:12 | 0.340000 | 0.340000 | 0.340000 | 0.340000 | 0.340000 | 0.340000 | 0.340000 | 40.450001 | 38.540001 | 38.470001 | ... |
| 2019-01-29 01:22:51 | 0.300000 | 0.300000 | 0.300000 | 0.300000 | 0.300000 | 0.300000 | 0.300000 | 0.300000 | 38.720001 | 38.369999 | ... |
| 2019-01-29 02:02:15 | 0.320000 | 0.320000 | 0.320000 | 0.320000 | 0.320000 | 0.320000 | 0.320000 | 42.570000 | 38.619999 | 39.180000 | ... |
| 2019-01-29 02:02:52 | 0.300000 | 0.300000 | 0.300000 | 0.300000 | 0.300000 | 0.300000 | 0.300000 | 0.300000 | 37.459999 | 37.990002 | ... |
| 2019-01-29 03:02:18 | 17.219999 | 17.219999 | 17.219999 | 17.219999 | 17.219999 | 17.219999 | 17.219999 | 17.219999 | 38.669998 | 39.099998 | ... |

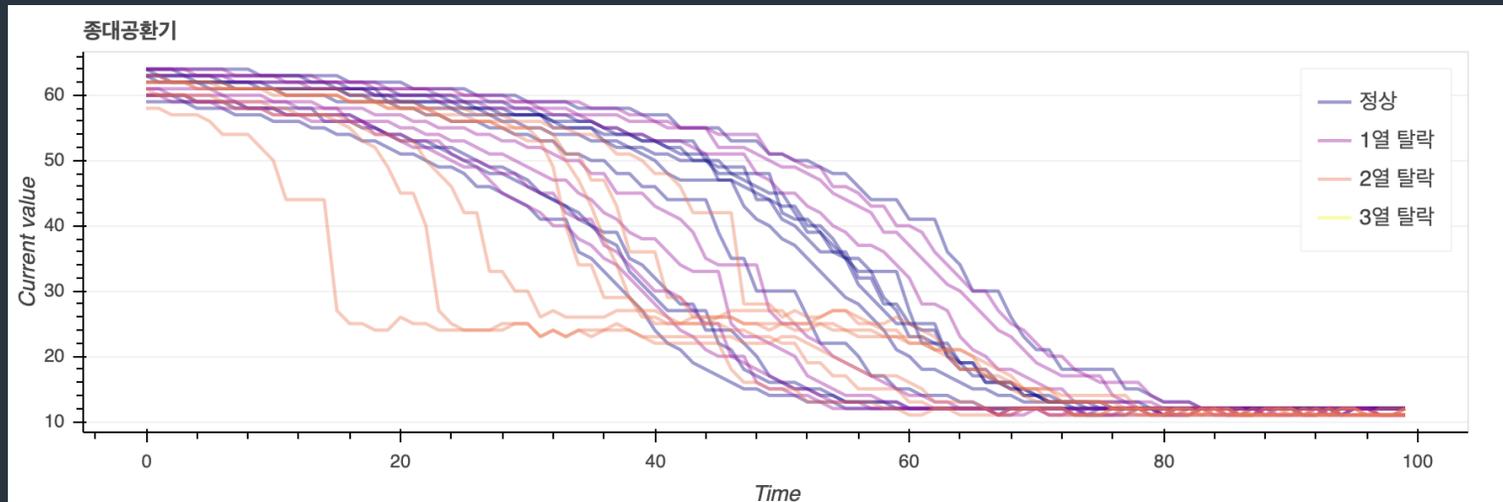
V벨트 탈락 감지 모델

- 시계열 데이터 동기화 기법을 활용한 공조기 작동 시점을 기준으로 1차원 데이터 분할(segmentation)



V벨트 탈락 감지 모델

- 다양한 공조기에 대한 **V벨트 탈락 감지 모델** 개발
- 가우시안 분포를 활용한 **이상 데이터 검출 모델**을 적용하여 비정상 데이터가 존재하지 않는 경우 이상 감지 (공조기 개수: 8천대)



추진 실적

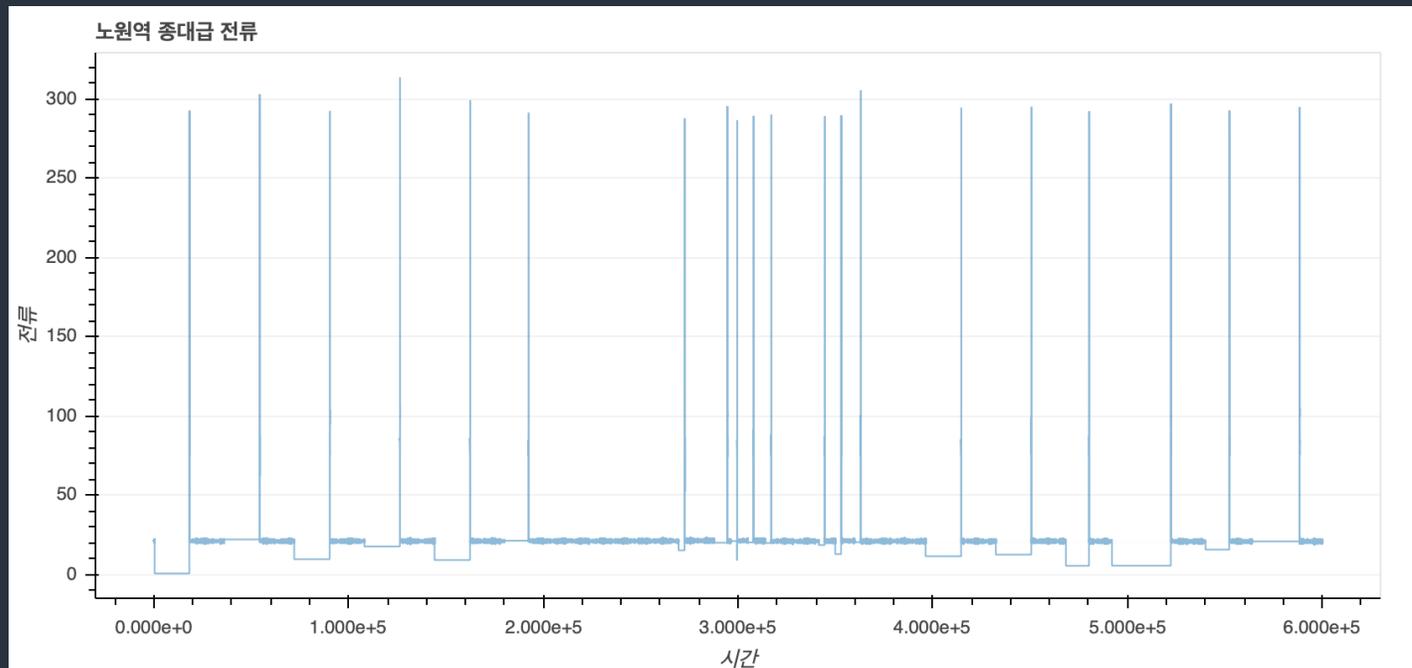
- **공조기 모터 주요 부품의 이상(상태) 감지(교체 시기 감지)**
 - V벨트, 베어링 교체 이전/이후 자동 분류 모델 개발
(샘플 장비 데이터 탐색 통해서 가능성 확인)

부품 이상 상태 감지 one stop 모델

- 공조기 작동 시점 기준으로 전류 데이터를 분할
- 교체 날짜 기준 이전과 이후 자동 라벨링
- 부품 교체 이전과 이후를 분류하는 모델 개발 및 적용

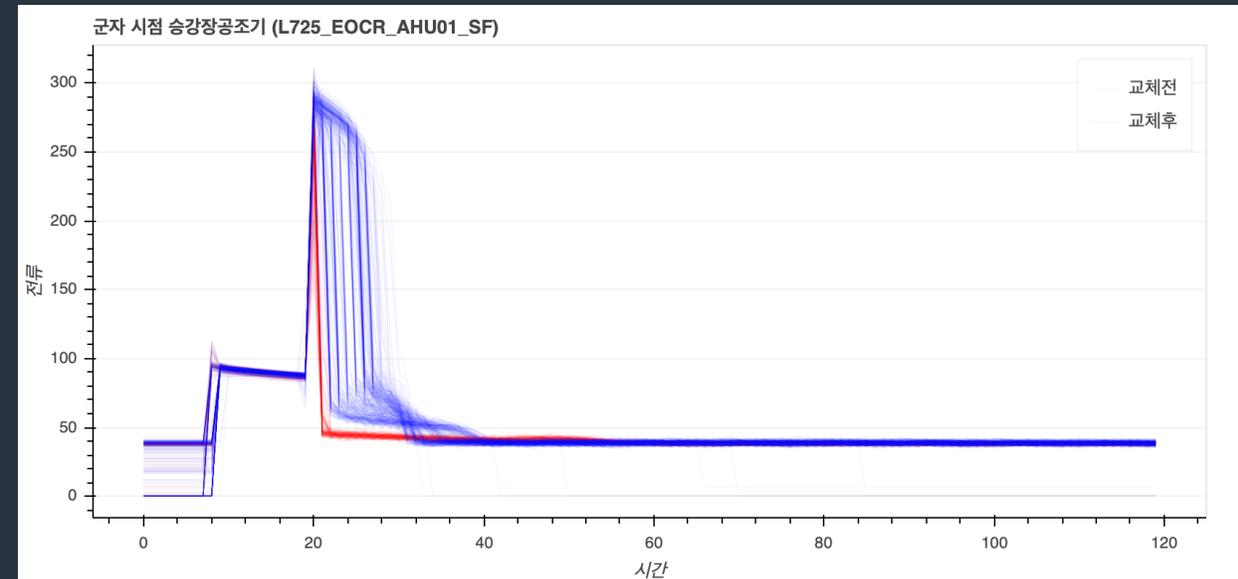
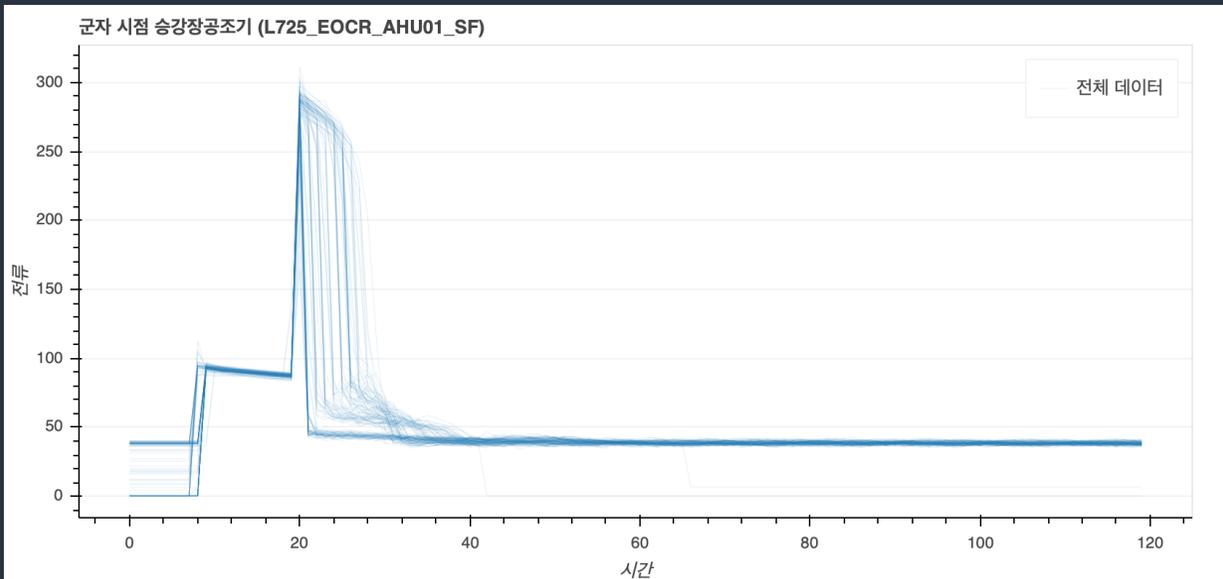
부품 이상 상태 감지 **one stop** 모델

- 공조기 작동 시점 기준으로 전류 데이터를 분할(기존 방법 고도화)



부품 이상 상태 감지 one stop 모델

- 교체 날짜 기준 이전과 이후 자동 라벨링



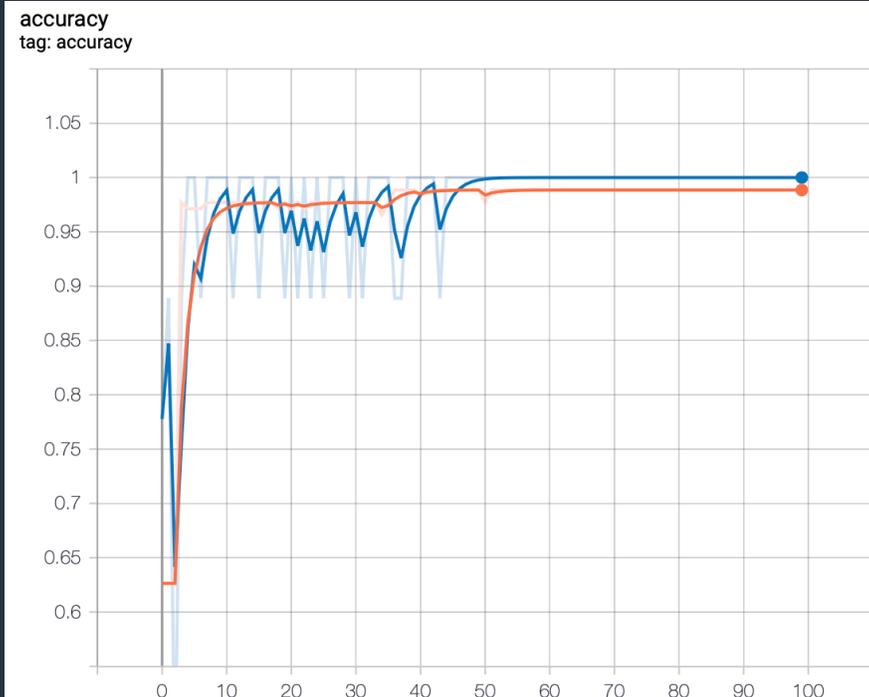
부품 이상 상태 감지 one stop 모델

- 부품 교체 이전과 이후를 분류하는 모델 개발 및 적용

accuracy_score: 0.9885057471264368

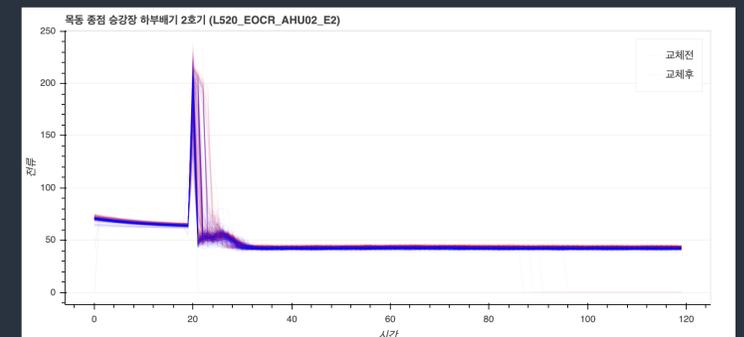
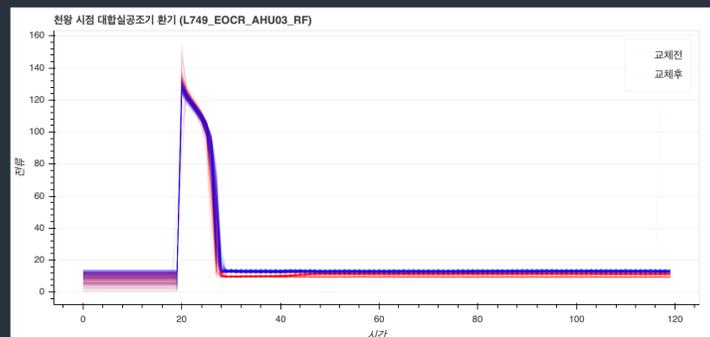
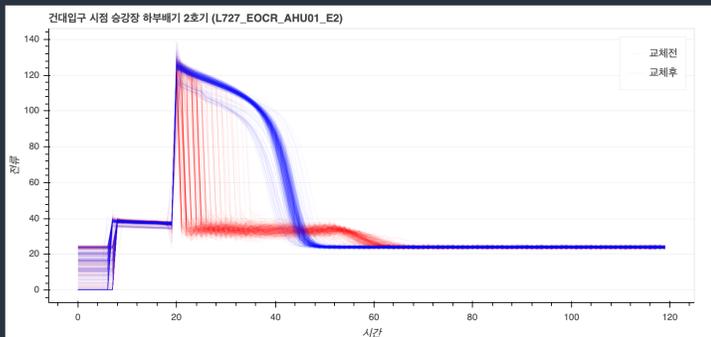
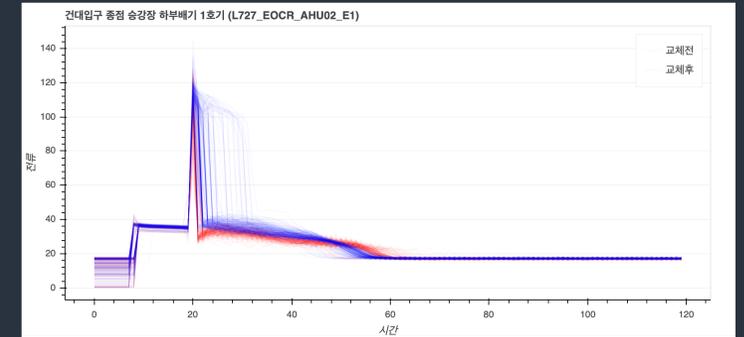
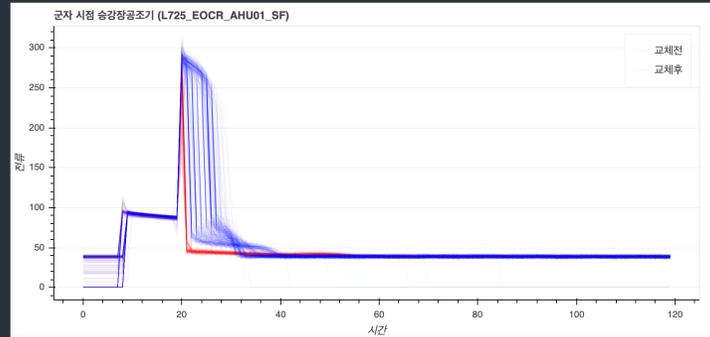
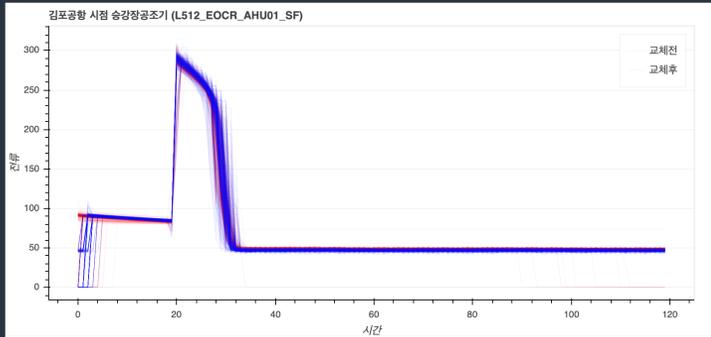
Report:

| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0.0 | 1.00 | 0.97 | 0.98 | 65 |
| 1.0 | 0.98 | 1.00 | 0.99 | 109 |
| accuracy | | | 0.99 | 174 |
| macro avg | 0.99 | 0.98 | 0.99 | 174 |
| weighted avg | 0.99 | 0.99 | 0.99 | 174 |



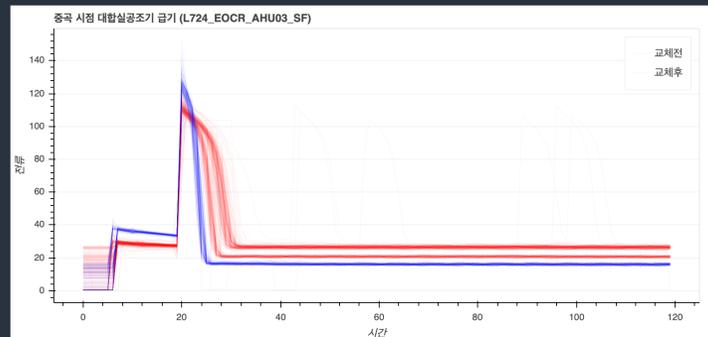
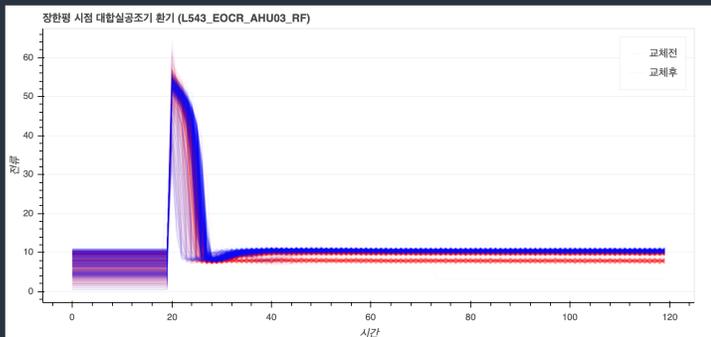
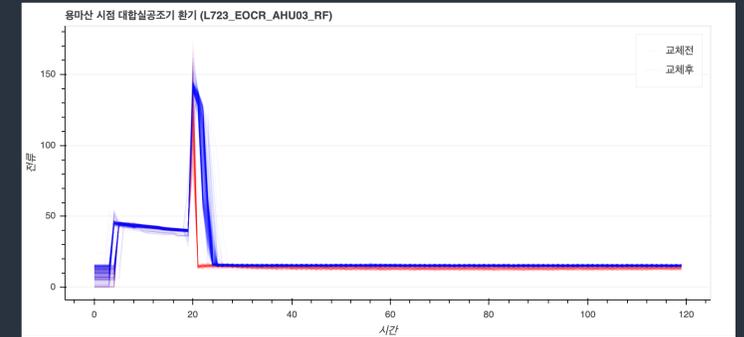
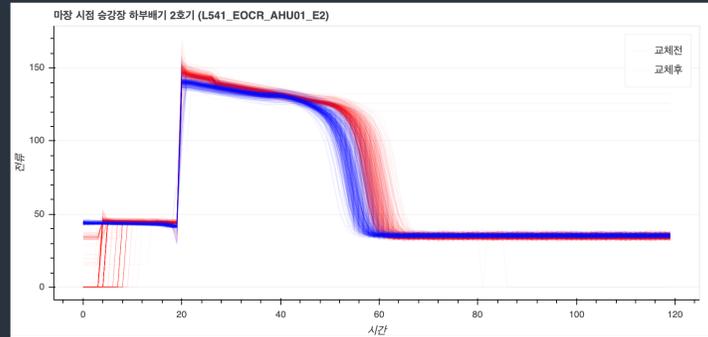
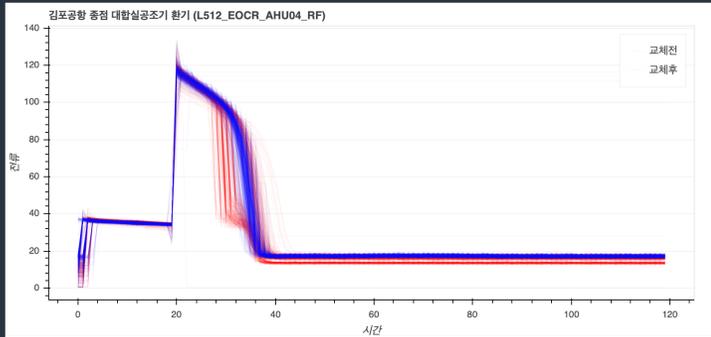
부품 이상 상태 감지 one stop 모델

- 자동 이전/이후 라벨링(V벨트)



부품 이상 상태 감지 one stop 모델

- 자동 이전/이후 라벨링(베어링)



정확도(V벨트)

| 공조기 | 교체 전(개수) | 교체 후(개수) | 교체날짜 | 정확도 | 정확도* |
|---------------------|----------|----------|------|------|------|
| 건대입구 시점승강장 하부배기 2호기 | 461 | 515 | 2/28 | 100% | 99% |
| 건대입구 종점승강장 하부배기 1호기 | 282 | 352 | 3/14 | 97% | 95% |
| 군자 시점 승강장공조기 | 262 | 434 | 3/22 | 100% | 97% |
| 김포공항 시점 승강장공조기 | 705 | 800 | 2/28 | 97% | 96% |
| 목동 종점 승강장 하부배기 2호기 | 576 | 437 | 3/30 | 85% | 82% |
| 상봉 B2 종점 대합실공조기 환기 | 547 | 636 | 3/30 | 98% | 93% |
| 천왕 시점 대합실공조기 환기 | 483 | 323 | 2/28 | 97% | 93% |

정확도(베어링)

| 공조기 | 교체 전(개수) | 교체 후(개수) | 교체 날짜 | 정확도 | 정확도* |
|--------------------|----------|----------|--------|------|------|
| 김포공항 종점 대합실공조기 환기 | 731 | 781 | 2월 28일 | 93% | 91% |
| 마장 시점 승강장 하부배기 2호기 | 700 | 727 | 2월 28일 | 100% | 99% |
| 용마산 시점 대합실공조기 환기 | 163 | 647 | 3월 22일 | 100% | 100% |
| 장한평 시점 대합실공조기 환기 | 1215 | 1048 | 2월 28일 | 97% | 96% |
| 중곡 시점 대합실공조기 급기 | 364 | 163 | 1월 31일 | 100% | 100% |

1d CNN 모델

| Layer | Input shape | Output shape | Param # |
|------------------------------------|----------------|----------------|-------------------|
| Reshape | (None, 120) | (None, 120, 1) | 0 |
| Conv1D | (None, 120, 1) | (None, 118, 8) | 32 |
| MaxPooling | (None, 118, 8) | (None, 59, 8) | 0 |
| BatchNormalization | (None, 59, 8) | (None, 59, 8) | 32 |
| Conv1D | (None, 59, 8) | (None, 57, 8) | 200 |
| MaxPooling | (None, 57, 8) | (None, 29, 8) | 0 |
| BatchNormalization | (None, 29, 8) | (None, 29, 8) | 32 |
| Flatten | (None, 29, 8) | (None, 232) | 0 |
| Dense | (None, 232) | (None, 8) | 1864 |
| Dense | (None, 8) | (None, 8) | 72 |
| Dense | (None, 8) | (None, 1) | 9 |
| Total(Trainable) parameters | | | 2241(2209) |

- Filter size: (3,1)
- Cost function: Cross entropy function
- Optimizer: Adam
- Learning rate: 1e-4
- Epoch: 300
- Batch size: 32

| Layer | Input shape | Output shape | Param # |
|------------------------------------|---------------|---------------|-----------------|
| Reshape | (None, 80) | (None, 80, 1) | 0 |
| Conv1D | (None, 80, 1) | (None, 78, 4) | 16 |
| MaxPooling | (None, 78, 4) | (None, 39, 4) | 0 |
| BatchNormalization | (None, 39, 4) | (None, 39, 4) | 16 |
| Conv1D | (None, 39, 4) | (None, 37, 4) | 52 |
| MaxPooling | (None, 37, 4) | (None, 19, 4) | 0 |
| BatchNormalization | (None, 19, 4) | (None, 19, 4) | 16 |
| Flatten | (None, 19, 4) | (None, 76) | 0 |
| Dense | (None, 76) | (None, 4) | 308 |
| Dense | (None, 4) | (None, 1) | 5 |
| Total(Trainable) parameters | | | 413(397) |

- Filter size: (3,1)
- Cost function: Cross entropy function
- Optimizer: Adam
- Learning rate: 5e-5
- Epoch: 300
- Batch size: 32

감사합니다.