Deep Learning

이상 탐지

(Anomaly Detection)

강사 양석환

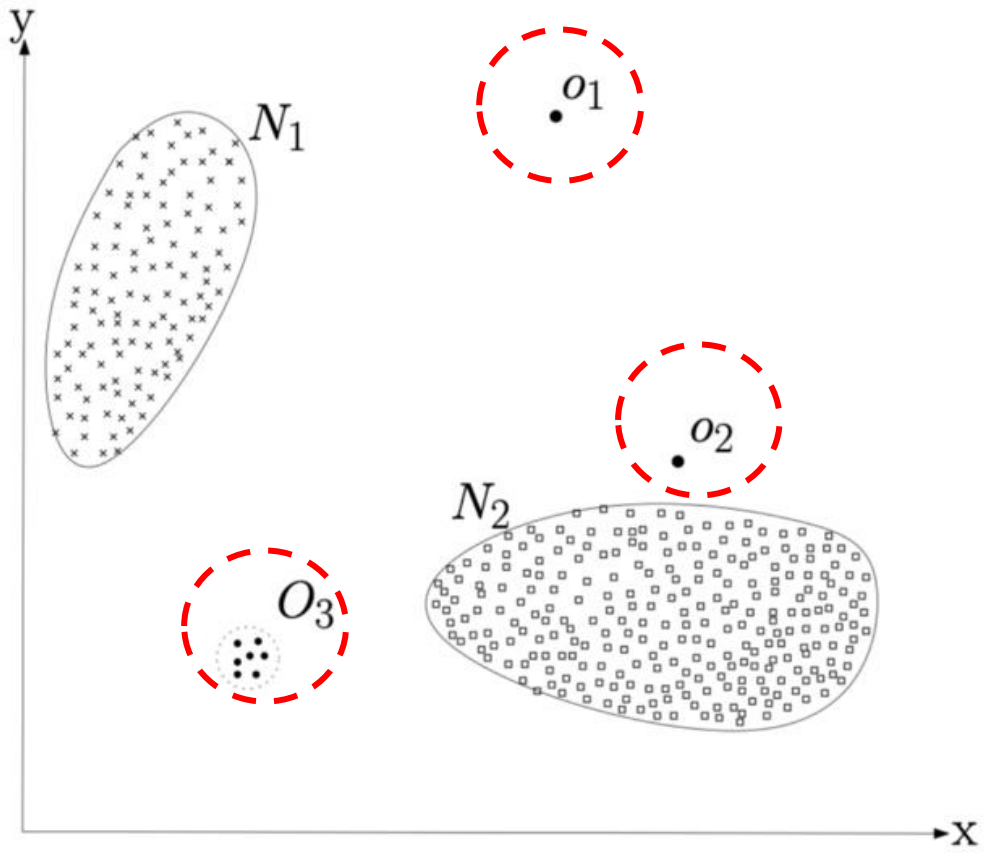
이상 탐지 개요

이상 탐지 (Anomaly Detection)

- 이상 탐지
  - 어떤 특정한 도메인에서 일반적으로 예상되는 특성을 따르지않는 데이터나, 정상(normal)으로 규정된 데이터와 다른 특징을 가지는 데이터를 찾아내는 것
  - 이상 탐지 기법을 써서 찾아내는 데이터
    - 불량, 오류, 악성코드, 가짜 데이터, 예외, 노이즈, 새로운 패턴 등의 데이터
    - 일반적으로 데이터셋에서 작은 비율의, 다른 데이터의 범주와 확연히 구분되는 데이터를 찾고 싶을때 그것을 outliers 혹은 anomalies라고 하며 이상 탐지 문제로 정의함

이상 탐지 (Anomaly Detection)

- **이상 데이터의 예**



- 2차원에서의 anomalies의 예.
- o1, o2 ,O3는 정상 영역 N1과 N2에서 떨어져 있다.

이상 탐지 (Anomaly Detection)

- 이상 탐지의 적용 사례
  - 신용카드 거래 내역을 조사하여 일반적이지 않은 패턴을 발견하여 신용카드 사기 행위를 잡아내는 문제
  - 제조업에서 공정 데이터를 분석하여 양품/불량을 분류하거나 공정 장비의 고장을 찾아내는 것
  - 의학 분야에서 영상, 뇌파 신호등의 의학 데이터을 통해 비정상적인 환자 상태를 탐지하여 진단을 내리는 것

머신러닝/딥러닝 기반의 이상 탐지

AiDA Lab.

Supervised anomaly detection

- 학습 데이터의 정상 데이터와 이상 데이터에 라벨링을 할 수 있다고 가정하면

  → 이상 탐지 문제를 지도 학습의 분류 문제로 간주할 수 있음

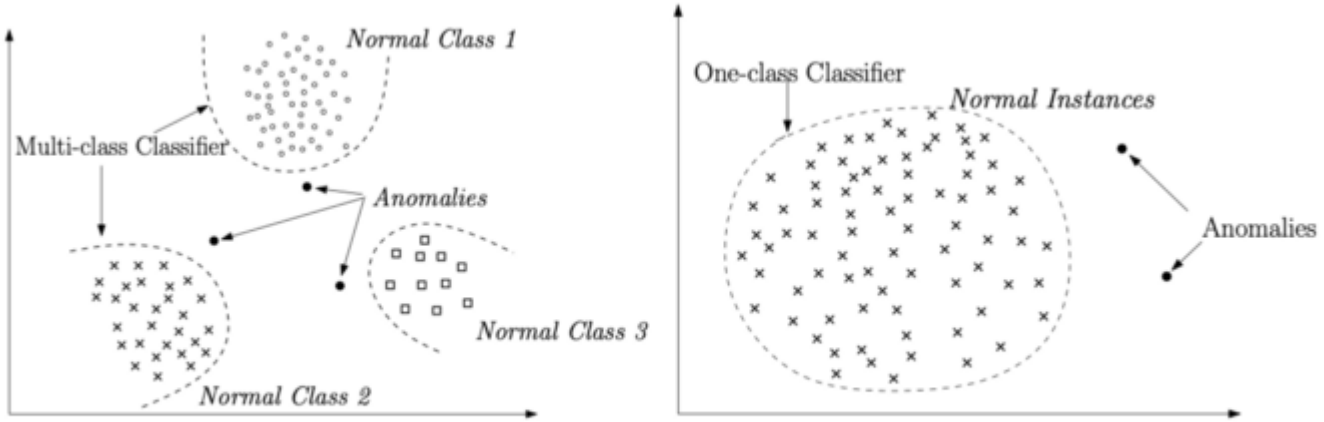
- 지도학습을 이용한 이상 탐지는 비교적 정확도가 높다는 것이 장점

- 이상 탐지 문제에서는 Imbalanced data가 자주 발생하는데, 이것은 분류 문제를 해결하는데 있어 큰 장애물이 됨

  - Imbalanced Data: 즉, 비정상적인 데이터가 정상 데이터에 비해 훨씬 적은 경우
  - 예: 제조업의 경우, 하나의 불량 데이터를 얻기 위해 최소 몇 백개에서 몇 천, 몇 만개의 양품 데이터를 얻어야 하는 경우

Supervised anomaly detection

- 분류 알고리즘을 성공적으로 학습시킬 수 있는 양의 비정상 데이터를 확보하기 위해서는 매우 거대한 사이즈의 데이터셋이 필요할 수 있음
- 이러한 데이터 불균형 문제를 해결하기 위해서 Over Sampling, Under Sampling, Reconstruction Loss 등 다양한 전략들이 연구 중



Classification을 이용한 Anomaly Detection

Supervised anomaly detection

- 지도학습 기반의 이상탐지 기법들
  - Support vector machine(SVM), Bayesian network, K-nearest neighbors 등

- 지도학습 기반의 이상탐지의 문제점
  - 이상 탐지 문제를 지도학습으로 해결하는 것은 많은 장점이 있지만 정상과 이상데이터를 정확히 구분하고 라벨링하는 것은 매우 어려움
  - 정상/이상 데이터 라벨링은 매우 시간이 걸리는 작업일 뿐만 아니라 정상과 이상의 경계를 나누는 것이 분명하지않을 때가 많음
  - 또한 데이터 불균형 문제도 상주

Unsupervised anomaly detection

- 비지도학습은 unlabeled data에서 데이터의 구조와 패턴을 알아내는 문제의 범주에 속함
- 데이터가 라벨링이 되어있지 않거나, 정상/이상에 대한 라벨링이 힘들 것으로 예상될 때 적용할 수 있고, 특히 데이터에 이상치가 매우 적게 나타날 때 유용한 기법

- 비지도 학습 기반의 이상탐지 기법들
  - Clustering, Principal Component Analysis (PCA), Autoencoder 등

Unsupervised anomaly detection

- **Clustering**
  - 대표적인 비지도학습 알고리즘
  - 유사한 데이터를 cluster로 그룹화하는 기법. 거리, 밀도 등의 다양한 기준을 사용하여 데이터를 그룹화 함
  - point anomalies를 탐지할 때 유용
  - Clustering을 이용해 anomalies를 찾아내는 방식은 Clustering의 결과에 따라 anomalies를 해석하는 방식이 여러가지이기 때문에 하나로 고정되어 있지 않음
    - 어떠한 cluster에도 속하지 않는 instance를 anomaly로 규정할 수도 있고
    - cluster의 중심에 가까울 수록 정상 데이터, 중심에서 멀어질 수록 이상 데이터일 확률이 커진다고 볼 수 있으며
    - 좀 더 크기가 크고 밀도가 높은 cluster를 정상 cluster, 크기가 작고 sparse한 cluster를 이상 cluster로 정의하기도 함

Unsupervised anomaly detection

- 합의된 해석 방식이 없는 것은
  - anomalies의 개념이 포괄적이며
  - 도메인 마다 anomalies의 정의가 다르기 때문
- Clustering을 이용해서 이상 탐지 문제를 풀 경우에는 이러한 anomalies에 대한 가정 중에서 문제를 가장 잘 설명하는 적절한 것을 선택하여 적용함

Unsupervised anomaly detection

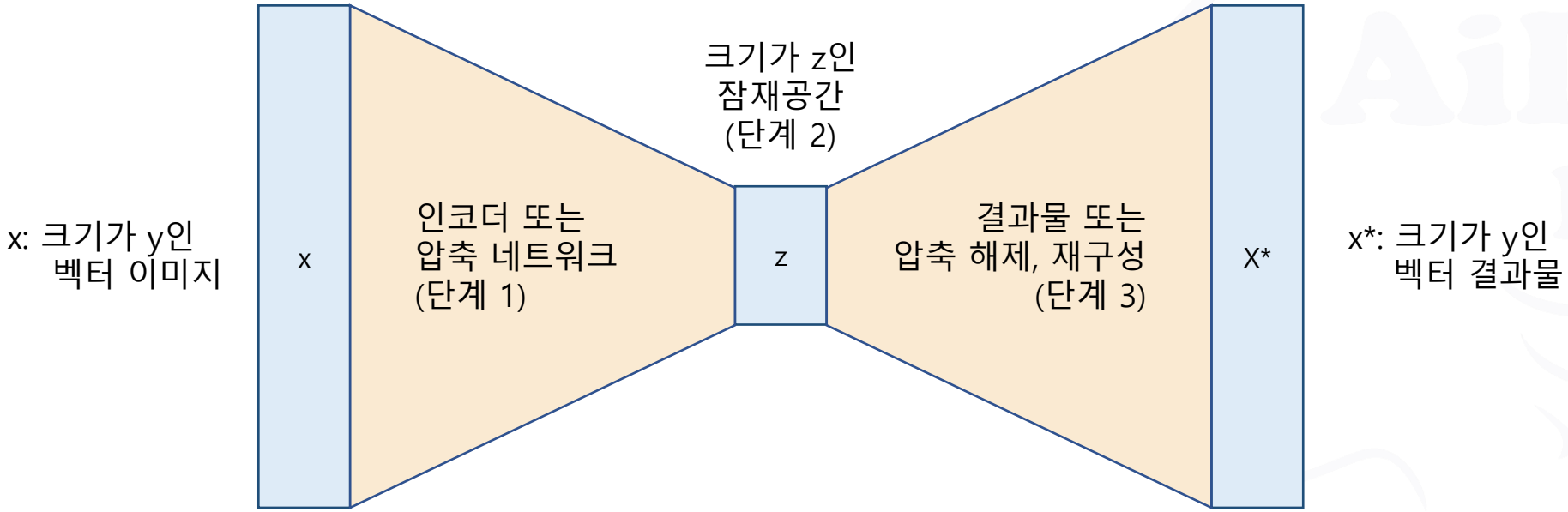
- PCA (Principal component analysis, 주성분 분석)
  - 고차원의 데이터를 저차원의 데이터로 환원시키는 기법
  - 서로 연관 가능성이 있는 고차원 공간의 표본 → 선형 연관성이 없는 저차원 공간(주성분)의 표본으로 변환
  - 데이터를 한 개의 축으로 사상 시켰을때 그 분산이 가장 커지는 축을 첫 번째 주성분, 두 번째로 커지는 축을 두 번째 주성분으로 놓이도록 새로운 좌표계로 데이터를 선형 변환
  - 첫째 주성분이 가장 큰 분산을 가지고, 이후의 주성분들은 이전의 주성분들과 직교한다는 제약 아래에 가장 큰 분산을 갖고 있다는 식으로 정의됨
  - 중요한 성분들은 공분산 행렬의 고유 벡터이기 때문에 직교함

Unsupervised anomaly detection

- 이와 같이 표본의 차이를 가장 잘 나타내는 성분들로 분해함으로써 데이터 분석에 여러가지 이점을 제공함
- PCA를 이용한 이상 탐지 → 데이터의 압축/복원 과정을 학습하면서 정상 데이터의 패턴을 학습하여 비정상 데이터를 검출

Unsupervised anomaly detection

- **오토 인코더 (Auto Encoder)**
  - **입력을 인코딩함으로써 압축한 정보를 잠재공간에 저장하고 이를 다시 디코딩하여 원본과 가깝게 복원하는 방식의 모델**
  - **인코더와 디코더로 구성됨**



Unsupervised anomaly detection

- 오토 인코더의 동작
  - 이미지 x를 오토 인코더에 입력
  - 재구성된 이미지 x*를 획득
  - x와 x*의 차이인 재구성 손실 측정
    - x와 x* 의 픽셀 간 거리(예, 평균 제곱 오차)를 함수에 사용
    - 경사 하강법 적용

Unsupervised anomaly detection

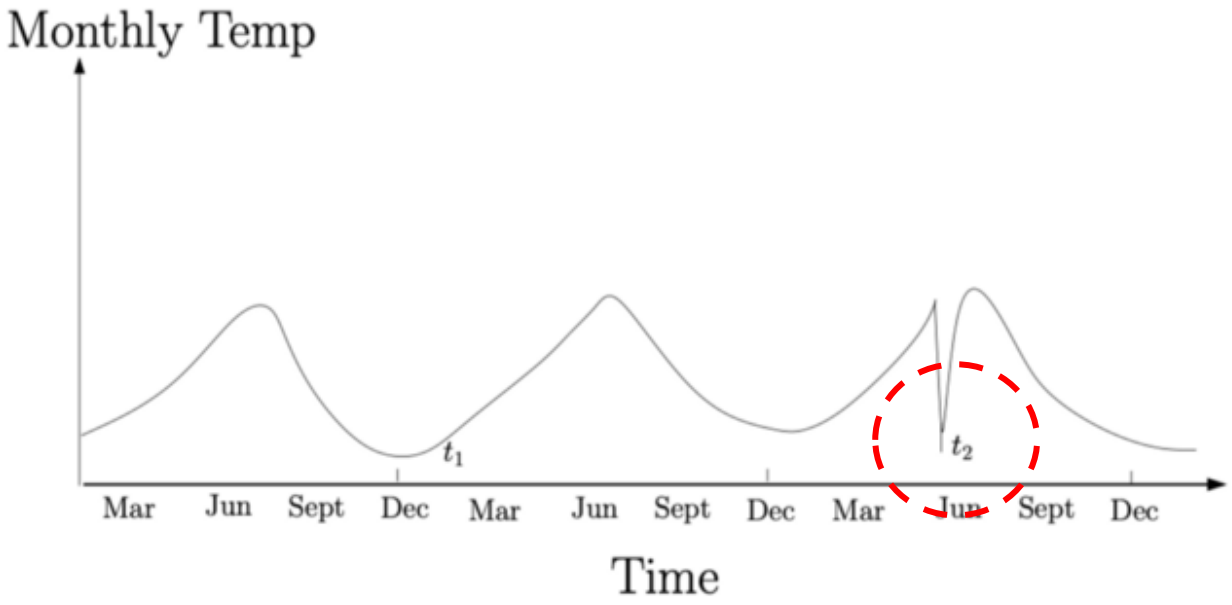
- **오토 인코더의 사용 방법**
  - 정상 데이터셋, 혹은 거의 대부분이 정상 데이터로 여겨지는 데이터셋을 오토 인코더 모델에 학습시킴
  - 알고리즘이 정보의 압축, 복원을 반복하게 되면 모델은 데이터를 실제와 가깝게 복원하기위해 정상 데이터의 패턴과 특징을 학습함
  - 최종적으로 학습이 잘 된 모델에
    - 정상 데이터를 입력하면 → 출력 값으로 원래 입력과 거의 비슷한 값을 도출해 냄
    - 비정상, 특이점을 가진 데이터를 입력하면 → 모델은 학습되지 않은 비정상 데이터를 복원하는 패턴을 배우지 않았음
      → 출력 값은 입력 값과 큰 차이가 나게 됨
      → 따라서 입력 값과 출력 값의 차이에 대한 기준을 설정해서, 그 기준보다 차이가 나는 데이터를 anomalies로 검출해 낼 수 있음

시계열 데이터의 이상 탐지

- 현실 세계에서 이상 탐지가 가장 많이 적용되는 데이터 타입중의 하나
- 다양한 Anomaly의 유형
  - Point Anomalies:
    - 단순히 개별적인 데이터 instance가 다른 데이터의 범주와 많이 벗어나는 것
  - Contextual Anomalies:
    - instance 자체로는 데이터의 값의 범주에서 벗어나지 않지만 특정한 context에서 변칙적인 것으로 여겨질 때
    - 실제 시계열 데이터에서 이상의 유형으로 contextual anomalies가 자주 발견되는데, 이때 ‘context’ 는 보통 시간적인 특성을 의미

시계열 데이터의 이상 탐지

- **시계열 이상 탐지의 예**
  - **기온에 대한 이상 탐지**



- t1 시간대의 기온과 t2 시간대의 기온은 거의 같지만
- context의 관점에서 볼 때 t2는 anomaly로 해석됨

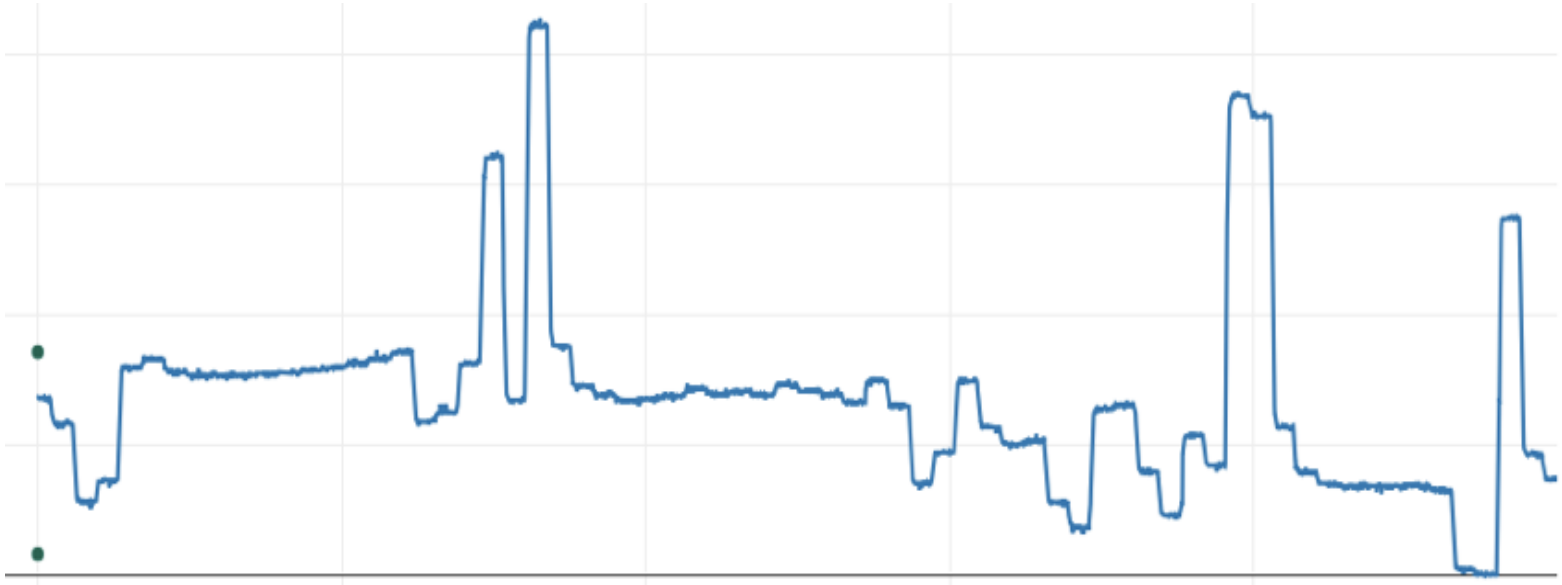
시계열 데이터의 이상 탐지

- 시계열 데이터의 전후 instance, context를 같이 고려해야하는 특성때문에
  - 일반적으로 시계열 데이터의 이상 탐지를 위해서는
  - Autoencoder 등의 이상 탐지 알고리즘과 RNN(Recurrent Neural Networks), LSTM(Long Short Term Memory)등의 시퀀스 데이터 처리가 가능한 알고리즘을 결합하여 사용
  - RNN 등의 네트워크는 입력 값으로 개별적인 instance 대신 sequence를 사용할 수 있어 시간적인 특성을 고려하여 이상치를 탐지할 수 있음

제조 설비에 대한 이상 탐지 예시

제조 설비 센서 데이터

- 센서 데이터 : 온도, 진동, 압력, 서보 모터의 전류 값/각도
- 제어 데이터 : PLC, CNC 데이터



제조 설비의 이상 동작

- 물리적 장치의 미지의 내적 값이 시간이 지나면 증가
- 특정 한계를 넘으면 외적인 현상이 나타난다.
  - 부러짐. 깨짐
- 이로 인해 장애나 불량품 발생

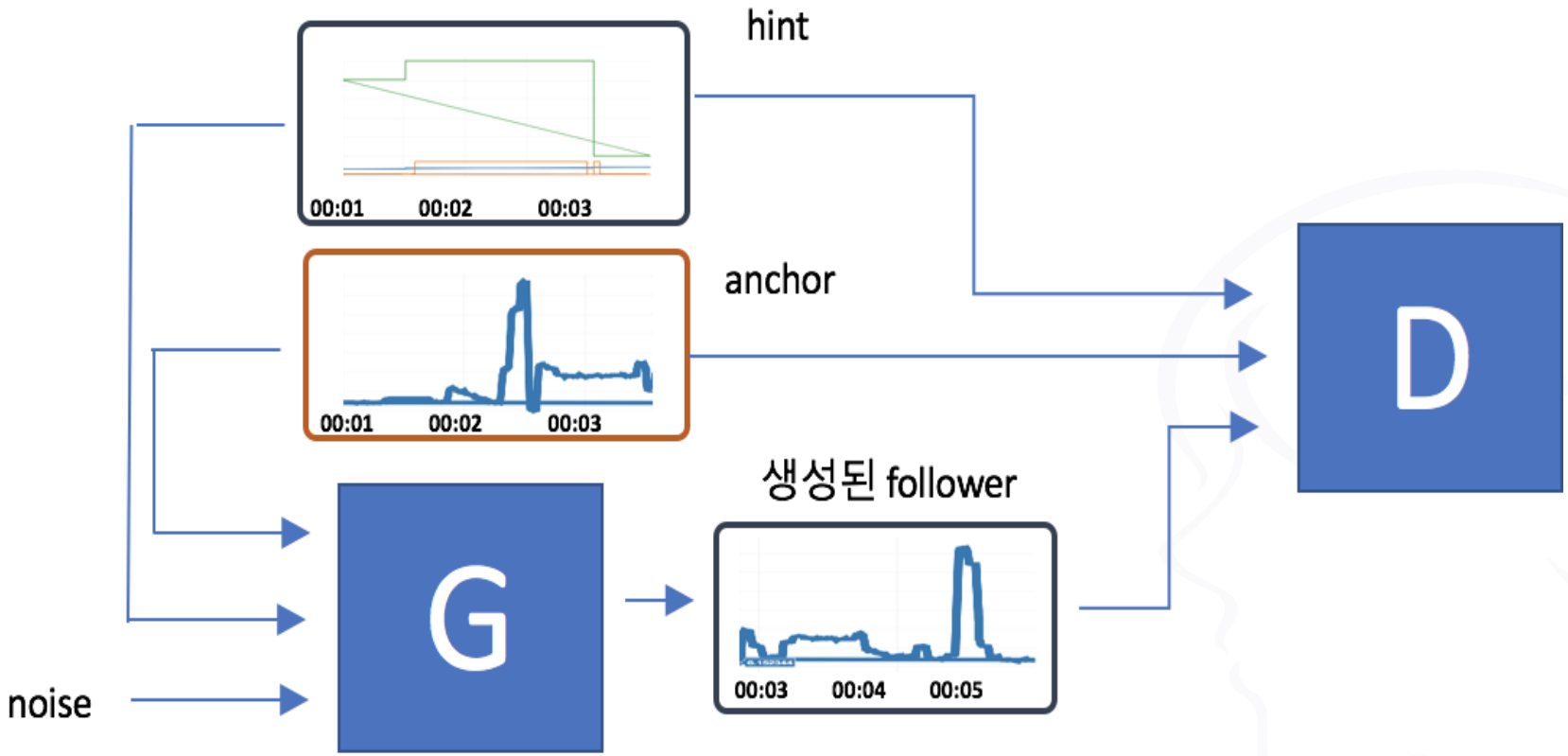
제조 설비의 이상 동작

- 이상 탐지의 어려움
  - 평소와 다른 데이터의 미묘한 변화를 사람이 탐지하기가 무척 어렵다.
  - 기존 통계적인 방법이나 기타 ML 방법으로도 어렵다.

- 제조 설비에 대한 딥러닝 적용의 어려움
  - 실제 제조설비에는 레이블링 데이터(이상 유무)가 거의 없다.
  - 있다 하더라도 정상/비정상의 비율이 극단적이다.

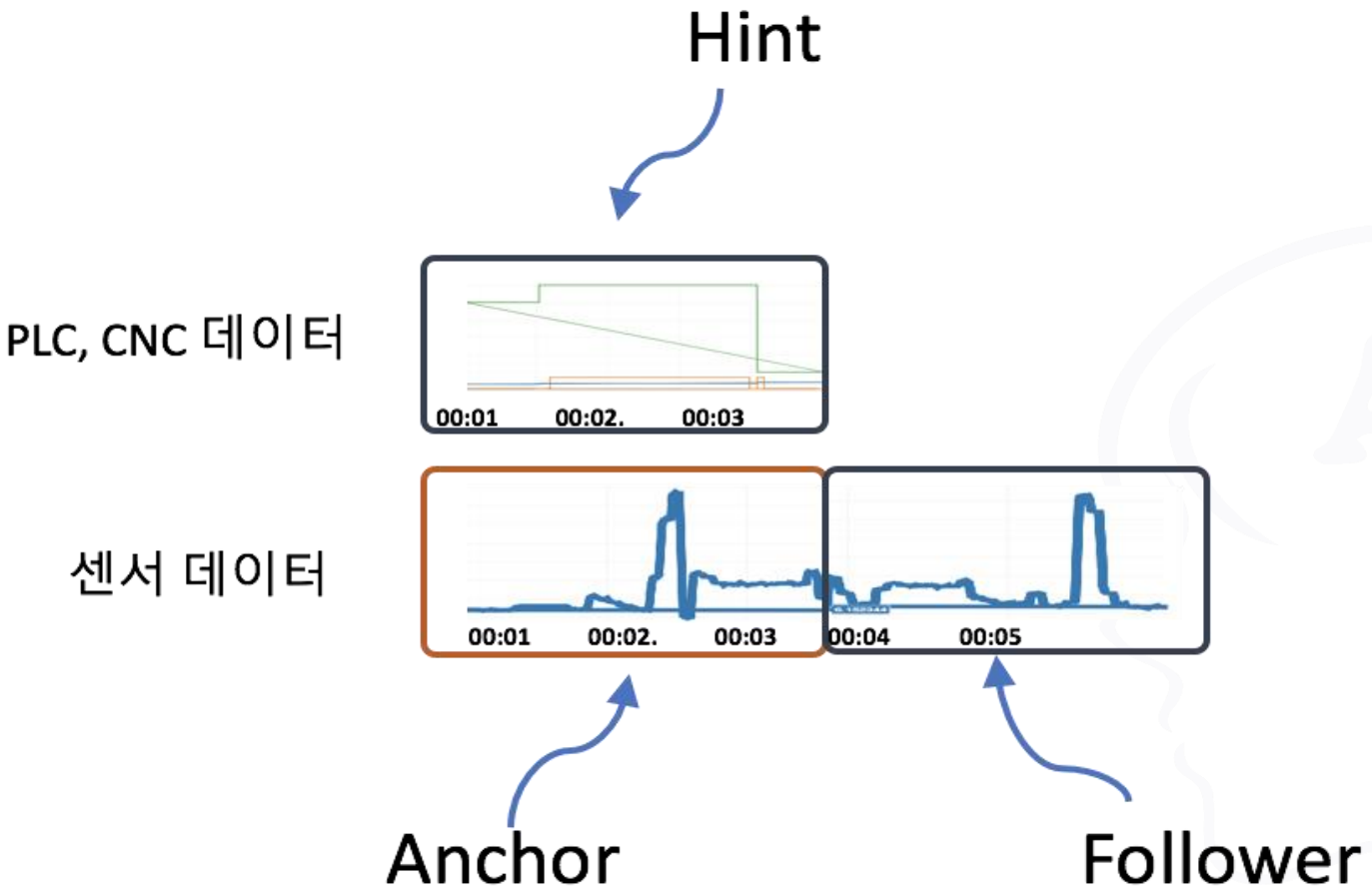
Conditional GAN 모델을 이용한 이상 탐지 예시

- **Conditional GAN 적용**



Conditional GAN 모델을 이용한 이상 탐지 예시

- **데이터의 구성**

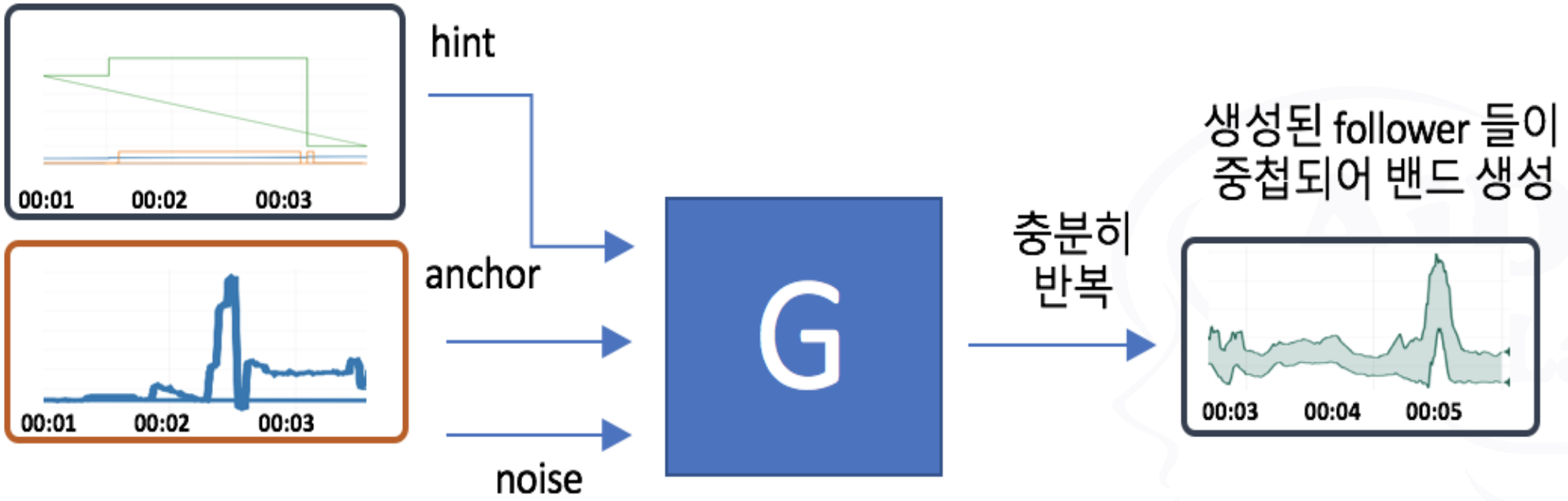


Conditional GAN 모델을 이용한 이상 탐지 예시

- 학습된 생성기 (G, Generator)
  - hint, anchor는 과거의 데이터
  - hint, anchor에 부합하는 미래의 follower 생성
  - 과거의 데이터에 대한 정상이라 할 수 있는 follower
  - GAN의 특징으로 무수한 follower의 생성이 가능하다.

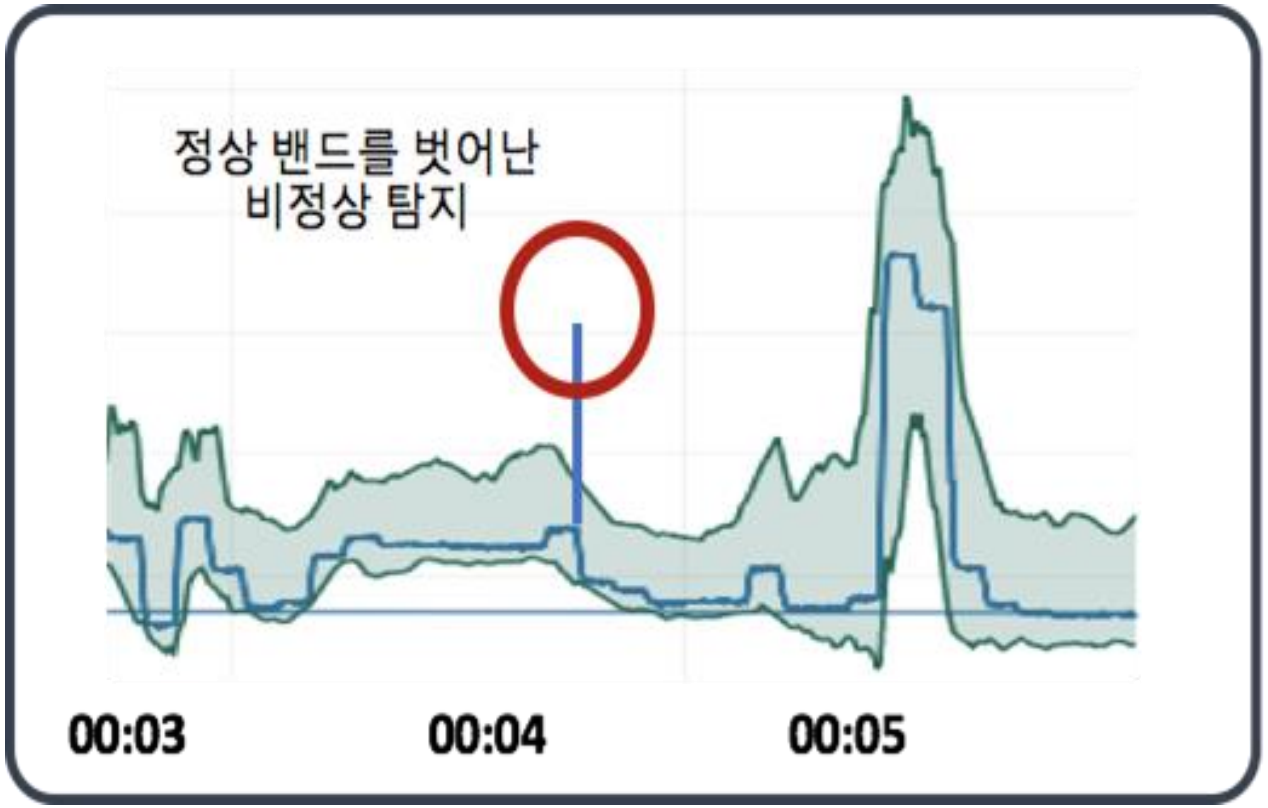
Conditional GAN 모델을 이용한 이상 탐지 예시

- 정상 데이터 밴드 생성



Conditional GAN 모델을 이용한 이상 탐지 예시

- **이상 탐지**



평소 데이터 학습

↓

현재 데이터로
미래 데이터 예측

↓

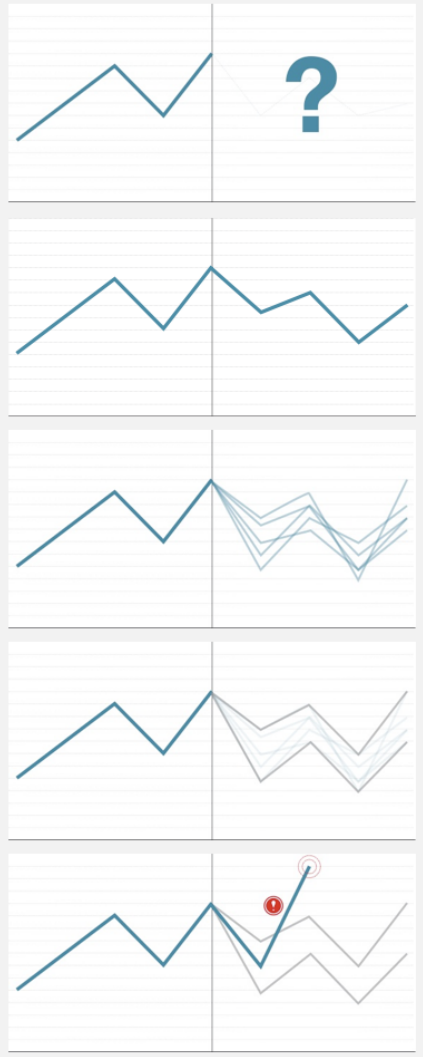
여러 개 예측

↓

중첩하여 밴드 생성

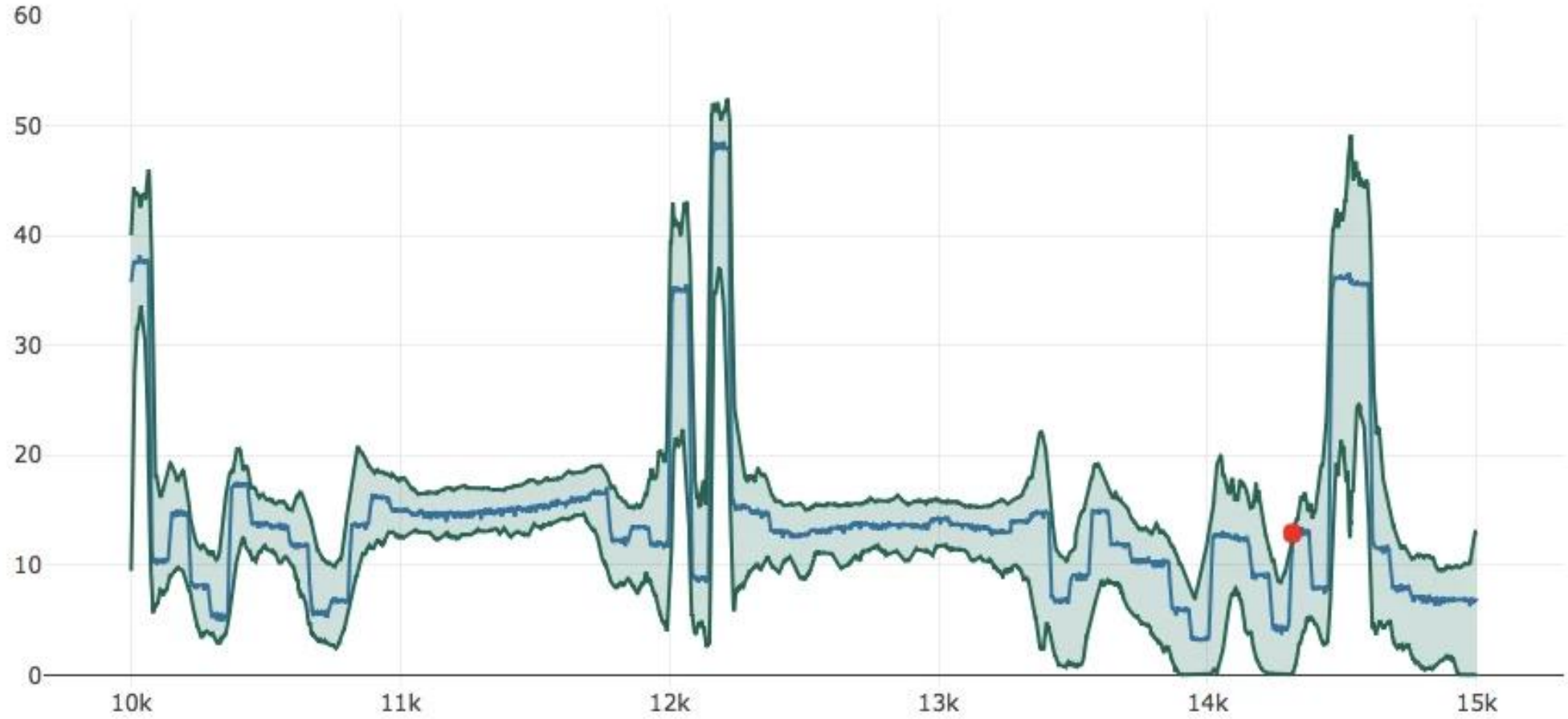
↓

밴드에서 벗어난 것이
비정상



Conditional GAN 모델을 이용한 이상 탐지 예시

- **이상 탐지 결과**



Conditional GAN 모델을 이용한 이상 탐지 예시

- **Conditional GAN 모델을 사용한 이상 탐지 성공의 의미**
  - **비싼 레이블링 데이터가 필요하지 않다.**
  - **타 제어 데이터도 사용**
  - **복수의 센서 데이터가 동시에 사용**