

# 산업용 회전체 예지보전을 위한 VAE 기반 이상 탐지 및 분류 모델

박현석, 김지현, 김수현, 이남경, 김희원, 오동진, 이건희\*  
에이치디씨랩스

{hyunskki, jh991219, soohyun.kim, nk\_lee, ive2go, djoh, Gunhee\_Lee}@hdc-labs.com

## A VAE-based Anomaly Detection and Fault Classification Model for Predictive Maintenance of Industrial Rotating Equipment

Jihyun Kim, Soohyun Kim, Namgyung LEE, Hyunseok Park, Heewon Kim, Dongjin Oh,  
Gunhee Lee\*  
HDC LABS

### 요약

본 논문은 산업 현장의 회전체 예지보전을 목적으로, 진동 센서 데이터를 활용한 이상 탐지 및 이상 유형 세부 분류 모델을 제안한다. 제안된 모델은 변분 오토인코더(VAE)를 기반으로 정상 패턴을 학습하며, 인코더의 잠재 표현을 활용하여 이상 여부를 이진 분류하고, 동시에 이상으로 판단된 데이터에 대해서는 베어링 이상, 정렬 불량 등 세부 고장 유형을 분류할 수 있도록 다중 분류기를 추가한 결합 구조로 설계된다. 다양한 이상 탐지 및 이상 분류 모델과의 성능 비교 결과 제안된 모델의 정상/비정상 탐지 정확도가 0.2%, 세부 분류 정확도는 1.33% 향상되는 결과를 보인다. 또한 실제 현장의 정상 설비 데이터로 실험 결과 가장 낮은 이상 탐지율을 보이며, 이는 제안하는 모델이 실제 설비에 대해서도 더 높은 신뢰도를 보임을 나타낸다.

### I. 서론

산업 설비의 고장이나 이상은 생산성 저하, 유지보수 비용의 증가, 나아가 인명 사고로 이어질 수 있어 이를 사전에 감지하고 대응하는 예지보전 기술의 중요성이 점차 커지고 있다. 특히 회전체나 모터 기반 시스템에서는 진동 센서 데이터를 통해 고장을 사전에 감지하려는 시도가 활발히 이루어지고 있으며[1], 다양한 이상 유형(베어링 이상, 정렬 불량 등)을 구분하는 것도 유지보수 전략 수립에 있어 핵심적인 요소이다.

기존의 이상 탐지 기법은 대부분 정상/비정상을 구분하는 이진 분류에 집중되어 있으며[2], 이상으로 판별된 후 그 원인을 파악하기 위해선 추가적인 전문가의 분석이 필요하거나 별도의 모델이 요구되는 경우가 많다. 이러한 구조는 실시간 진단 및 자동화 측면에서 한계를 가지는 문제가 있다.

본 연구에서는 이런 문제를 해결하고자, 변분 오토인코더(VAE) 기반의 이상 탐지 모델에 잠재 표현을 활용한 세부 고장 유형 분류기를 결합한 결합 구조의 예지보전 모델을 제안한다. 제안된 모델은 VAE를 통해 정상/비정상을 판단하며, 동시에 잠재 표현을 바탕으로 고장 유형을 세분화 분류함으로써 탐지와 세부 분류를 동시에 수행할 수 있는 구조를 갖는다. 학습 시에는 클래스 불균형 문제를 해결하기 위해 SMOTE 기법을 적용하고, 전체 손실 함수에는 재구성 손실, KL 발산, 이진 분류 손실, 세부 분류 손실을 통합하여 잠재표현이 원본을 복원하는 표현이자, 세부분류도 알아 낼 수 있는 표현으로 학습되고 이를 통해서 성능을 극대화한다.

본 연구의 주요 기여는 다음과 같다.

1. 이상 탐지와 고장 원인 추정을 동시에 수행하는 결합 구조의 예지보전 모델을 제안한다.
2. 진동 센서 기반의 데이터셋의 클래스 불균형 보정 및 다중 손실 학습 구조를 통하여 이상 탐지 및 분류 성능을 향상시킨다.
3. 실제 현장의 고장 설비 데이터로 실험을 진행하여 실험실 수준에서의 성능이 아닌 실제 세계 데이터에도 적용 가능함을 보인다.

### II. 본론

본 연구에서는 진동 기반 설비 이상 탐지 및 고장 유형 분류를 위해, 공공 진동 데이터 셋인 CWRU[3]와 [4]에서 제공하는 데이터를 통합하여 학습에 사용한다. 두 데이터 셋은 다양한 고장 유형과 다양한 진동 데이터를 보이며, 이를 통합하여 새로운 통합 진동 진단 데이터셋을 구축하였다. 데이터 간 진폭 범위의 차이를 스케일 정규화를 통해 맞추고 모델의 입력으로 사용된다.

구축된 데이터셋은 학습과 평가 데이터로 분리되며, 이중 학습 데이터셋에 대해서 고장 유형 간 샘플 수의 불균형 문제를 해결하기 위해 SMOTE 기법을 적용한다. 이를 통해 학습 과정에서 저빈도 고장 유형에 대한 분류 성능 저하를 줄인다.

제안하는 모델은 VAE를 기반으로 이상 여부 분류와 이상 세부 유형 분류를 결합 구조로 통합한 모델이다. 전체 구조는 다음과 같은 세가지 주요 컴포넌트로 구성된다:

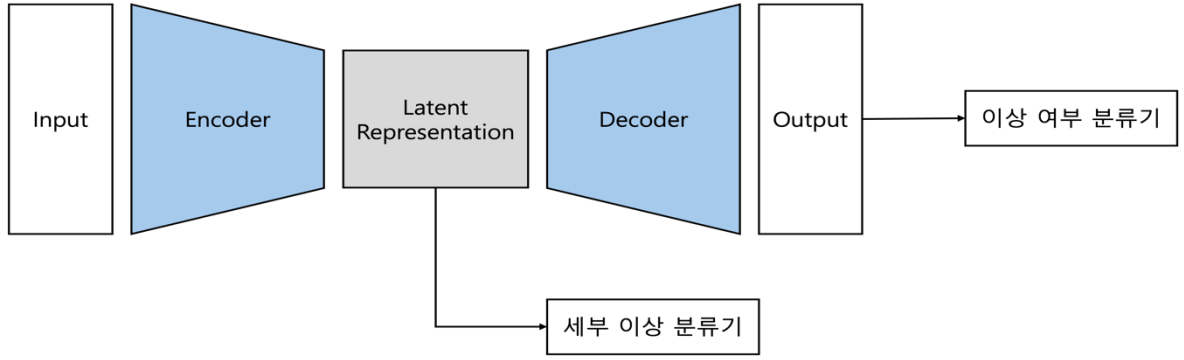


그림 1. 제안 방법의 구조

#### 1. VAE 기반 인코더-디코더

VAE 는 입력 진동 데이터를 인코딩하여 잠재 표현 벡터  $z$ 로 압축하고, 이를 원본으로 복원하는 디코딩 과정이 존재한다. 이 과정은 입력과 복원값의 차이를 최소화하는 재구성 손실과 잠재 분포의 정규화를 위한 KL 발산 손실을 통해 학습된다.

#### 2. 이상 탐지를 위한 이진 분류기

VAE 에서 복원된 값은 정상/비정상 여부를 판단하는 이상 여부 분류기로 입력되고, 이 분류기는 이상 여부를 이진 분류로 판단한다.

#### 3. 고장 유형 세부 분류기

잠재표현  $z$  는 별도의 다중 분류기를 통해 고장 유형을 분류하는 데 사용된다. 이 분류기는 고장 유형에 대한 세부 이상 분류를 진행한다.

그림 1. 은 제안하는 모델의 구조를 나타낸다. 전체 모델은 다음의 결합 손실을 기반으로 학습된다:

$$L_{total} = L_{recon} + \lambda_{kl} * L_{KL} + \lambda_{bin} * L_{binary} + \lambda_{cls} * L_{fault}$$

이 수식에서  $L_{recon}$  은 입력과 복원값의 차이를 계산하는 재구성 손실,  $L_{KL}$  은 잠재 공간 정규화를 위한 KL 발산 손실,  $L_{binary}$  은 정상/비정상 판단을 위한 손실,  $L_{fault}$  은 세부 고장 유형 분류를 위한 손실이고,  $\lambda$  값들은 각 손실 항의 중요도를 조정하는 초매개변수이다.

	이상분류 F1-score	세부분류 F1-score
MLP	0.9609	0.9034
Random Forest	0.9452	0.9064
Gradient Boosting	0.9469	0.9040
Decision Tree	0.9469	0.9067
결합 구조 모델 (Ours)	<b>0.9628</b>	<b>0.9154</b>

표 1. 모델별 지표

제안하는 방법의 성능을 평가하기 위해 해당 통합 데이터셋에 대해 정상/비정상 분류(Binary), 세부분류(Multi-class)에 대한 F1-score 를 평가한다. 표 1.은 모든 모델에 비해, 정상/비정상 탐지 성능과 세부분류 성능 모두 다 좋게 나타났으며, 이 결과는 결합 구조 방식의 모델이 이상을 탐지하고 더 나아가 무슨 이상인지 분류하는데 더 적합하다는 것을 알 수 있다.

또한 HDC 아이파크몰의 실제 현장의 비정상 설비 데이터 2 주치로 정상/비정상 이진 분류 테스트를 해본 결과, Random Forest 는 61.8%, MLP 는 72.5%의 결과를 보인 반면, 제안하는 방법은 84.7%로 제안하는 방법이 가장

높은 이상 탐지율을 보이며 이는 실제 현장의 데이터에도 다른 모델 대비 가장 좋은 이상탐지 성능을 보임을 나타낸다.

### III. 결론

본 연구에서는 진동 센서 기반의 회전체 설비에 대해 이상 탐지와 세부 고장 유형 분류를 동시에 수행할 수 있는 결합 구조의 예지보전 모델을 제안한다. 제안된 모델은 VAE 를 기반으로 정상 패턴을 학습하고, 잠재 표현을 활용하여 정상/비정상 이진 분류 및 고장 유형 다중 분류를 동시에 수행하는 구조로 설계되었다. 실험은 CWRU 및 공개 진동 진단 데이터셋을 통합한 실험 데이터를 사용했으며, 기존의 모델들과 비교하여 정상/비정상 탐지 정확도와 세부 분류 정확도가 향상된 결과를 보였다. 또한 실제 현장 설비 데이터 실험 결과에서도 더 좋은 성능을 보이며 이는 실제 산업 설비 환경에서의 예지보전 시스템 구현에 매우 유의미한 기반 기술로 활용될 수 있다.

마지막으로, 제안하는 모델은 단순 이상 탐지를 넘어서 실제 고장 원인까지 식별할 수 있는 하나의 구조를 갖추고 있어, 향후 자동화 및 MLOps 시스템 구현에 접근성이 높게 활용될 수 있다.

### 참 고 문 헌

- [1] 이수형, 고상근 and 이수안. (2022). 기계장비 진동 데이터를 이용한 딥러닝 기반의 고장 분류 모델. 한국차세대컴퓨팅학회 논문지, 18(2), 36-46.
- [2] 서운호, 마평식, 우정환, 선경환, 김병욱, 김봉석. (2021). Variational Autoencoder 를 이용한 데이터 및 지식기반 회전기기 이상 진단. 한국소음진동공학회논문집, 31(3), 289-297. 10.5050/KSNVE.2021.31.3.289.
- [3] D. Neupane and J. Seok, "Bearing Fault Detection and Diagnosis Using Case Western Reserve University Dataset With Deep Learning Approaches: A Review," in *IEEE Access*, vol. 8, pp. 93155-93178, 2020,
- [4] Atmaja, Bagus Tris, et al. "Lab-scale vibration analysis dataset and baseline methods for machinery fault diagnosis with machine learning." *Journal of Vibration Engineering & Technologies* 12.2 (2024): 1991-2001.