

LSTM-VAE를 활용한 기계시설물 장치의 이상 탐지 시스템

서재홍* · 박준성** · 유준우** · 박희준***†

* 연세대학교 산업공학과 석사과정

** 연세대학교 산업공학과 박사과정

*** 연세대학교 산업공학과 정교수

Anomaly Detection System in Mechanical Facility Equipment: Using Long Short-Term Memory Variational Autoencoder

Seo, Jaehong* · Park, Junsung** · Yoo, Joonwoo** · Park, Heejun***†

* Industrial Engineering, Yonsei University, M.S. candidate

** Industrial Engineering, Yonsei University, Ph.D. candidate

*** Industrial Engineering, Yonsei University, Professor

ABSTRACT

Purpose: The purpose of this study is to compare machine learning models for anomaly detection of mechanical facility equipment and suggest an anomaly detection system for mechanical facility equipment in subway stations. It helps to predict failures and plan the maintenance of facility. Ultimately it aims to improve the quality of facility equipment.

Methods: The data collected from Daejeon Metropolitan Rapid Transit Corporation was used in this experiment. The experiment was performed using Python, Scikit-learn, tensorflow 2.0 for preprocessing and machine learning. Also it was conducted in two failure states of the equipment. We compared and analyzed five unsupervised machine learning models focused on model Long Short-Term Memory Variational Autoencoder(LSTM-VAE).

Results: In both experiments, change in vibration and current data was observed when there is a defect. When the rotating body failure was happened, the magnitude of vibration has increased but current has decreased. In situation of axis alignment failure, both of vibration and current have increased. In addition, model LSTM-VAE showed superior accuracy than the other four base-line models.

Conclusion: According to the results, model LSTM-VAE showed outstanding performance with more than 97% of accuracy in the experiments. Thus, the quality of mechanical facility equipment will be improved if the proposed anomaly detection system is established with this model used.

Key Words: Multivariate Time Series Data, Unsupervised Learning, Machine Learning, Anomaly Detection, Long Short-Term Memory Variational Autoencoder

● Received 25 November 2021, 1st revised 28 November 2021, accepted 1 December 2021

† Corresponding Author(h.park@yonsei.ac.kr)

© 2021, Korean Society for Quality Management

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-Commercial use, distribution, and reproduction in any medium, provided the original work is properly cited

1. Introduction

이상 탐지(Anomaly Detection)란, 정상(normal)과 구분되는 이상(anomaly)을 탐지하는 것으로 Novelty Detection, Outlier Detection이라고 불리기도 한다(E.S. Lee et al., 2020). 이상 탐지는 크게 비지도 이상 탐지(Unsupervised Anomaly Detection)와 지도 이상 탐지(Supervised Anomaly Detection)로 나뉜다. 대부분 이상 데이터는 정상 데이터에 비해 그 수가 적게 발견되기 때문에 학습 과정 시 정상 데이터만을 활용하는 비지도 이상 탐지 방법들이 지도 이상 탐지 방법에 비해 보다 집중 연구되고 있다(Perera P et al., 2019; S. Akcay S et al., 2019).

주로 활용되는 분야로는 침입탐지시스템(IDS), 의료건강(Healthcare), IoT 이상탐지(Hasan M et al., 2019), 시계열 데이터분석, 영상분석, 이상금융거래탐지시스템(FDS), 산업 분야 이상탐지 등이 있다(Chalapathy R and Chawla S, 2019).

기계시설물의 데이터란 기계의 정상 상태와 고장 상태의 데이터를 말하며, 이 데이터는 이상 탐지를 통해 기계의 이상유무 파악에 활용 될 수 있다. 최근 인공지능의 급격한 발달로 기존의 전문가의 경험에 의존하던 전통적인 이상 탐지 방법에서 탈피하여 다양한 머신러닝 기법들이 등장하기 시작했다(K.H. Sun et al., 2020). 특히, 오토인코더(Autoencoder; AE)(Zong B et al., 2018; Baur C et al., 2018)와 적대적 생성 네트워크(Generative Adversarial Network; GAN)(Li D et al., 2018; N.Y. Choi and W.H. Kim, 2019) 방식을 기반으로 한 딥러닝 기법들이 가장 활발히 연구되고 있다.

본 연구에서 사용하는 데이터는 대전광역시 도시철도공사의 각 역사에 설치된 기계장치의 진동 및 전류센서로부터 수집되었다. 주로 송풍 및 환기를 담당하는 기계장치의 모터에 대한 데이터이다. 기계시설물의 모터가 고장 나면 산업, 교통, 급배수, 소방 방재, 전력 계통 등 전체 시스템에도 악영향을 줄 수 있다. 특히 이러한 산업 현장에서 갑작스러운 셧다운(shutdown) 현상이라도 발생하게 되면 손해비용이 하루에 5,000만원가까이 발생할 만큼 치명적이라고 알려져 있다(DataRPM, 2017). 따라서 미리 고장을 예측하고 사전에 예방할 수 있는 기술이 필요하다. 현재 유럽의 철도차량 기술 사양인 TSI(Technical Specification for Interoperability)은 철도차량의 실시간 모니터링을 통해 탈선감지 및 오작동 방지를 위해 노력하고 있으며 핀란드에서는 RFID(Radio-Frequency Identification) 솔루션이 적용된 철도 인식 파일럿 시스템을 개발하여 2012년부터 철도차량의 모니터링을 실시하고 있다. 국내에서는 서울 지하철 2호선에 진동검지방식으로 이상검사장치를 운용 중에 있으나, 다양한 환경에서 적용 가능한 이상탐지 시스템 개발을 통해 전국으로 상용화하기 위해서는 보다 활발한 연구가 필요하다.

본 연구의 목적은 기계시설물 장치 데이터를 사용하여 이상 탐지에 대표적으로 사용되는 알고리즘들의 성능을 비교하고, 산업 시계열 데이터에서 시간 정보 반영의 중요성에 대해 탐구하는데 있다. 또한 LSTM-VAE(Long Short-Term Memory Variational Autoencoder)를 활용한 기계시설물 장치의 이상 탐지 시스템을 제안하여 본 연구의 결과를 통해 사물인터넷(IoT)을 적용한 지하철 역사 내의 설비고장 진단 인공지능 시스템을 구축하는 것이다. 이를 활용해 고장을 예측하고 유지보수 계획을 수립하는데 도움을 주어 궁극적으로 기계시설물 장치의 품질향상을 이끌어낼 수 있을 것으로 기대된다.

2. Related works

2.1. Anomaly detection

이상탐지의 정의는 다양하다. 진정 확률 밀도(True probability density)가 낮은 데이터를 이상으로 보기도 하고 (Harmeling S et al., 2006), 예상한 패턴을 따르지 않는 데이터를 이상으로 보기도 한다(Chandola V et al., 2009). 이처럼 예상한 패턴이나 행동에 맞지 않는 특이한 데이터를 찾아내는 방법들을 말하며, 다른 말로 Novelty Detection, Outlier Detection이라고 불리기도 한다.

이상탐지의 모델들은 과거부터 현재에 이르기까지 다양하게 연구되고 있다. 과거에는 통계적 분석과 회귀분석법이 주로 사용되었으며 대표적으로 확률적 모델과 선형 모델들이 있다. 현재는 기계학습(machine learning) 방법과 인공신경망을 활용한 연구들이 가장 활발하게 진행되고 있다. 크게 데이터의 라벨의 유무에 따라 지도(Supervised), 반지도(Semi-supervised), 비지도(Unsupervised)로 나눌 수 있다. 지도 학습 기반의 이상탐지는 학습 데이터의 정상과 비정상 라벨이 모두 존재하는 경우를 말한다. 모든 데이터의 라벨링이 되어 있기 때문에 다른 학습 방법들에 비해 높은 정확도를 보이는 장점이 있다. 그러나 실제 산업 현장에서는 비정상 데이터가 정상 데이터에 비해 수집되기 어려운 점 때문에 사용하는데 한계가 있다. 반지도 학습 기반의 이상탐지는 정상 데이터만을 가지고 학습하는 방법으로 서포트 벡터머신(Support Vector Machine: SVM)과 AE(Autoencoder)(Zong B et al., 2018) 그리고 GAN(Generative Adversarial Net-work)(Li D et al., 2018) 등을 활용해 활발히 연구가 이루어지고 있다. 정상 데이터의 라벨만을 가지고도 학습을 진행할 수 있는 장점이 있지만, 데이터 패턴의 변화나 설정 변수 등에 민감하고 지도학습에 비해 정확도가 떨어진다는 단점이 있다. 마지막으로 비지도 이상탐지는 데이터의 라벨이 없는 경우에 학습을 진행하는 방법으로 주로 주성분 분석(Principal Component Analysis: PCA)과 AE(Autoencoder)(Baur C et al., 2018; M.K. Seo and W.Y. Yun, 2019)를 사용한 연구가 있다. 라벨이 필요 없다는 장점이 존재하나, 데이터의 분포가 비선형적일 경우 정확도가 떨어지며 하이퍼 파라미터(Hyper Parameter)에 민감하다는 단점이 있다.

이상탐지는 다양한 분야에서 활용되고 있다. 침입탐지시스템(IDS), 의료건강(Healthcare)(Wei Q et al., 2018), IoT(Hasan M et al., 2019), 시계열 데이터분석, 영상분석, 이상금융거래탐지시스템(FDS), 자율 자동차 주행, 컴퓨터 네트워크 해킹, 산업 분야 이상탐지 등이 있다(Hodge V and Austin J, 2004; M. J. Kim et al., 2019; J.S. Kim et al., 2021; Chalapathy R and Chawla S, 2019).

최근 딥러닝을 사용해 많은 이상탐지 연구가 진행되고 있으며, 센서데이터에 대한 연구로는 LSTM-AD(Malhotra P et al., 2015)와 GAN(N.Y. Choi and W.H. Kim, 2019)알고리즘을 활용한 연구 등이 있다. 본 연구의 기계시설물 장치의 이상탐지는 시설물에 부착된 센서로부터 측정된 진동 및 전류 데이터로 이루어져 있으며, 산업현장 특성상 비정상 데이터 수집에 어려운 점을 반영해 비지도 학습기반 이상탐지 방법을 사용하였다.

2.2. Anomaly score

비정상 데이터의 라벨이 없는 경우에 이상탐지를 위해서는 정상과 비정상을 구분하는 기준이 필요하다. 이상탐지에서는 이 기준선이 되는 임계값(threshold)을 정해 이보다 높거나 낮음을 판단해 이상치를 판별한다. 따라서 데이터에 이상 수준을 의미하는 이상 점수(Anomaly score)를 부여하여 정상과 비정상 데이터를 구분한다(Chalapathy R and Chawla S, 2019).

$$Anomaly\ Score = -\log p(x_t; \mu_{x_t}, \Sigma_{x_t}) \quad (1)$$

μ_{x_t}, Σ_{x_t} : mean and covariance of the reconstructed distribution, $N \sim (\mu_{x_t}, \Sigma_{x_t})$

$$\begin{cases} \text{if } f_s(x_t, \emptyset, \theta) > \text{threshold,} & \text{anomaly} \\ \text{otherwise,} & \text{normal} \end{cases} \quad (2)$$

본 연구에서 사용하는 모델은 LSTM-VAE(Long Short-Term Memory Variational Autoencoder)로 Formula (1)과 같이 $N \sim (\mu_{x_t}, \Sigma_{x_t})$ based 관측치의 negative log likelihood를 사용해서 이상 점수(Anomaly score)를 정의할 수 있다. μ_{x_t} 와 Σ_{x_t} 는 각각 모델을 통해 계산된 재구성 분포의 평균과 공분산을 나타내며 Formula (2)에서 $f_s(x_t, \emptyset, \theta)$ 는 이상 점수 추정 값을 의미한다. 그리고 이상 진단 모델은 아래 Formula (4)에서 정의된 손실 함수(Loss function)를 이상 점수로 사용할 수 있다. 손실 함수를 작게 만들면 정상 데이터와의 오차를 줄이면서 데이터의 확률 분포를 유사하게 형성하는 것이기 때문에 손실 함수의 값이 임계값(threshold) 이상이 되면 이상이 있다고 판단할 수 있다. 결국 이상 점수(Anomaly score)가 높으면 LSTM-VAE(Long Short-Term Memory Variational Autoencoder) 모델로부터 재구성(reconstruction)이 잘 되지 않았다는 것을 의미하게 되는 것이다(Park D et al., 2018). 여기서 임계값은 정상 데이터로 학습했음을 고려했을 때 누적 확률 분포(Cumulative Probability Density)가 99% 이상인 지점으로 결정할 수 있다.

2.3. Autoencoder(AE)

이상탐지에서 Autoencoder는 인공지능망을 활용한 비지도 학습 기반의 방법 중 하나이다(Bank D and et al., 2020). Figure 1은 Autoencoder의 구조이다. 인코더와 디코더로 구성되어 있으며, 입력과 출력 시의 노드 개수가 같다. Input layer에서 데이터를 받아 Hidden layer로 갈수록 그 수가 줄어들면서 정보를 압축하고 다시 Output layer로 복원하는 과정을 거쳐 재구성 오류를 생성하는데, 이때 발생하는 이 재구성 오류의 정도를 이상 점수(Anomaly score)로 사용하여 임계값(threshold)과 비교하여 데이터의 정상유무를 판단한다(An J and Cho S, 2015).

오토인코더는 데이터의 라벨이 필요하지 않으며, 입력 데이터와 출력 데이터가 유사한 방향으로 학습이 진행된다. 이때 입력 데이터의 특징을 추출하여 Hidden layer를 통하여 잠재공간(Latent space)으로 저장한 후 Autoencoder의 나머지 한 부분인 디코더에서 원래 데이터를 복원하는 과정을 거친다. 보통 잠재공간으로 가는 과정동안 데이터의 손실이 발생하게 되는데 이 때문에 복원된 데이터는 원본 데이터와 차이를 보이게 된다. 학습이 끝난 후 새로 입력된 데이터는 원본 데이터에 비해 주요하지 않은 특징들은 사라지고 기존 학습 데이터와 유사한 방향으로 복원한다.

본 연구에서 사용하는 기계시설물 데이터는 센서데이터라는 측면에서 시스템 모니터링 분야와 유사한 특징을 가진다. 현재 해당 분야에서 Autoencoder를 활용한 연구들이 활발히 진행되고 있다. 또한 센서 데이터는 다차원의 시계열 데이터라는 주요한 특징이 있는데, 기존 연구는 이러한 특징을 활용하기 위해 순환신경망(RNN)이라고 할 수 있는 LSTM을 Autoencoder에 적용하기도 했다. 또한 VAE(Variational Autoencoder)(Guo Y et al., 2018)와 GAN(Generative Adversary Network)(Son V H et al., 2019)를 활용한 연구들도 활발히 이루어지고 있다.

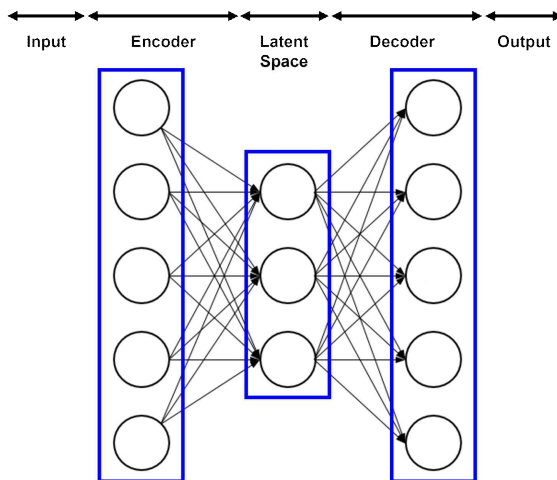


Figure 1. Architecture of Autoencoder

3. Experiment

3.1. Anomaly Detection System of Mechanical Facility Equipment

본 논문에서 제안하는 이상탐지 시스템 방법은 Figure 2와 같다. ① 각 시설물에 설치된 센서(진동, 전류, 온도, 습도, 미세먼지센서 등)로부터 무선 네트워크를 통해 원격 데이터를 수집한다. 통신서버, 수집서버, 분류 서버로 구성된 서버로부터 각 데이터들을 목적에 맞게 전송한 후 데이터베이스에 저장한다. ② 수집된 데이터는 이상값, 제한값, 결측값 등을 삭제하고 정상과 비정상 데이터로 구분하여 1차적으로 데이터 정제의 과정을 거친다. ③ 정제된 데이터는 데이터 탐색, 파라미터 검사, 학습 및 검증 데이터셋의 생성을 통해 데이터 검증을 한다. ④ 데이터분석 과정에서는 AI학습을 위해 필요한 전처리 과정, 학습 그리고 모델링을 통해 기계시설물의 고장을 예측하고 원인을 분석한다. ⑤ 데이터 분석의 결과로부터 예측된 결과를 활용하여 원격모니터링과 운용계획 및 유지보수 계획을 수립한다.

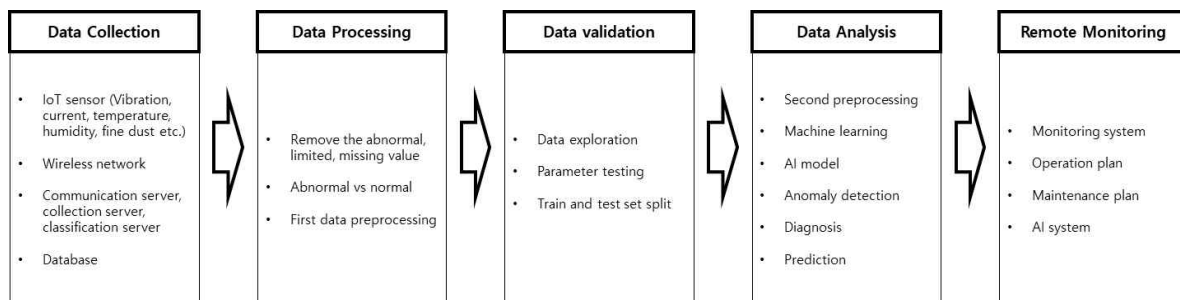


Figure 2. Anomaly detection system of mechanical facility equipment

3.2. Data

본 연구에서 사용하는 데이터는 기계시설물 데이터이다. 기계시설물 데이터란 기계의 정상 상태와 고장 상태의 데이터를 의미하며, 이러한 데이터는 기계학습을 통해 현재 기계의 상태를 예측하는데 활용가능하다. 사용된 데이터는 대전광역시 도시철도공사에서 구축한 기계시설물 고장 예지 센서 데이터로 대전광역시도시철도공사, 한국기계연구원, (주)에스씨솔루션에서 가공에 참여하였고 데이터메이커에서 검수를 마쳤다. 주로 송풍 및 환기를 담당하는 기계장치의 모터에 대한 데이터로 대전역, 시청역, 갑천역 역사의 공조실에 설치된 41개의 모터에 대해서 출력별로 분류하여 수집되었다. 진동센서와 3상 전류센서를 통해 측정된 정상상태, 베어링 불량, 회전체불평형, 축정렬 불량 그리고 벨트 느슨함 총 5가지 유형의 부하별 데이터 구축량은 Table 1과 같다.

Table 1. Total volume of data collected

Load (KW)	Volume of data collected (Current/Vibration)								
	Failure data					Normal Data			
	Unbalanced rotating body	Loose belt	Axis alignment failure	Bearing failure	Subtotal		Subtotal		
					Quantity	Ratio(%)	Quantity	Ratio(%)	
2.2	30,000	60,000	30,000	40,000	160,000	15.5	270,000	25	
3.7	30,000	-	30,000	40,000	100,000	9.7	100,000	9.3	
3.75	-	-	30,000	-	30,000	2.9	30,000	2.8	
5.5	30,000	30,000	30,000	40,000	130,000	12.6	120,000	11.1	
7.5	-	30,000	30,000	40,000	100,000	9.7	80,000	7.4	
11	30,000	30,000	30,000	40,000	130,000	12.6	150,000	13.9	
15	30,000	30,000	-	40,000	100,000	9.7	80,000	7.4	
18.5	-	30,000	-	40,000	70,000	6.8	60,000	5.6	
22	30,000	30,000	30,000	-	90,000	8.7	70,000	6.5	
30	-	-	30,000	-	30,000	2.9	30,000	2.8	
37	-	-	30,000	-	30,000	2.9	30,000	2.8	
55	30,000	30,000	-	-	60,000	5.8	60,000	5.6	
Sub total	Total	210,000	270,000	270,000	280,000	1,030,000		1,080,000	
	Ratio(%)	20.4	26.2	26.2	27.2	100			

본 연구에서는 2.2kw 모터의 회전체불평형과 축정렬불량 두 가지 상황을 가정하여 실험을 구성하였다. 먼저 학습에 사용될 정상 데이터는 10,000초 동안 수집된 전류값(R상, S상, T상)과 진동 가속도값의 RMS(root mean square) 데이터를 1초 단위로 계산하여 사용하였다. 학습된 모델의 검증에 사용할 검증 데이터는 회전체불평형과 축정렬불량 상황 각각에서 2,000초 동안 측정된 데이터와 2,000초 동안의 정상 상태에서 측정된 데이터를 결합하여 4,000초 동안 정상 상태에서 불량상태로 변하는 두 가지 상황을 가정하였다.

3.3. Preprocessing

기계시설물 데이터를 머신러닝 모델에 효과적으로 학습시키기 위해서는 전처리가 필수적이다. 데이터는 3개의 전류 값과 1개의 진동 값으로 4가지 속성으로 구성된다. 1초마다 측정된 RMS(root mean square)값을 사용하였으며 정상 데이터는 0, 이상 데이터는 1로 라벨링 하였고, 데이터의 각 속성의 값을 일정한 범위로 만들어주기 위해서 최소값과 최대값이 0에서 1사이의 범위를 따르도록 표준화를 진행해주었다. 본 실험의 평가를 위해 회전체불평형과 축정렬불량 두 가지 고장 상황을 가정하여 실험을 진행했다. Figure 3은 정상 상태의 데이터를 도식화 한 것이고, Figure 4는 회전체불평형 고장이 발생했을 때 데이터를 그래프로 나타낸 것이다. 마지막으로 Figure 5는 축정렬불량 고장 시 상황을 가정하여 만든 테스트 데이터를 나타낸 것이다.

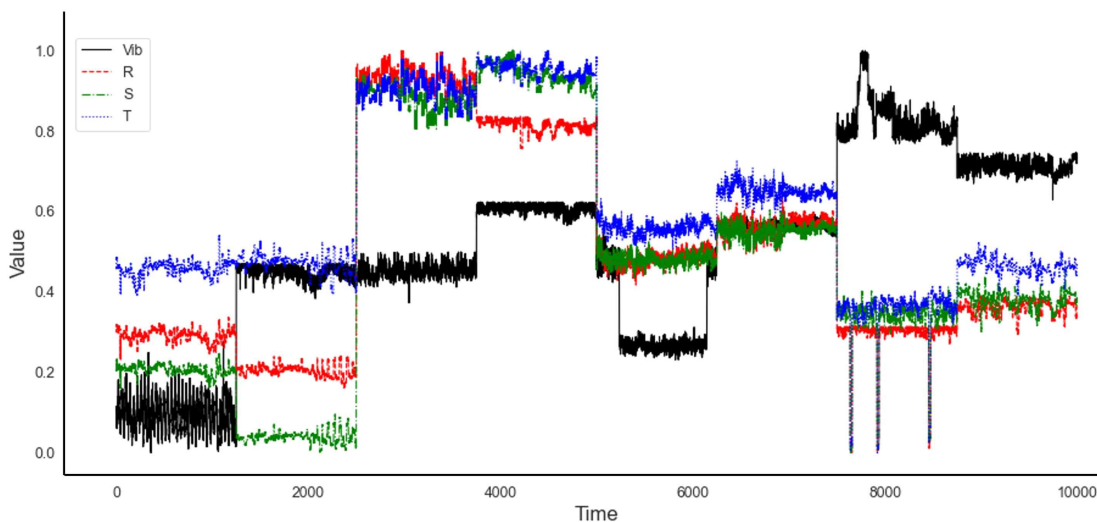


Figure 3. Vibration and current value of normal data

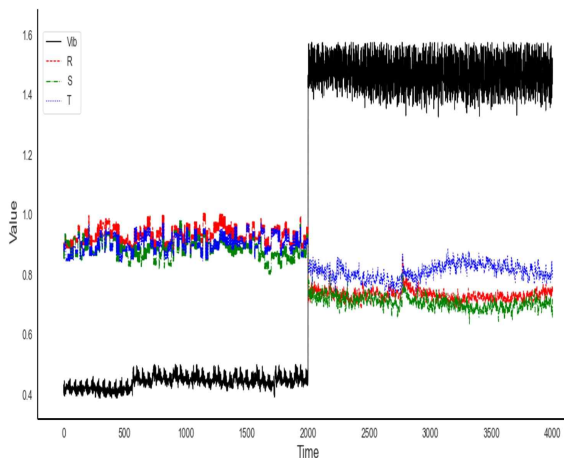


Figure 4. Data of rotating body failure

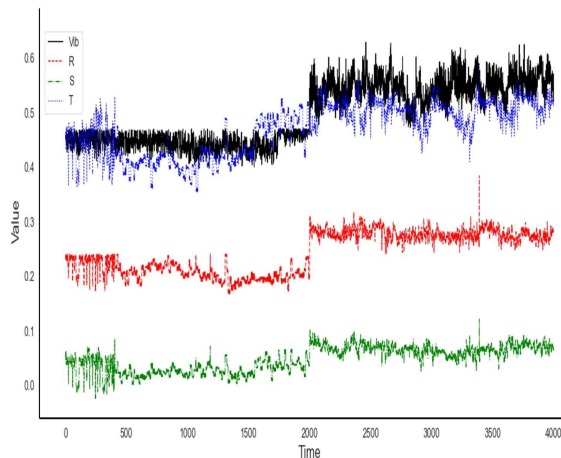


Figure 5. Data of axis alignment failure

먼저 기계학습을 위해 필요한 학습 데이터 세트는 정상 상태에서 1초에 한 번 측정된 진동과 전류 데이터 10,000개를 추출하여 만들었다. 학습된 모델을 평가하기 위한 첫 번째 실험은 1초에 한 번 측정된 정상상태의 데이터 2,000개와 측정렬불량 상태에서의 데이터 2,000개를 합쳐서 정상에서 불량상태로 전환되는 4,000초 동안의 데이터로 구성했다. 두 번째 실험은 1초에 한 번 측정된 정상상태의 데이터 2,000개와 회전체불평형 상태에서의 데이터 2,000개를 합쳐서 정상에서 불량상태로 전환되는 4,000초의 데이터로 구성했다.

3.4. Model

이상탐지를 위한 방법으로 통계 기법, 머신러닝, 딥러닝 등 여러 가지 기법들이 있는데, 최근에는 머신러닝, 딥러닝과 같은 인공지능 알고리즘을 사용한 이상 탐지 기법이 활발하게 연구되고 있다. 실험을 위해 총 5가지의 이상탐지 모델을 도입하였다. 먼저 Random은 무작위적 분류모델로 다른 알고리즘 모델들의 성능 평가 기준이 되어준다. 둘째로, IF(Isolation Forest)와 OSVM(One-Class Support Vector Machine)모델은 이상탐지 성능을 비교하기 위해 선정하였다. 두 모델의 경우 널리 사용되는 이상탐지 모델이나 시간 정보를 포함하지 않는다. 마지막으로 시간 정보를 포함한 알고리즘으로 AE(Autoencoder)를 선정하여 LSTM-VAE(Long Short-Term Memory Variational Autoencoder)와의 성능 비교를 통해 알고리즘의 정확도 향상 정도를 확인하였다.

선정된 모델들의 파라미터는 Table 2와 같다. 동일한 조건에서 LSTM-VAE와 AE의 성능을 비교하기 위하여 Latent space의 크기는 4로 고정했다.

Table 2. Model Parameter

Model	Parameter Setup
IF	n_estimators = 1,000, contamination = auto, max_features = 1, max_samples = auto
OSVM	kernel = rbf, gamma = 0.1, nu = 0.01
AE	Hidden layer = (32, 16, 4)
LSTM-VAE	Hidden layer = (32, 16), Latent space = 4

해당 모델들을 제외하고도 이상 탐지 알고리즘에 모델에 대한 다양한 연구가 존재한다. 특히, 앞서 2장에서 언급했던 것처럼 AE(Autoencoder)와 GAN(Generative Adversarial Network)기반의 모델들이 가장 활발하게 연구되고 있다. 그러나 본 연구에서 사용하는 기계시설물 데이터는 하루에도 수십만 개 이상 수집되는 시간정보가 담긴 빅데이터라는 특성이 있다. 이 같은 특성을 가지는 산업 시계열 데이터는 복잡한 계산 구조를 가진 알고리즘을 사용할 경우 기하급수적으로 학습 시간이 늘어나게 된다. 따라서 오랜 학습 시간을 요구하는 GAN기반 모델이나 여러 가지를 융합한 하이브리드 모델들을 사용하기보다 데이터의 특성을 잘 반영할 수 있는 AE기반의 LSTM-VAE를 사용하여 실험을 진행했다.

본 연구에서 사용하는 주요 이상탐지 알고리즘은 LSTM-VAE이다. VAE(Variation Autoencoder)는 Formula (3)의 손실 함수(Loss function)을 최소화하여 모델을 얻을 수 있다.

$$L(\varnothing, \theta, x_i) = -E_{q_{\varnothing}}[\log p_{\theta}(x_i|z)] + D_{KL}[q_{\varnothing}(z|x_i)|p_{\varnothing}(z|x_i)] \quad (3)$$

VAE의 디코더는 데이터의 사후확률 $p(z|x)$ 를 학습한다. 그러나 사후확률의 계산이 어렵기 때문에 이를 보다 다루기 쉬운 분포 $q(z)$ 로 근사하는 변분추론의 방법을 사용한다. 변분추론은 $p(z|x)$ 와 $q(z)$ 사이의 KLD(Kullback-Leibler divergence)를 계산하고, KLD가 줄어드는 방향으로 $q(z)$ 를 조정해서 최종 $q(z)$ 를 얻는 것이다. Formula (3)의 우변 $-E_{q_\phi}[\log p_\theta(x_i|z)]$ 는 Reconstruction Term으로 인코더가 데이터 x 를 받아서 q 로부터 z 를 뽑아내고, 디코더는 인코더가 만든 z 를 받아서 원 데이터 x 를 복원할 때 이 둘 사이의 크로스 엔트로피를 나타낸다. 즉, 입력(Input)과 출력(Output)의 오차(Error)의 기댓값을 나타내기 때문에 이상적인 샘플링 함수로부터 얼마나 잘 복원했는가를 의미한다. 또한 좌변 $D_{KL}[q_\phi(z|x_i)|p_\theta(z|x_i)]$ 는 Reconstruction Term으로 z 가 zero-mean Gaussian이라고 가정한 인코더의 실제 사후 확률 분포 ($p_\theta(z|x_i)$)와 근사화 확률 분포($q_\phi(z|x_i)$)의 유사한 정도를 나타내는 KLD이다. 따라서 이상적인 sampling 함수가 최대한 prior과 같도록 만들어 주는 역할을 한다.

$$L(\phi, \theta, x_i) = \frac{1}{2} \sum_{j=1}^D \left[\log(\sigma'_{i,j}) + \frac{x_{i,j} - \mu'_{i,j}}{\sigma'_{i,j}} \right] - \frac{1}{2} \sum_{j=1}^J (1 + \log \sigma_{i,j}^2 - \mu_{i,j}^2 - \sigma_{i,j}^2) \quad (4)$$

사전 확률 분포와 출력의 확률 분포를 가우시안(Gaussian) 분포로 가정했을 때, Formula (3)은 Formula (4)과 같이 나타낼 수 있다. Formula (4)에서 $\mu'_{i,j}$ 와 $\sigma'_{i,j}$ 는 각각 출력 데이터의 평균과 분산을 의미하며, D 는 입/출력 데이터의 차원(벡터)이다. 또한 $\mu_{i,j}$ 와 $\sigma_{i,j}^2$ 는 각각 은닉 변수의 평균과 분산을 나타내고, J 는 은닉 변수의 차원(벡터)을 의미한다.

아래 Figure 6은 LSTM-VAE의 구조이다. LSTM과 그 뒤에 오는 linear들로 z (평균, 분산)를 구한다. 이때 Progress-based Prior들의 분포가 $\Sigma_p = 1$ 인 $N(\mu_p, \Sigma_p)$ 로 각각의 평균을 모두 다르게 설정한다. 이렇게 각 time step 마다 분포를 다르게 하여 시계열 데이터의 시간 의존성(temporal dependency)을 반영할 수 있다. 이렇게 인코더에서 생성된 평균 μ_{x_t} 와 분산 Σ_{x_t} 의 분포에서 z 를 랜덤하게 샘플링하여 디코더의 LSTM으로 전달된다. 최종적으로 Loss function은 Reconstruction Term과 Regularization Term의 합으로 구성되어 이 값들을 최소화 하는 방식으로 최적의 ϕ, θ (파라미터)를 찾는 것이 목표이다. 이렇게 학습된 파라미터를 바탕으로 Formula (1)을 계산하여 Anomaly score를 구하고 그 값이 임계값(threshold)보다 크면 anomaly로 그렇지 않으면 normal로 판단한다(Park D et al., 2018).

결과적으로 VAE는 Autoencoder(AE)와 비슷한 구조를 갖지만 AE는 차원을 줄여 잠재공간인 하나의 값으로 저장하고 다시 생성하는 방식인데 반해, VAE는 가우시안 확률 분포를 만들어 내어 특징을 저장하고 거기서 디코더를 통해 데이터를 다시 생성하는 방식에서 차이를 보인다.

4. Results

Figure 7은 LSTM-VAE(Long Short-Term Memory Variational Autoencoder) 모델의 손실함수(loss function)값의 분포를 나타낸다. 이를 통해 이상을 식별하기 위한 적절한 임계값(threshold)을 결정하기 위해 손실 함수의 누적 확률 분포가 99%이상이 되는 지점인 0.04를 임계값(threshold)으로 설정했다. 테스트 데이터를 통해서 재

구성 손실을 계산하였을 때 해당 임계값을 넘을 경우 이상으로 판별한다.

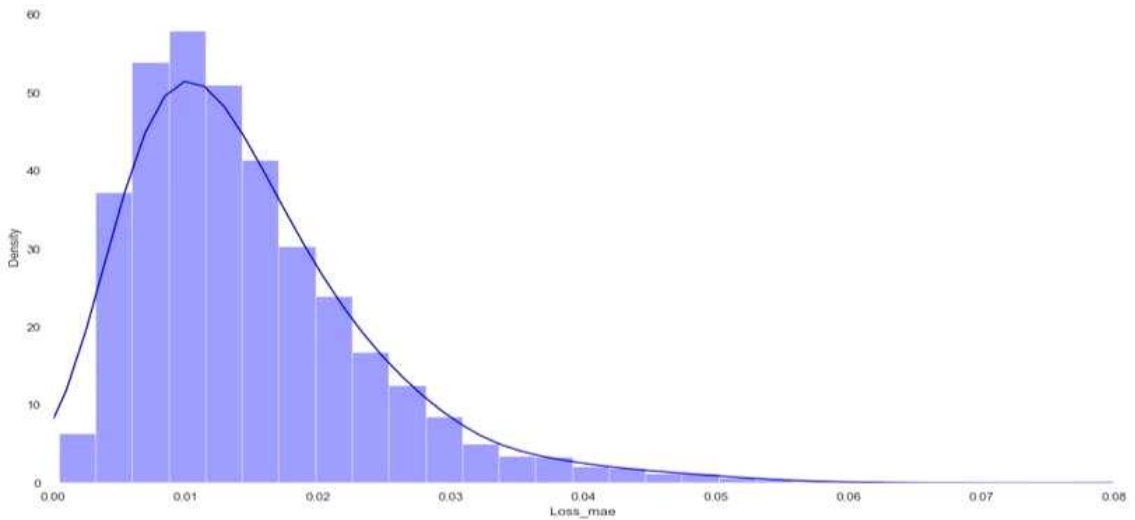


Figure 7. Distribution of loss function value

Figure 8 회전체불평형 실험과 Figure 9 축정렬불량 실험의 Loss value를 보면 2,000초가 지난 시점부터 값이 증가하였고 빨간색 선으로 표시한 임계값(threshold)을 뛰어넘어 이상(anomaly)이 관측됨을 알 수 있다.

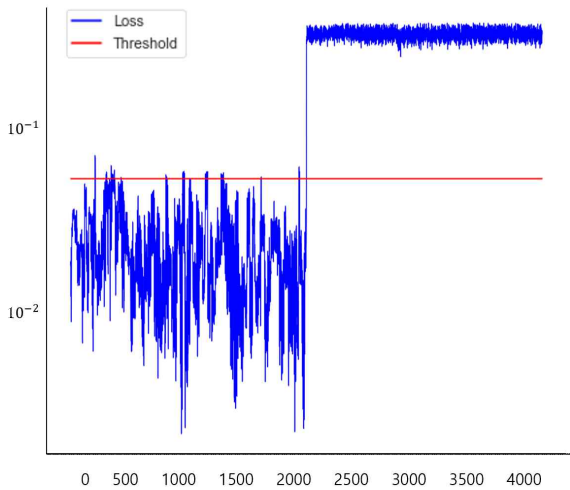


Figure 8. Loss value of rotating body failure

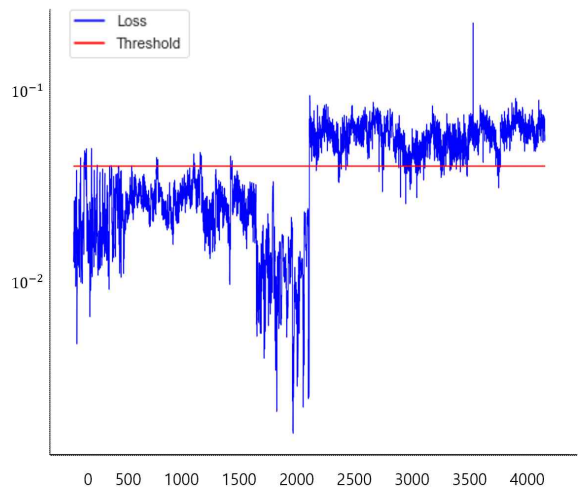


Figure 9. Loss value of axis alignment failure

마지막으로 모델의 성능을 평가하기 위해 다음 4개의 base-line 모델들을 추가해서 비교했다. ① Random은 클래스에 가중치를 부여하며 무작위로 분류한다. ② IF(Isolation Forest)는 데이터를 의사결정나무(Decision Tree)의 형태로 표현하여 분류하는 밀도기반의 이상탐지 모델이다. ③ OSVM(one-class support vector machine)은 영역기반 학습방법으로 결정경계에 초점을 맞춰 테스트 데이터가 학습된 영역 내에 존재하면 정상으로 그렇지 않으면 이상으로 판단한다. ④ AE(Autoencoder) 정상 데이터를 오토인코더를 통하여 원래의 데이터보다 작은 차원의 잠재

공간(latent space)에 압축하고 이를 디코더를 통해 복원해 원본과의 차이를 비교함으로써 이상탐지를 한다.

본 실험에서는 분류 모델의 성능평가 지표로 Accuracy, Precision, Recall, F1-score를 사용하였다. Accuracy는 정확도로 전체 예측한 것 중에 올바른 예측을 얼마나 했는지를 의미하고, Precision은 정밀도로 Positive로 예측한 것 중 실제로 Positive가 얼마나 되는지를 보여주는 지표로 얼마나 정밀한지를 알 수 있다. 또한 Recall은 실제 Positive인 것 중 모델이 Positive라 분류한 비율로 실제 Anomaly를 Anomaly로 예측하는 비율을 나타낸다. 마지막으로 F1-score은 Precision과 Recall의 조화평균을 구한 것으로 데이터가 불균형한 분류문제에서 Accuracy로 성능을 평가하기엔 편향성이 너무 크게 나타나기 때문에 주요 평가지표로 사용된다.

기계시설물 장치의 이상탐지에서는 시설물의 고장 발생 시 막대한 손해비용이 발생하기 때문에 이상 데이터를 분류하는 것이 정상 데이터를 분류하는 것보다 중요하다. 따라서 이상 데이터의 성능평가 지표를 기준으로 평가했다. 실험에서는 클래스별로 균일하게 데이터 수를 조정하였기 때문에 Accuracy로 성능을 보는 것이 좋지만 실제 현장에서는 클래스별 데이터 수의 불균형 문제가 지속적으로 발생하기 때문에 F1-score도 함께 고려하는 것이 좋다.

Table 3. Model performance evaluation of Random, IF, OSVM, AE and LSTM-VAE

Model	Experiment1 (Rotating Body Error)				Experiment2 (Axis Alignment Error)			
	precision	recall	accuracy	f1-score	precision	recall	accuracy	f1-score
Random	0.5218	0.5197	0.5214	0.5207	0.5238	0.5089	0.5185	0.5162
IF	0.8906	0.7849	0.8310	0.8344	0.7720	0.8897	0.8597	0.8267
OSVM	0.7988	0.8103	0.7492	0.8045	0.8007	0.6989	0.7315	0.7463
AE	0.8879	0.8963	0.8906	0.8921	0.9305	0.9158	0.9274	0.9231
LSTM-VAE	0.9791	0.9605	0.9700	0.9697	0.9799	1.0000	0.9898	0.9899

Table 3은 앞서 가정한 두 가지 실험에 대한 결과를 나타낸다. Experiment1은 처음 2,000초 동안 정상 모터로부터 데이터가 수집된 후 그 뒤 2,000초는 회전체 불평형(Rotating Body Error)이 발생하여 고장 난 상황을 가정하였고, Experiment2는 2,000초 동안의 정상 상태에서의 상황 이후 2,000초 동안의 축정렬 불량(Axis Alignment Error)이 발생한 고장 상황을 가정하였다. Table 3의 결과를 보면 이상 데이터의 분류 성능은 Accuracy와 F1-score 모두 Random, IF, OSVM, AE보다 LSTM-VAE 모델의 이상 탐지 성능이 더 우수한 결과를 보였다.

5. Conclusions

기계시설물 장치가 설치된 현장에서는 고장 상태의 데이터보다 정상 상태의 데이터 수집이 용이하기 때문에 클래스 불균형 상태에서도 적용할 수 있는 비지도 학습 방법이 사용되었다. 또한 기계시설물의 이상 탐지를 위해 지하철역사 내 모터 장비의 시계열 데이터를 활용하여 이상 탐지를 실시했다. 모터의 고장은 교통, 소방 방재, 급배수 등의 전체 시스템에도 악영향을 줄 수 있기 때문에 고장을 탐지하고 예지하는 기술이 필요하다. 모터의 주요 원인 고장은 베어링, 회전체, 벨트, 축 등에서 발생하며 본 연구에서는 회전체와 축에 문제가 있는 상황에 대해서 실험하였다.

회전체불평형은 회전체의 질량이 균일하지 못해 발생하는 일반적인 고장으로 질량의 불균형으로 인해 원심력이 크게 발생한다. 따라서 이상 발생 시 진동 가속도가 큰 폭으로 증가하였으며 전류값의 크기는 줄어든 것을 알 수 있

다. 측정렬불량은 보통 열화에 의해 느슨해지거나 모터와 공조팬 등의 정렬이 불량하여 발생하는 고장유형으로 고장 발생 시 진동 가속도와 전류값 모두 증가하는 것을 볼 수 있다. 또한 LSTM-VAE(Long Short-Term Memory Variational Autoencoder) 모델의 손실함수(loss function)값도 고장 발생 시 큰 폭으로 증가한 것을 확인 할 수 있다.

실험 결과를 보면 LSTM-VAE가 시간 정보를 반영하지 못하는 Random, IF, OSVM보다 뛰어난 성능을 보인 것을 알 수 있다. 이를 통해 시설물 데이터에서 시간 정보를 반영하는 것의 중요성을 알 수 있었다. 또한 AE는 시간정보를 반영하지만 이전 데이터를 기억하여 다음 데이터에 반영하는 시간의 연속성은 반영하지 못한다. LSTM-VAE는 LSTM 구조를 도입하여 이를 반영할 수 있었기에 AE보다 높은 성능을 나타낸 것으로 보인다. 결과적으로 LSTM-VAE 모델은 두 번의 실험에서 Accuracy가 각각 97%와 98.98%로 다른 base-line 4개의 모델들 보다 뛰어난 성능을 보였다. 이는 다차원 신호의 분포를 모델링 하고 예상 분포로 신호를 재구성하여 다변량 시계열 데이터의 특징을 효과적으로 추출했음을 의미한다.

본 연구를 통해서 모터는 고장이 발생하거나 성능이 저하되면 공통적으로 진동과 전류값의 변화가 발생하는 것을 확인했다. 따라서 모터에서 발생하는 진동, 전류 데이터를 기반으로 기계 장치의 이상 진단을 하면 안정적인 운용계획을 수립하고 시스템의 관리 효율성을 증대하는 효과를 기대할 수 있다. 또한 주요 사물인터넷(IoT) 솔루션 제공업체들과 협업을 통해 인공지능 솔루션을 개발하여 시설물 생애주기 중 불필요한 유지 보수비 발생을 방지하고 향상된 정확성으로 시스템을 구축할 수 있다. 이렇게 수작업으로 진행되던 점검 및 정비 인력 운영의 비효율성을 감소하여 산업의 생산성 증대 및 운영 경쟁력을 강화할 수 있는 효과가 기대된다.

본 연구는 비지도 기계학습 방법 중 하나인 LSTM-VAE 모델을 활용해 이상 탐지 시스템을 제안하였지만 다음과 같은 한계점을 지니고 있다. 현재 실험에서 사용한 데이터의 종류는 진동과 전류 데이터뿐이다. 시설물에 고장이 발생하면 진동과 전류뿐만 아니라 온도, 습도 그리고 미세먼지 등 다양한 변화가 발생 할 수 있다. 따라서 보다 다양한 센서로부터 수집된 다변량의 시계열 데이터를 활용해 이상 탐지를 수행하는 연구가 필요하다. 또한 이상 탐지 분야에서는 AE(Autoencoder)모델 기반의 방법들도 많이 연구되고 있지만, GAN(Generative Adversarial Network), CNN(Convolutional Neural Network), RNN(Recurrent Neural Network)등을 응용한 다양한 모델들이 소개되고 있다. 따라서 앞으로 다양한 센서로부터 수집된 데이터를 기반으로 다양한 종류의 기계학습 모델들과 비교를 통하여 학습 속도가 빠르고 정확도가 높은 모델을 찾는 연구가 진행 되어야 할 것이다.

REFERENCES

- Akçay S, Atapour-Abarghouei A, Breckon T P. 2019. Skip-ganomaly: Skip connected and adversarially trained encoder-decoder anomaly detection. 2019 International Joint Conference on Neural Networks (IJCNN). IEEE.
- An J. and Cho S. 2015. Variational Autoencoder based AnomalyDetection using Reconstruction Probability. Technical Report. SNU Data MiningCenter. 1-18.
- Bank D., Koenigstein N., and Giryas R. 2020. Autoencoders. arXiv eprints.
- Baur C., Wiestler B., Albarqouni S., and Navab N. 2018. Deep autoencoding models for unsupervised anomaly segmentation in brain MR images. In Proceeding of International MICCAI Brainlesion Workshop. 11383:161-169.
- Chalapathy R. and Chawla S. 2019. Deep Learning for Anomaly Detection: A Survey. CoRR, (abs/1901.03407).
- Chandola V., Banerjee A., and Kumar V. 2009. Anomaly detection : A survey, ACM computing surveys (CSUR), 41(3):15.

- DataRPM, P. 2017. Anomaly Detection & Prediction Decoded: 6 Industries. Copious Challenges. Extraordinary Impact. Technical Report.
- E.S. Lee, H.C. Bae, H.J. Kim, H.N. Han, Y.K., and Lee, J.Y. Son. 2020. Trends in AI Technology for Smart Manufacturing in the Future. *Electronics and Telecommunications Trends*. 35(1):60–70.
- Guo Y., Liao W., Wang Q., Yu L., Ji T., and Li P. 2018. Multidimensional Time Series Anomaly Detection: A GRU-based Gaussian Mixture Variational Autoencoder Approach. *Proceedings of the 10th Asian Conference on Machine Learning*, PMLR. 95:97–112.
- Harmeling S., Dornhege G., Tax D., Meinecke F., and Müller K R. 2006. From outliers to prototypes : ordering data. *Neurocomputing* 69(13–15):1608–1618.
- Hasan M., Islam M M, Zarif M I I, and Hashem M M A. 2019. Attack and anomaly detection in IoT sensors in IoT sites using machine learning approaches. *Internet of Things*. 7:100059
- Hodge V. and Austin, J. 2004. A survey of outlier detection methodologies. *Artificial Intelligence Review* 22(2):85–126.
- J.S. Kim, C.K. Lee, K.N. Kim, C.W. Kim, H.M. Song, S.S. Ahn, J.W. Oh, H.S. Jo, and S.S. Han. 2021. A Case Study on the Target Sampling Inspection for Improving Outgoing Quality. *Journal of Korean Society for Quality Management* 49(3):421–431.
- K.H. Sun, H Huh, Tama, B A, S.Y Lee, J.H. Jung, and S. Lee, 2020. Vision-Based Fault Diagnostics Using Explainable Deep Learning with Class Activation Map. *IEEE Access* 8:129169–129179.
- Li D, Chen D, Goh J, and Ng S K. 2018. Anomaly detection with generative adversarial networks for multivariate time series. *arXiv preprint arXiv:1809.04758*.
- M.J. Kim, Y.H. Park, T.K. Kim, J.S. Jung. 2019. A Study on the Prediction Diagnosis System Improvement by Error Terms and Learning Methodologies Application. *Journal of Korean Society for Quality Management*. 47(4):783–793.
- M.K. Seo, and W.Y. Yun. 2019. Condition Monitoring and Diagnosis of a Hot Strip Roughing Mill Using an Autoencoder. *Journal of Korean Society for Quality Management* 47(1):75–86.
- Maesschalck R D, Jouan-Rimbaud D, and Massart D L. 2000. The Mahalanobis Distance. *Chemometrics and Intelligent Laboratory Systems* 50(1):1–18.
- Malhotra P, Vig L, Shroff G, and Agarwal P. 2015. Long short term memory networks for anomaly detection in time series. In *Proceedings of Presses Universitaires de Louvain* 89:89–94.
- Munawar A, Vinayavekhin P, and Magistris G. D. 2017. Limiting the reconstruction capability of generative neural network using negative learning. 2017 *IEEE 27th International Workshop on Machine Learning for Signal Processing (MLSP)*. IEEE.
- N.Y. Choi, and W.H. Kim. 2019. Detecting user behavior anomalies using Generative Adversarial Networks. *Intelligence Information Research* 25(3):43–62.
- Park D, Hoshi Y, and Kemp C C. 2018. A Multimodal Anomaly Detector for Robot-Assisted Feeding Using an LSTM-Based Variational Autoencoder. *IEEE Robotics and Automation Letters* 3(3):1544–1551.
- Perera P, Nallapati R, and Xiang B. 2019. Ogan: One-class novelty detection using gans with constrained latent representations. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.
- Pokrajac D, Lazarevic A, and Latecki L. 2007. Incremental Local Outlier Detection for Data Streams. *CIDM Conference*.
- Rousseeuw P. and Leroy A. 2003. *Robust Regression and Outlier Detection*. Wiley.
- Son V H, Daisuke U, Kiyoshi H, Kazuki M, Pranata S, and Shen S. M. 2019. Anomaly detection with adversarial dual autoencoders. *arXiv eprint*.

- Wei Q, Ren Y, Hou R, Shi B, Lo J Y, and Carin L. 2018. Anomaly detection for medical images based on a one-class classification. In *Proceeding of Medical Imaging 2018: Computer-Aided Diagnosis*.
- Yamanaka Y, Iwata T, Takahashi H, Yamada M., and Kanai S. 2019. Autoencoding binary classifiers for supervised anomaly detection. *Pacific Rim International Conference on Artificial Intelligence*.
- Zong B, Song Q, Min M R, Cheng W, Lumezanu C, Cho D, and Chen H. 2018. Deep autoencoding gaussian mixture model for unsupervised anomaly detection. In *Proceeding of International Conference on Learning Representations*.

저자소개

- 서재홍** 연세대학교에서 컴퓨터와 산업공학 학사과정을 졸업하고, 동대학원에서 산업공학 석사과정에 재학 중이다. 주요 관심분야는 플랫폼, 데이터사이언스, 이상탐지, 기계학습이다.
- 박준성** 연세대학교 산업공학과를 졸업하고, 동대학원에서 박사과정에 재학 중이다. 주요 관심분야는 소비자 행동, 플랫폼 이론 등이다.
- 유준우** 연세대학교 산업공학과를 졸업하고, 동대학원에서 박사과정에 재학 중이다. 주요 관심분야는 기술경영, 경영관리이다.
- 박희준** George Washington University에서 공학경영 전공으로 박사학위를 취득하고 Marymount University 경영학대에 재직하였으며, 현재 연세대학교 산업공학과에 재직 중이다. 주요 관심 분야는 플랫폼 기반의 혁신전략, 비즈니스모델 개발, 전략 수립 및 성과평가 방법론 개발 등이다.