

기계학습을 활용한 레이더 공기건조기 이상 탐지

김요진¹, 강필성^{2*}, 염동원¹, 김건우¹
¹한화시스템, ²고려대학교 산업경영공학과

Anomaly Detection of Air Dryer for Radar based on Machine Learning Algorithms

Yojin Kim¹, Pilsung Kang^{2*}, Dong-Won Yeom¹, Kun-Woo Kim¹
¹Hanwha System
²School of Industrial Management Engineering, Korea University

요약 무기체계가 첨단화됨에 따라 기존의 사후정비나 예방정비를 통해 무기체계를 운용 및 정비하는 것은 선제적 관리측면에서 한계점이 존재한다. 이를 극복하기 위해 국방부에서는 「국방개혁 2020」을 통해 빅데이터 기반의 총수명주기 관리체계의 중요성을 강조하였고, 국방업무 전분야에서 상태기반정비(CBM: Condition Based Maintenance, 이하 CBM)를 적용하는 사업이 추진되고 있다. 하지만 보안규제 및 아날로그식 장비운용 등으로 인해 데이터 수집 및 활용에 제약이 있고 이로 인해 빅데이터 기반의 CBM을 적용하는데 어려움을 겪고 있다. 본 연구에서는 이러한 제약으로 인해 충분히 확보되지 않은 장비의 운용데이터를 이용하여 레이더의 고장방지를 위한 공기건조기의 이상탐지 모델 구축 사례를 제시하며, 운용데이터와 고장이력이 충분하지 않은 여건에서의 무기체계의 이상탐지 방법을 제안한다. 6개 부대의 운용데이터로 확보한 정상데이터를 이용하여 비지도학습 알고리즘 기반의 여러 가지 이상탐지 모델을 구축한 후 최적의 모델을 선정하는 작업을 수행하였다. 또한, 부대별로 정상상태에서의 오알람율을 낮추고 최대한 빠르게 이상을 탐지할 수 있도록 최적의 임계치 선정 및 고장알람기준을 정의하는 작업을 진행하였다. 본 연구결과는 무기체계의 운용데이터를 이용하여 기계학습 기반의 CBM 모델을 확보하는 실증사례를 확보함으로써 추후 다른 무기체계에서도 인공지능 적용을 통한 기계학습 방법론의 적용 가능성을 확인하고 기반을 마련하였다.

Abstract As weapon systems become advanced, it is difficult to operate and maintain them through existing post-maintenance or preventive maintenance methods. To overcome this, the Ministry of National Defense has emphasized the importance of "management of the total life cycle system" through "Defense Reform 2020," and a project is being promoted to apply big-data-based state-based maintenance (CBM) in all areas of defense work. However, due to security regulations and the operation of analog facilities, data collection and use are limited, and it is difficult to apply big-data-based CBM. This research shows a case of constructing an abnormal detection model of an air dryer to prevent radar failure using the operation data of facilities that are not sufficiently secured due to the restrictions mentioned. An abnormal detection method is proposed for a weapon system under conditions with insufficient operation data and failure history. After acquiring normal operational data from six military units, an anomaly detection model was made using various anomaly detection techniques from unsupervised learning, and an optimal model was selected. In addition, an optimal threshold and alarm criteria were selected to decrease the false alarm rate in normal conditions for each unit and to detect abnormalities as quickly as possible. The results of this study confirmed the possibility of applying anomaly detection through AI application to another weapon system and laid the foundation of securing empirical cases, CBM models using AI, and data analysis techniques with actual operation data.

Keywords : Unsupervised Learning, Anomaly Detection, Weapon System, CBM, Machine Learning, AI

*Corresponding Author : Pilsung Kang(Korea University)

email: pilsung_kang@korea.ac.kr

Received December 22, 2022

Revised January 25, 2023

Accepted March 3, 2023

Published March 31, 2023

1. 서론

1.1 연구 배경

최근 무기체계의 구조와 기능이 다양해지고 복잡해지면서 기존의 정비체계와 인력만으로는 새로운 무기체계를 효율적으로 운용하고 정비하는데 한계가 있다[1]. 이를 극복하기 위해서 사후정비나 예방정비가 아닌, 장비의 상태를 분석하여 고장을 예방하고 고장시점을 예측하는 상태기반정비의 도입을 국방 전분야에서 고려하고 있다[2]. 상태기반정비(CBM)란 실시간으로 수집된 데이터를 기반으로 장비의 정상작동 여부를 판단하고 고장이 발생하기 전에 정비가 이뤄질 수 있도록 하는 정비형태로 장비의 가동률향상과 정비비용 절감효과를 얻을 수 있다[1,3].

최근 국방부에서 「국방개혁 2020」을 발표하면서 국방업무 전분야에서 빅데이터 기반 총수명주기관리 등의 사업이 추진되었고, 방위사업 품질관리규정 개정 및 총수명주기관리업무훈령이 제정되면서 무기체계에서의 상태기반정비의 중요성은 날로 높아지고 있다[4-7].

지금까지 무기체계에 적용되고 있는 CBM은 주로 함정과 항공의 엔진, 터빈등과 같은 기계적 결합 검출에 집중되어 있어, 레이더와 같은 전자 장비에 대한 상태기반 정비연구 및 사례가 부족한 실정이다[2,4,8]. 전자장비에 대한 고장방지 연구를 위해 레이더 도파관의 결빙을 방지하여 고장을 예방하는 CBM 연구가 진행되었다[8]. 해당 연구에서는 도파관의 결빙을 발생시키는 공기건조기의 고장유형을 분석하고 각 고장 유형별로 모의실험을 통해 정상데이터와 고장데이터를 획득한 후 정상과 이상을 구분하는 고장분류모델을 확보하였다[8].

하지만 실제 운용데이터의 경우 고장이력데이터가 정상데이터 대비 부족한 경우가 대부분이기 때문에 실제 운용데이터를 사용하여 고장분류모델을 구축할 경우, 정상과 고장간의 데이터불균형(data imbalance) 문제가 발생하게 되어 분류모델이 제대로 학습하지 못하는 문제가 발생한다[9].

무기체계의 경우 보안규제 및 이날로그식 장비 운영으로 데이터 수집이 어렵고 데이터 활용에도 한계가 있다. 본 연구의 대상인 천마무기체계 또한, 보안규제 등으로 인해 지금까지 데이터 수집이 이뤄지지 않아 연구가 진행되는 동안 수집된 운용데이터를 이용하여 장비의 고장을 방지하는 이상탐지 모델을 확보하여야 했다. 이런 경우 고장데이터를 수집기간동안 획득하는데 한계가 있으므로 아래 Fig. 1(a)과 같이 정상과 고장을 학습하는 지

도학습(supervised learning)기법의 분류(classification) 기법을 적용하기보다는 아래 Fig. 1(b)와 같이 확보 가능한 정상데이터만을 학습하는 비지도학습(unsupervised learning) 방법의 이상탐지(anomaly detection) 기법을 적용하는 것이 적절하다[10]. 따라서 본 연구에서는 무기체계의 CBM의 방안으로 지도학습 기법이 아닌 비지도학습 기반의 이상탐지 기법을 제안하고자 한다.

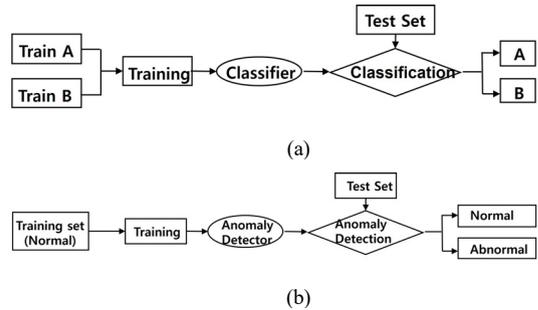


Fig. 1. Supervised learning vs Unsupervised learning
(a) Supervised learning (b) Unsupervised learning

1.2 연구대상

단거리 지대공 유도무기 천마(K-SAM, Soream Surface to Air Missile system)는 한국군 방공의 취약성을 보완하기 위해 개발한 최초의 국산 방공 유도무기이다. 탐지 및 추적 장치와 유도탄 8기, 사격통제장치를 단일궤도형 장갑차량에 탑재한 유도무기로 아래 Fig. 2과 같은 구조로 구성되어 있다. 탐지레이더는 이동하는 표적을 탐지해 위협평가 및 탐지 간 추적을 실시하고 추적레이더는 탐지레이더가 찾아낸 표적을 추적하며, 유도탄 발사 후 유도탄 추적과 유도명령 송신 기능을 수행한다.

천마체계는 1문당 약 000억 원이며 000여 부대에 설치되어 있다. 따라서 천마를 대상으로 CBM 적용 및 확대 전파할 경우, 큰 경비 절감을 기대할 수 있다.

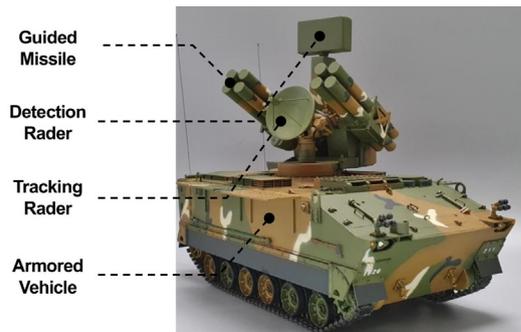


Fig. 2. K-SAM

천마의 상태기반정비를 위하여 CBM의 적용대상을 선정하는 작업을 진행하였다. 실제 정비이력과 군부대의 요구사항을 토대로 고장발생의 주요원인이 되는 공기건조기를 CBM의 대상으로 삼았다.

공기건조기는 천마레이다의 도파관내 습기를 제거하고 이물질 유입을 방지하는 역할을 하는 장비로 고장으로 인해 정상운용이 어려운 경우, 불안정한 온/습도 제어로 인해 천마의 안테나 조립체, 고주파헤드, 고주파전환스위치 등의 주요품목 고장과 도파관내 결로를 발생시켜 천마체계 운용에 치명적인 영향을 끼치게 된다. 따라서, 공기건조기의 상태를 감시하여 적기에 고장을 식별하고 사전예측을 통한 운용성을 확보할 수 있도록 이상탐지방법을 확보하는 것을 본 연구의 목적으로 삼았다.

본 연구에서는 6개 부대에 설치된 천마무기체계 공기건조기의 운용데이터를 수집하여 정상데이터만을 추출한 후 이상탐지의 여러 기법을 적용하여 이상탐지모델을 구축한 후, 최적의 알고리즘을 선정하는 작업을 수행하였다. 또한, 각 부대별로 최적의 시점에서 이상 알람이 제공될 수 있도록 임계치 선정 작업과 고장판단기준을 선정하는 작업을 수행하여 그 결과를 확인하였다.

본 논문의 구성은 다음과 같다. 2장에서는 천마공기건조기 장비에 대한 소개와 여러 가지 이상탐지기법에 대해 살펴본다. 3장에서는 공기건조기의 이상탐지를 위한 방안을 설명하고 4장에서는 연구수행결과를 확인한 후 각 부대별 확보한 최종모델결과를 살펴본다. 마지막으로 5장에서는 본 연구결과를 요약하면서 한계점을 살펴보고 추후 연구방안에 대해 기술하였다.

2. 이론 연구

2.1 이상탐지(anomaly detection) 방법론

이상탐지 방법은 데이터를 기반으로 정상과 이상을 구분하는 기계학습(machine learning)의 한 범주로서 정상과 고장에 대한 정답(label)이 모두 포함된 데이터 셋을 이용하여 정상과 고장을 분류하는 지도학습 기반 방법론과 이상데이터가 없는 상황에서 정상데이터만을 활용하여 정상 패턴을 학습한 뒤 테스트데이터에 대해 정상과 다른 패턴의 데이터를 고장으로 탐지하는 비지도 학습의 방법론으로 나눌 수 있다[10].

대부분의 실제 이상탐지 적용사례를 살펴보면 다수의 정상데이터 셋과 소수의 고장데이터 셋으로 데이터가 이뤄져 있다. 예를 들어 제조업의 경우 고장발생 시 일정

및 비용 측면에서 큰 손해가 발생하므로 일정한 주기로 설비를 점검하는 등의 정비활동을 통해 설비의 정상상태를 지속적으로 유지하고자 한다. 따라서 장비의 대부분의 데이터 셋은 정상상태의 데이터이고 불가피하게 발생된 고장에 대해서만 소수의 고장 데이터 셋을 가지게 된다. 이렇게 각 범주별 데이터의 절대적인 수의 차이가 큰 경우 데이터 불균형 문제가 발생하게 되어 분류모델은 다수의 데이터에 편향된 분류경계선을 갖게 된다. 이런 경우 다수의 범주의 정상데이터는 정상으로 잘 분류하지만 소수범주의 고장데이터를 정상으로 오분류할 가능성이 커지게 된다[9,10]. 또한, 본 연구대상처럼 데이터 수집기간 동안에 고장이 발생하지 않아 고장데이터가 없는 경우에는 학습 시 고장 데이터 셋을 요구하는 지도학습 기법을 적용하기 어렵다.

이에 반해, 비지도학습의 이상탐지 기법은 정상 데이터 셋의 패턴(pattern)을 모델이 학습하여 학습된 정상 데이터 셋과 다른 패턴의 데이터를 고장으로 탐지하는 기법이다. 따라서 고장이력 데이터가 없는 경우도 모델의 학습이 가능하여 본 연구대상인 천마 공기건조기에 적합한 이상탐지 방법론이다.

비지도학습의 이상탐지 기법을 장비에 적용 시 다변량 분석을 통해 장비의 이상 정도를 이상스코어(anomaly score)로 도출하고 고장 수준에 대한 이상스코어의 임계치(threshold)을 설정하여 임계치보다 낮은 이상스코어를 갖는 장비의 상태를 정상, 임계치보다 높은 이상스코어를 가질 때를 이상으로 탐지한다. 실제 실무에 적용된 비지도학습의 이상탐지기법은 고장이 발생하기 전에 최대한 빨리 이상을 탐지하고 알람을 제공하여 조치를 취할 수 있는 시간을 제공하는 것을 목적으로 한다.

2.2 기계학습 기반 이상탐지 기법

천마 공기건조기에 이상탐지를 위해 여러가지 비지도 학습의 이상탐지 기법을 적용하여 최적의 모델을 찾는 작업을 진행하였다.

2.2.1 Isolation Forest(IF)

Isolation Forest는 다수의 범주에 속하는 정상 데이터는 다른 데이터들과 고립시키기 어렵고 이에 반해 소수범주의 이상데이터는 다른 데이터들과 고립시키기 쉽다는 원리를 이용하여 이상탐지를 하는 모델로 랜덤포레스트의 의사결정나무와 같이 iTree를 여러 번 반복 사용하여 앙상블하는 기법이다[11]. iTree를 이용하여 관측치를 고립시키는데 이때 이상관측치의 경우 적은 횟수의

분기만으로도 고립이 가능하여 루트노드(root node)로부터 관측치의 말단노드(leaf node)까지 짧은 경로길이를 갖고, 정상관측치는 고립시키기 위해서는 더 많은 분기점이 필요하므로 더 긴 경로의 길이를 갖는다. 각 iTree의 평균 경로길이를 이용하여 이상스코어로 변환한다. 짧은 경로길이의 이상관측치는 큰 값의 이상스코어를 갖고, 긴 경로길이의 정상관측치는 작은 값의 이상스코어를 갖게 된다.

2.2.2 Local Outlier Factor(LOF)

관측치를 포함한 군집과 주변의 정상의 이웃데이터로 이뤄진 군집의 상대적인 밀도차이를 평가하여 지역적인 outlier를 탐지하는 방법론이다.

이웃 데이터와 절대적 거리가 비교적 가까워 정상치로 분류할 수 있는 경우라도 이웃 데이터들 간의 군집의 밀도를 비교해 봤을 때, 관측치를 포함한 군집의 밀도가 이웃 데이터들로 이뤄진 군집보다 상대적으로 낮은 밀도의 군집에 속하게 되면 이를 이상치(outlier)로 판단하는 기법이다[12].

본 방법론은 LOF라는 factor로 이상스코어를 산출한다. LOF의 factor값은 Eq. (1)과 같다.

$$LOF_k(p) = \frac{\sum_{o \in N_k(p)} \frac{lrd_k(o)}{lrd_k(p)}}{|N_k(p)|} \quad (1)$$

where, lrd denotes local reachability, N denotes number of neighbors of p.

Eq. (1)에서 lrdk는 군집의 Local Reachability Density를 이야기하며, lrdk(o)는 이웃 데이터로 이뤄진 군집의 밀도값이고 lrdk(p)는 관측치를 중심으로 하는 군집의 밀도값으로 이웃의 주변 데이터의 상대적 밀도보다 관측치의 밀도가 상대적으로 낮은 경우 큰 값의 LOF 값을 갖게 된다.

LOF는 전체 데이터분포를 고려하여 outlier를 탐지하는 것이 아닌, 관측치 이웃 데이터와의 상대적인 관계를 고려하여 outlier를 탐색할 수 있는 방법론으로 지역적인 특징을 고려할 수 있는 장점을 가지고 있다.

2.2.3 One-Class SVM (OC SVM)

OC-SVM은 정상데이터를 커널을 이용하여 feature space로 차원을 변환(mapping)한 후 선형의 분류 경계면을 찾고 해당 경계면이 원점에서 최대한 멀어지게 모델을 학습 시켜 정상데이터의 분류 경계면을 얻는다. 이

렇게 획득한 분류 경계면 기준으로 정상 영역 이외의 영역의 데이터를 이상으로 탐지하는 방법론이다[13-15].

$$\begin{aligned} \min_W \frac{1}{2} \|W\|^2 + \frac{1}{\nu l} \sum_{i=1}^l \xi_i - \rho \quad (2) \\ s.t. W \cdot \Phi(X_i) \geq \rho - \xi_i \\ i = 1, 2, \dots, l, \xi_i \geq 0 \end{aligned}$$

where, W, ν , ξ , ρ , l denotes parameters.

One-Class SVM의 목적함수는 Eq. (2)와 같고 cost의 역수인 ν 가 작을수록 margin을 벗어나는 데이터에 대해 더 큰 페널티를 부여하여 서포트 벡터의 개수와 margin은 줄어들고, 반대로 ν 값이 클수록 margin이 넓어지며 서포트 벡터의 개수가 늘어난다. 따라서 ν 값이 작을수록 경계면의 복잡도가 높아지고 ν 값이 클수록 경계면의 복잡도가 낮아지는 성질을 가진다[14].

2.2.4 Principal Component Analysis(PCA)

PCA기법이란 데이터의 분산이 최대가 되는 축(주성분)을 찾아 데이터를 projection 시킨 후, feature extraction하는 기법이다[16,17].

이러한 PCA 기법을 이용하여 정상의 데이터를 대상으로 차원을 변환시킨 후, 다시 데이터를 복원시키는 과정을 모델이 학습하게 하여 정상과 다른 관측치가 입력되었을 때에는 모델이 제대로 복원하지 못하는 원리를 이용하여 이상탐지를 하는 방법론이다. 이상스코어는 원래의 데이터와 복원된 데이터 사이의 차이로 산출되고 해당 차이를 복원오차(reconstruction error)라 말하며, 이상데이터 일수록 복원오차가 크고 큰 이상스코어를 갖는다[17].

2.2.5 LSTM Auto Encoder(LSTM-AE)

LSTM-AE란 Encoder-Decoder 구조를 통해 정상데이터를 latent vector로 압축시키고, 다시 복원하는 과정을 모델이 학습하게 하는 기법으로 이때 시간에 따른 데이터 변화를 학습할 수 있도록 LSTM 네트워크 구조를 적용한 기법이다.

정상데이터만을 학습데이터 셋으로 사용하여 모델을 입력데이터와 복원데이터 사이의 차이가 최소화 되도록 학습시켜 정상관측치에는 낮은 복원오차를 도출하고 정상과 다른 패턴의 이상 관측치가 입력될 경우에는 높은 복원오차를 갖는 원리를 이용하여 이상을 탐지하는 방법론이다[18].

3. 공기건조기 이상탐지 방안

3.1 공기건조기 구조 및 고장원인 분석

공기건조기는 도파관에 유입되는 공기를 건조시키는 건조제를 담은 2개의 건조통 유닛(unit)과 건조통의 건조와 재생의 작동을 교체시켜주는 캠축, 그리고 습기를 흡수한 건조통의 수분을 제거하는 재생작업을 도와주는 히터로 구성되어 있다.

아래 Fig. 3과 같이 두 개의 건조 유닛 중 한 개의 건조 유닛에서는 외부 공기의 습기를 건조제를 이용하여 흡수한 후 도파관에 건조된 공기를 주입하는 건조(흡수) 작업을 진행하고, 다른 건조 유닛에서는 흡수 작업 후의 건조제의 수분을 히터를 이용하여 제거하는 재생작업을 진행한다. 각 건조 유닛은 두 시간 간격의 캠축의 교체작업을 통해 건조/재생을 반복하여 공기건조기의 습도를 일정하게 유지해 준다.

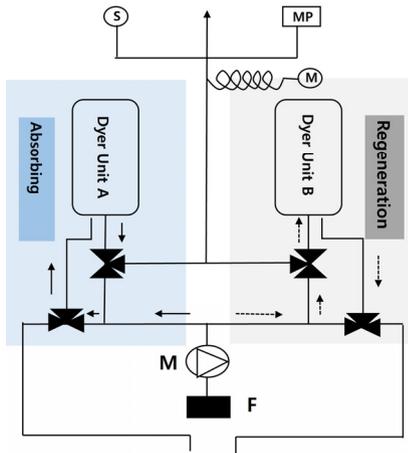


Fig. 3. Structure of Air dryer

공기건조기의 주요고장유형으로는 건조통 내부 건조제 노후화로 인한 건조통 불량, 제어모터 불량으로 인한 캠축고장, 히팅기 및 모터고장으로 인한 히터고장 등 있다.

3.2 수집 데이터 현황

총 여섯 개의 부대의 공기건조기의 운용데이터를 1월부터 8월까지 수집하였고, 실시간 측정을 위해 1초 단위의 데이터를 수집하였다.

공기건조기의 구조를 변경하지 않고 수집 가능한 인자 중 건조기의 이상에 직접적인 영향을 받는 내부 온습도

인자를 데이터 수집대상으로 삼았다. 또한, 내부 온습도는 외부환경에 영향을 받는다는 점을 고려하여 외부 온습도도 함께 수집대상으로 삼았다.

아래 Fig. 4는 건조기 운용시의 내/외부 온습도 그래프이다.

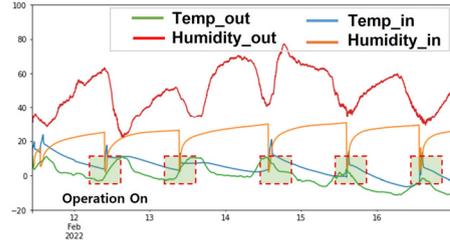


Fig. 4. Data pattern

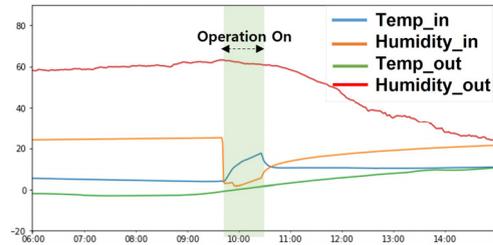


Fig. 5. Dry-on pattern

각 인자들의 패턴을 좀 더 자세히 살펴보면 위의 Fig. 5과 같이 건조기의 작동을 시작할 때, 내부 습도는 크게 떨어지고 내부 온도는 증가하는 것을 확인할 수 있으며, 건조기의 작동을 멈출 때는 습도는 크게 증가하고 내부 온도는 크게 감소하는 것을 확인할 수 있다.

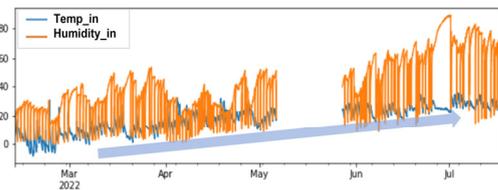


Fig. 6. Inner temperature/Inner humidity(example)

위의 Fig. 6은 실제 부대의 공기건조기의 내부 온습도 그래프이다. 위의 Fig. 4와 Fig. 6을 통해 알 수 있듯이 각 부대에서는 결로 방지를 위해 공기건조기를 1~2시간 정도 일정하지 않은 주기로 운용하고 있고 계절의 변화에 따라 온습도가 트렌드를 가지고 변화되고 있음을 확인할 수 있다.

또한, 센서 이상 등으로 인한 수집 문제로 발생한 긴 시간의 결측 구간은 보간하는 것이 오히려 데이터의 신뢰성을 낮춘다고 판단하여 해당 결측구간에 대해 보간하지 않고 모두 분석에서 제외하였다.

작동중인 건조기의 이상탐지를 위해서는 건조기 작동 시점의 데이터가 필요하다. 하지만 천마 공기건조기의 경우 Fig. 5과 같이 작동 on/off 구분 없이 데이터가 수집되고 있고 별도의 건조기의 작동 on/off 신호가 수집되고 있지 않은 상황이므로, 데이터 패턴을 분석하여 작동 on/off를 분리하는 작업이 필요하였다.

3.3 작동 on 데이터 셋 분리

내부 온습도는 공기건조기의 작동여부에 따라 크게 변화하는 것을 확인할 수 있었다. 아래의 Fig. 7과 Fig. 8를 살펴보면 건조기가 작동을 시작할 때 내부 습도의 변화율은 매우 큰 값의 음의 값을 갖고, 건조기가 작동을 멈출 때에는 내부 습도의 변화율이 주변보다 조금 더 큰 값의 양의 값을 갖는다는 것을 확인할 수 있다.

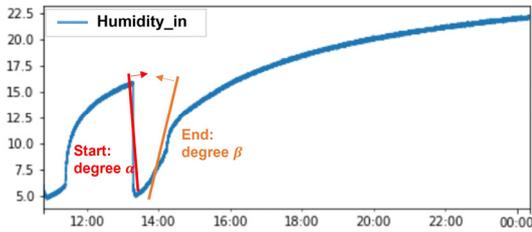


Fig. 7. Inner humidity pattern on dryer on

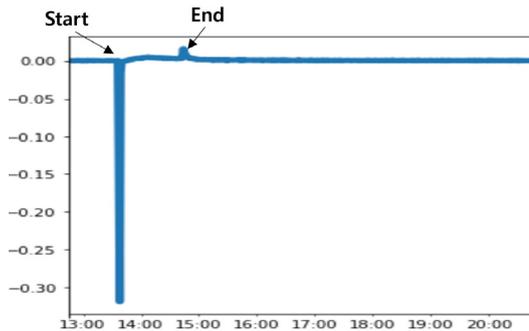


Fig. 8. Degree of inner humidity

따라서 아래와 같은 조건에 의해 작동구간을 분리하였다.

- 1) 건조기 on 시작 시점 : 기울기 $< \alpha$ ($\alpha = \mu - 3\sigma$)
 - 2) 건조기 on 멈춤 시점 : 기울기 $> \beta$ ($\beta = \mu + \sigma$)
- ※ μ : 내부습도기울기의 평균,
 σ : 내부습도기울기의 표준편차

3.4 정상데이터 셋 확보

비지도학습의 이상탐지는 정상을 학습하여 정상의 패턴을 벗어나는 관측치를 이상으로 탐지하므로 학습 셋인 정상패턴에 대한 정제 정도에 따라 이상탐지 성능이 좌우된다. 따라서 확보된 작동 on 데이터 셋에서 이상 작동으로 판단되는 시점들을 제외하여 정상데이터 셋을 확보하는 작업을 수행하였다. 아래의 Fig. 9~11과 같이, 이상 작동으로 보이는 작동 중 건조통 교체나 짧은 주기의 작동 on/off 및 이상치로 보이는 hunting 등을 제거하는 작업을 진행하였다.

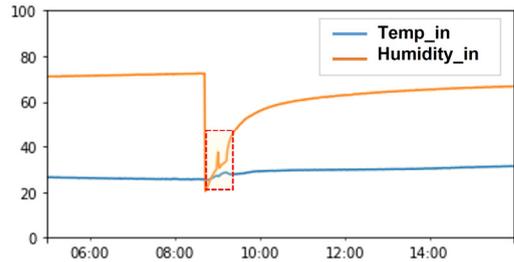


Fig. 9. Changing dryer unit on operation

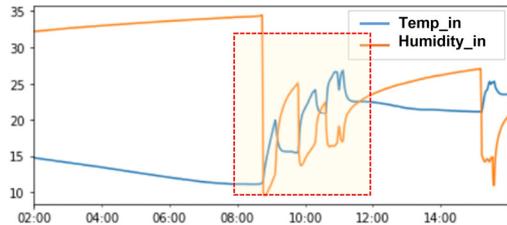


Fig. 10. Short cycle of on/off operation

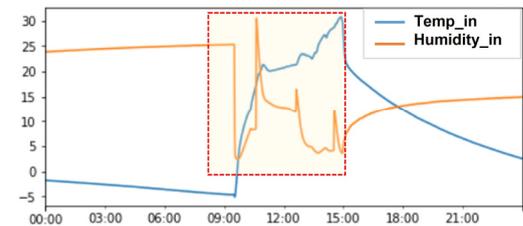


Fig. 11. Hunting

또한, 위에서 나열된 이상시점을 제거한 후에도 아래의 Fig. 12와 같이 정상패턴을 따르지 않은 데이터들이 존재하는 것을 확인하였다. 좀 더 정제된 정상패턴을 확보하기 위해 내부 온도는 정상 작동 시 지속적으로 온도가 증가한다는 성질을 이용하여 내부온도 변화율이 0보다 작은 시점을 제거하는 작업을 수행하였다. 그 결과 아

래의 Fig. 12과 같이 정제된 정상패턴을 가진 데이터 셋을 확보할 수 있었다.

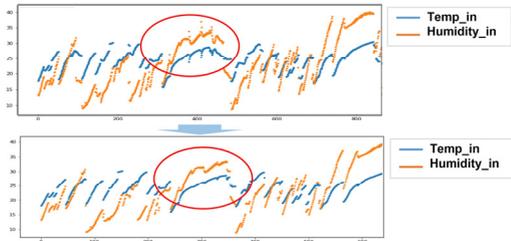


Fig. 12. Data cleansing by inner temperature change rate

3.5 변수 선정 (Feature selection)

수집된 변수들에 대해 분석 대상 변수 선정을 진행하였다. 일반적으로 변수 선정은 데이터의 사이즈를 줄여 계산 시간을 줄이고 모델 성능 관점에서 노이즈(noise)나 모델과 관계성이 적은 인자를 제거하여 모델의 성능을 향상시키는 기법이다[19]. 높은 성능의 이상탐지 모델을 확보하기 위해 고장 메커니즘과의 관련성, 데이터의 신뢰성, 변수 간 상관관계 등을 고려하여 내부 온습도, 외부 온습도 변수 중 분석 대상으로 삼을 변수를 선정하는 작업을 수행하였다.

공기건조기의 고장 메커니즘과 직접적인 연관성이 있는 내부 습도와 내부 온도를 주요변수로 먼저 선정하였다.

외부 온도의 센서는 외벽에 외부환경으로부터 보호하기 위해 제작된 투명 케이스 안에 설치되었다. 그런데 센서 케이스의 위치가 햇볕이 잘 드는 곳에 위치하여 직사광선으로 인한 온실효과의 영향으로 3월 이후 30℃이상의 고온의 데이터가 수집되어 데이터 자체의 신뢰성을 잃었고, 이상탐지 모델에 외부 온도를 변수로 포함시켰을 때 노이즈로 작용하여 모델의 성능이 오히려 낮아지는 것을 확인하였다.

외부 습도의 경우 내부 습도가 외부 습도의 영향을 받아 높은 상관관계를 가질 것이라고 예상하였지만 운용데이터를 확인해본 결과 일정 시간대에 지속적으로 건조기가 작동되는 것이 아닌 불규칙적인 작동운용으로 인해 외부 습도의 분포를 규정하는 것이 어려웠고 아래의 Fig. 13과 같이 실제 내부 습도와외 피어슨 상관계수(pearson correlation)의 값이 0에 가까운 값을 가지는 것으로 확인되었다. 따라서 주요변수와의 관련성이 낮은 외부 습도를 분석 대상에서 제외하였다.

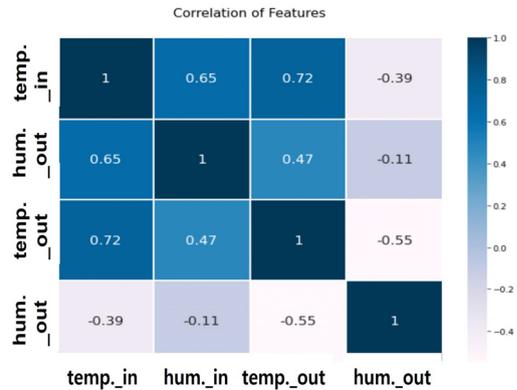


Fig. 13. Pearson correlation (ooo army unit)

위의 변수 선택 과정을 통해 공기건조기의 내부 온습도만을 분석변수로 사용하였다.

3.6 학습/테스트 데이터 셋 확보

이상탐지모델의 학습 및 성능검증을 위해 확보된 정상 데이터 셋의 80%를 학습에 사용하고 나머지 20%를 학습된 모델이 정상을 잘 탐지하는지 판단하는 테스트데이터 셋으로 사용하였다. 테스트 셋에는 정상데이터 셋뿐만 아니라 고장을 잘 탐지하는지를 평가할 고장데이터 셋이 필요하다. 하지만 본 연구의 데이터 수집기간인 1월~8월동안 6부대에서 공기건조기의 고장이 발생하지 않아 실제 고장이력을 확보하지 못하였다. 대신 실험실에서 공기건조기의 주요고장유형인 건조통 불량, 캠모터 고장, 히터고장 등의 상황을 모사하여 대표적 고장유형에 대한 고장데이터를 확보하였다. 3가지 고장유형의 데이터는 아래 Fig. 14~Fig. 16과 같다.

아래의 Fig. 14는 두 개의 건조통 중 한 개의 건조통 불량으로 인해 건조통 교체가 이뤄진 후 불량 건조통이 다른 건조통에 비해 낮은 습도 흡수 성능을 가지는 것을 확인할 수 있다.

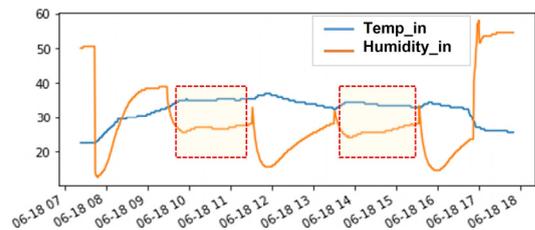


Fig. 14. Failure of dryer unit

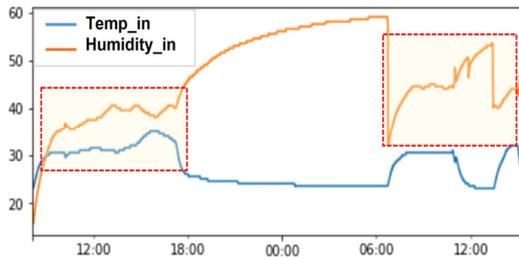


Fig. 15. Failure of cam motor

또한, 위의 Fig. 15의 캠모터의 고장으로 건조통의 교체가 되지 않은 경우, 한 개의 건조통이 재생작업이 없이 건조에만 쓰이게 되면서 습도가 낮아지지 않고 지속적으로 증가되는 것을 확인할 수 있다.

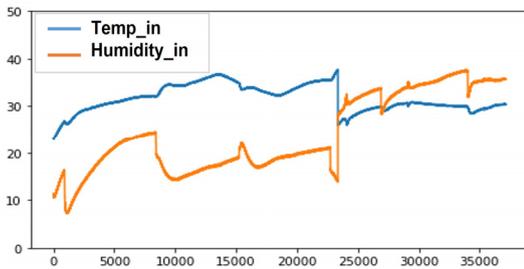


Fig. 16. Heater off

위의 Fig. 16의 히터 off를 살펴보면 초기 시점에는 이전에 재생작업을 마친 건조통이 제대로 건조작업을 수행하여 낮은 습도를 갖지만 이후 교체작업이 이뤄진 후 히터에 의한 재생작업이 수행되지 못한 건조통으로 인해 습도가 낮아지지 못하고 점점 증가하는 것을 확인할 수 있다.

총 6부대의 정상데이터 셋에 3가지의 고장유형을 개별로 붙여(concatenate)하여 총 18가지의 테스트 셋을 구축하였다.

3.7 알고리즘별 모델 구축

앞서 2장의 연구 배경에서 살펴본 다섯 가지의 이상탐지 기법을 천마 공기건조기에 적용 및 모델 구축한 후 성능검증을 통해 최적의 이상탐지 알고리즘을 선정하고자 하였다. 각 알고리즘의 파라미터(parameter) 설정은 아래 Table 1과 같다.

Table 1. Algorithm parameter

Algorithm	Parameters
IF	estimators = 100, Max sampling = 1,024
LOF	k = 100
OC SVM	kernel = RBF, $\nu=0.01$, gamma =0.01
PCA	n_principal components = 1
LSTM AE	dimension of latent vector = 1, hidden layers = 2, window size = 30, step size=15, adam optimizer, learning rate=0.01

3.8 이상스코어 임계치(Threshold) 선정

이상탐지 모델은 이상이 발생하기 전에 빠르게 이상을 탐지하여 알람을 제공하는 것을 목적으로 한다. 따라서 최적의 이상스코어 임계치를 선정하여 관측치가 임계치보다 높은 이상스코어를 갖는 경우에 해당 관측치를 이상으로 구분한다. 이를 위해 부트스트랩(Bootstrap)기법을 이용하여 각 부대별로 최적의 임계치를 선정하는 작업을 진행하였다. 아래의 Fig. 17과 같이 부트스트랩 기법을 이용한 임계치 설정은 정상상태의 이상스코어 표본 집합의 값들을 복원추출을 통해 여러 개의 배치(batch)를 생성한 후, 각 배치의 상위 99%의 이상스코어의 평균을 최적의 임계치로 결정하는 기법이다. 정상상태에서의 상위 99%의 이상스코어를 임계치로 가지므로 대부분의 정상데이터를 정상으로 구분하면서 정상을 벗어나는 이상 데이터에 대해 최대한 빠르게 이상을 감지할 수 있다. 본 연구에서는 100개의 batch를 사용하였다.

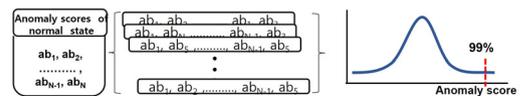


Fig. 17. Best threshold by using bootstrap method

3.9 이상 연속발생횟수(frequency) 기준 정의

센서의 오류나 환경 등의 이유로 장비가 정상상태인데도 불구하고 간헐적으로 데이터가 튀는 현상(hunting)이 발생할 수 있다. 따라서 데이터의 이상스코어가 임계치를 한번만 초과했을 때를 이상으로 판단하지 않고 연속으로 일정 기준횟수 이상을 초과할 때를 이상으로 판단하여 고장에 대한 알람을 제공하여야 한다. 연속발생횟수에 대한 기준이 없는 경우 아래의 Fig. 18과 같이 정상상태에서 간헐적으로 발생하는 이상데이터에 대해 오알

람 (false alarm)이 제공되어 현업의 장비관리자에게 혼란을 야기할 수 있다.

오알람 방지를 위해 연속발생횟수의 값을 너무 크게 설정할 경우, 진짜 이상이 발생했을 때 알람 제공 시간이 그만큼 늦어지는 단점이 있으므로 오알람 방지와 고장탐지 성능을 모두 고려하여 적절한 기준을 설정하는 작업이 중요하다.

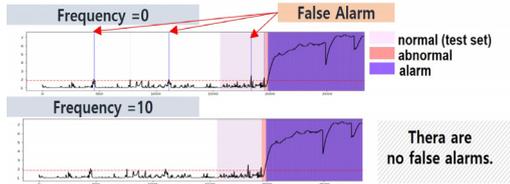


Fig. 18. Effect of setting the continuous frequency

4. 연구 결과

4.1 알고리즘별 성능검증

여러 번의 실험을 걸쳐 알고리즘별 천마 공기전조기에 적절한 파라미터(parameter)를 선정하였다. 이중 가장 모델 성능이 좋은 알고리즘을 최종 모델의 알고리즘으로 삼았다. 아래의 Fig. 19과 같이 정상상태의 학습구간과 테스트구간에서 낮은 이상스코어를 도출하고 고장구간에서는 정상구간과 확연히 차이가 나는 높은 이상스코어를 도출하는지를 기준으로 삼아 모델 성능을 검증하였다.

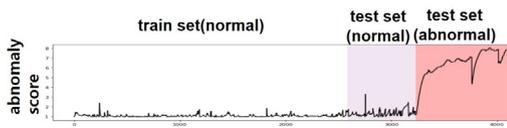


Fig. 19. Example of good performance model

각 알고리즘별 모델성능검증 결과는 아래의 Table 2와 같다. 테스트 셋의 정상과 고장구간 모두를 적절하게 탐지하는 경우 1점의 점수를 부여하고, 정상 또는 이상 탐지의 성능이 떨어지는 경우를 0.5점, 모두 탐지 못하는 경우에는 0점의 점수를 부여하여 알고리즘별 최종 성능 점수를 산출하였다.

산출된 결과에 따르면 총 18개의 테스트 셋에서 LOF가 17점의 점수를 획득하여 최종 모델의 알고리즘으로 선정되었다.

Table 2. Performance score of algorithms

Algorithm	Performance Score
IF	0
LOF	17
OC SVM	13
PCA	9.5
LSTM AE	11

4.2 부대별 최종 모델 확보

최종 이상탐지 알고리즘으로 선정된 LOF 모델을 이용하여 부대별 최종 모델을 확보하기 위해 최적의 임계치 선정과 이상 연속발생횟수 선정 작업을 각 부대로 진행하였다.

실제 고장이력을 확보한 경우에는 정상구간에서 이상구간으로 변화하는 전조구간을 확인할 수 있고, 전조구간에서의 이상을 탐지하여 알람을 제공할 수 있다. 또한, 고장 발생 시간으로부터 이상 알람을 제공하는 시점 사이의 차이를 모델 성능의 평가지표로 삼아 시점의 차이가 클수록 좋은 모델로 성능평가를 할 수 있다. 하지만 본 연구에서는 데이터 수집기간동안 공기전조기의 고장 발생이력이 없어 전조구간에서의 이상탐지가 불가하였다. 따라서 모의실험을 통해 확보한 고장데이터를 이용하여 실제의 정상과 실험실의 고장을 잘 분류하는지를 모델의 성능의 척도로 삼았다. 평가지표로는 고장구간을 고장으로 잘 구분하는지를 평가하는 정알람(true alarm/재현율)과 정상구간에서의 정상을 고장으로 분류하는지를 평가하는 오알람(false alarm/1-특이도)을 사용하였다.

각 부대로 최대 정알람율을 높이고 오알람율을 낮추는 임계치와 이상 연속발생횟수를 설정하여 최종 모델을 확보하였고, 아래 Table 3에서 확인할 수 있듯이 대부분의 부대에서 정알람율은 90% 이상의 값을 가지고, 오알람율은 0에 가까운 값을 갖는 것을 확인할 수 있다.

아래의 Table 3을 보면 고장유형 3번째의 히터 오프에서 정알람율이 낮게 나오는 것을 확인할 수 있다. 히터 오프의 경우 히터가 꺼진 시점의 건조 유닛의 건조통은 이전 단계에서 재생단계를 마친 후의 건조통이므로, 정상패턴과 비슷한 습도 흡수 패턴을 가지게 된다. 따라서 고장 초기구간에서는 건조통의 건조작업이 잘 작동하여 다른 고장에 비해 상대적으로 낮은 정알람율을 갖게 되었다.

또한, 고장데이터가 실제데이터가 아니고 실내 환경에서 확보된 데이터인 점을 감안해야 한다. 실험실의 고장

데이터는 각 부대별 특징 및 계절 영향을 반영하지 못하는 데이터이므로 부대별로 학습된 이상탐지 모델의 성능이 실제 데이터를 테스트 셋으로 사용하는 것보다 더 낮은 수밖에 없음을 고려하면 Table 3의 결과는 뛰어난 성능의 이상탐지 모델을 각 부대별로 확보했다고 이야기할 수 있다.

Table 3. Performance of best model on each army unit

Army Unit	abnormal type	Threshold	Frequency	False Alarm(%)	True Alarm(%)
1	ab1 ¹⁾	2.38	5	0	95.8
	ab2 ²⁾			0	98.0
	ab3 ³⁾			0	90.6
2	ab1	2.06	5	0	98.8
	ab2			0	99.5
	ab3			0	98.8
3	ab1	2.19	15	0	92.1
	ab2			0	95.0
	ab3			0	81.0
4	ab1	1.54	10	0.7	92.3
	ab2			0.7	94.3
	ab3			0.7	75.9
5	ab1	1.89	20	0.7	92.1
	ab2			0.7	95.7
	ab3			0.7	91.1
6	ab1	1.82	10	0	93.6
	ab2			0	96.1
	ab3			0	77.7

Note) ab1: failure of dryer unit, ab2: failure of cam motor, ab3: heater off

5. 결론

본 연구는 부대에서 운용하고 있는 무기체계 장비의 실제 데이터를 수집하고 데이터에 대한 데이터 탐색 및 분석을 통해 부대별 이상을 탐지하는 최적의 모델을 구축하였다.

기존의 지도학습 기반의 고장과 정상을 분류하는 이상탐지 모델에서 벗어나, 비지도학습 기반의 이상탐지 모델 확보방안을 통해 고장이력 수집이 어려운 실제 설비를 대상으로 이상을 탐지하는 방안을 제안하였고 제안된 방법론을 통해 오일람을 줄이면서 빠르게 이상을 탐지하는 모델을 확보할 수 있었다.

또한, 데이터의 변수가 적고 작동의 on/off 구분이 없는 무기체계의 장비도 데이터 탐색을 통한 데이터 셋 구축과 여러 가지 통계적 기법과 머신러닝, 딥러닝 기법을 적용하여 장비의 상태를 기반으로 이상을 탐지하는 최적

의 모델을 확보하였다.

이러한 연구적 성과는 국내의 열악한 조건의 다른 무기체계에서도 데이터분석을 활용한 상태기반정비의 도입의 가능성을 확인할 수 있는 유의미한 연구라고 말할 수 있다.

본 연구는 천마 공기건조기의 온습도 데이터를 활용하여 이상탐지 모델을 구축하였다. 이미 운용되고 있는 장비를 대상으로 장비의 구조를 바꾸지 않고 시범적으로 쉽게 획득할 수 있는 변수 중 고장과 가장 연관성이 높은 온습도 변수를 수집 대상으로 삼았다. 하지만 온습도 변수의 경우 고장의 직접적인 원인을 대표하는 변수가 아니므로 실제 장비의 진동이나 기계적 결함 등을 대표하는 변수를 분석대상으로 삼아 이상탐지뿐만 아니라 고장의 원인까지 도출하는 연구를 진행하게 된다면 보다 효율적인 정비체계를 갖추는데 큰 도움이 될 것이다.

또한, 각 건조통별 센서 및 작동 on/off 센서가 설치되고 고장이력을 포함한 장기간의 데이터 수집이 이뤄진다면 설비의 수명예측 연구도 함께 이뤄질 수 있고 이러한 연구는 정비체계 효율화에 큰 도움이 될 것이다.

References

- [1] I. H. Song, "Reliability improvement method in weapon systems through field failure data analysis", *Journal of the Korea Academia-Industrial cooperation Society*, vol. 19, no. 12 pp. 110-117, 2018. DOI: <https://doi.org/10.5762/KAIS.2018.19.12.110>
- [2] S. H. Kim, Y. M. Chung, H. J. Koh, "A Case Study on the Effect of Maintenance Cost Reduction in Weapon-Systems with CBM", *Korea Association of Defense Industry Studies*, vol.17, no.1, June 2010.
- [3] B. Sun, S. Zeng, R. Kang, and M. Pecht, "Benefits and challenges of system prognostics", *IEEE Transactions on Reliability*, Vol.61, No.2, pp.323-335, 2012.
- [4] H. S. Jeong, "A Study on the Condition Based Maintenance Development for the Expansion of CBM Adoption in Weapon Systems", *Journal of the Korea Academia-Industrial cooperation Society*, vol.22, no.8, pp. 631-638, Aug. 2021. DOI: <https://doi.org/10.5762/KAIS.2021.22.8.631>
- [5] MND(Ministry of National Defense), Directive on defence force development, 2022.
- [6] MND(Ministry of National Defense), Directive on total life cycle management, 2022.
- [7] J. W. Ahn, S. W. Noh, T. H. Kim, "An Empirical Study on the Prediction of Future New Defense Technologies in Artificial Intelligence", *Journal of the*

Korea Academia-Industrial cooperation Society, vol. 21, no.9 pp. 458-465, 2020.

DOI: <https://doi.org/10.5762/KAIS.2020.21.9.458>

- [8] Y. G. K, M. J. Son, S. C. Noh, S. J. Kim, "A Study on Prevention of Condensation and Freezing in Radar Waveguide using Condition Based Maintenance", *Journal of the Korea Academia-Industrial cooperation Society*, vol.23, no.7, pp. 229-239, 2022.
DOI: <https://doi.org/10.5762/KAIS.2022.23.7.229>
- [9] H. He, E. A. Garcia, "Learning from Imbalanced Data", *IEEE Transaction on Knowledge and Data Engineering*, vol.21, no.9, Sep. 2009.
DOI: <https://doi.org/10.1109/TKDE.2008.239>
- [10] V. Chandola, A. Banerjee, "Anomaly Detection : A Survey", *ACM Computing Surveys*, vol.41, Issue 3, no.15, pp.1-58, Jul. 2009.
DOI: <https://doi.org/10.1145/1541880.1541882>
- [11] F. T. Liu, K. M. Ting, Z. H. Zhou, "Isolation Forest", *IEEE International Conference on Data Mining*, IEEE, Pisa, Italy, pp.413-422, Dec. 2008.
DOI: <https://doi.org/10.1109/ICDM.2008.17>
- [12] M. M. Breunig, H.-P. Kriegel, R. T. Ng, J. Sander, "LOF : Identifying Density Based Local Outliers," *Proceedings of the ACM SIGMOD Conference*, ACM, Dallas, Texas, New York, vol.29, issue 2, pp.93-104, June 2000.
DOI: <https://doi.org/10.1145/335191.335388>
- [13] C. Cortes, V. Vapnik, "Support-vector networks". *Machine Learning*, 20, no.3 , 273-297, 1995.
DOI: <https://doi.org/10.1023/A:1022627411411>
- [14] B. Scholkopf, J. C. Platt, J. Shawe-Taylor, A. J. Smola, and R. C. Williamson, "Estimating the Support of a High Dimensional Distribution", *Neural computation*, 13(7):1443-1471, 2001.
DOI: <https://doi.org/10.1162/089976601750264965>
- [15] S. M. Erfani, S. Rajasegarar, S. Karunasekera, C. Leckie, "High-dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning", *Pattern Recognition*, vol.58, pp.121-134, Oct. 2016.
DOI: <https://doi.org/10.1016/i.patcog.2016.03.028>
- [16] M. Imani and H. Ghassemian, "Principal component discriminant analysis for feature extraction and classification of hyper spectral images", *Iranian Conference on Intelligent Systems (ICIS)*, IEEE, Bam, Iran, Feb. 2014.
DOI: <https://doi.org/10.1109/IranianCIS.2014.6802535>
- [17] F. Harrou, F. Kadri, S. Chaabane, C. Tahon , Y. Sun "Improved principal component analysis for anomaly detection: Application to an emergency department", *Computers & Industrial Engineering*, vol.88, pp. 63~77, Oct. 2015.
DOI: <https://doi.org/10.1016/j.cie.2015.06.020>
- [18] P. Malhotra, A. Ramakrishnan, G. Anand, L. Vig, P. Agarwal, G. Shrof, LSTM-based Encoder-Decoder for Multi-sensor Anomaly Detection", *ICML 2016 Anomaly Detection Workshop*, New York, USA, 2016.

DOI: <https://doi.org/10.48550/arXiv.1607.00148>

- [19] M. Shardlow, An Analysis of Feature Selection Techniques.

김 요 진(Yojin Kim)

[정회원]



- 2008년 2월 : 서울대학교 기계항공공학부 (기계공학 학사)
- 2021년 9월 ~ 현재 : 고려대학교 산업경영공학부 (산업경영공학 석사)
- 2019년 7월 ~ 현재 : 한화시스템 ICT data 인텔리전스 과장

<관심분야>

이상탐지, 머신러닝

강 필 성(Pilsung Kang)

[정회원]



- 2003년 2월 : 서울대학교 산업공학과 (산업공학 학사)
- 2010년 2월 : 서울대학교 산업공학과 (산업공학 박사)
- 2015년 3월 ~ 현재 : 고려대학교 산업경영공학부 교수

<관심분야>

머신러닝, 딥러닝

염 동 원(Dong-Won Kim)

[정회원]



- 2015년 2월 : 동아대학교 전자공학 학사 (전자공학학사)
- 2015년 2월 ~ 현재 : 한화시스템 MRO사업팀 대리

<관심분야>

국방, 정보통신

김 건 우(Kun-Woo Kim)

[정회원]



- 2011년 2월 : 경북대학교 전자공학
학사 (전자공학학사)
- 2011년 2월 ~ 현재 : 한화시스템
MRO사업팀 과장

〈관심분야〉

국방, 정보통신