

# 딥러닝 알고리즘을 활용한 디스플레이 정전척 FAB 공정의 불량 탐지 및 원인 분석

김서연<sup>1</sup> · 이지윤<sup>1</sup> · 목충협<sup>1</sup> · 김상훈<sup>1</sup> · 문석호<sup>1</sup> · 경윤영<sup>2</sup> · 진영길<sup>2</sup> · 이우진<sup>2</sup> · 최종명<sup>2</sup> · 김성범<sup>1\*</sup>

<sup>1</sup>고려대학교 산업경영공학과 / <sup>2</sup>삼성 디스플레이

## Defect Detection and Cause Analysis on Display Electrostatic Chuck FAB Process Using Deep Learning Algorithms

Seo-Yeon Kim<sup>1</sup> · Jiyeon Lee<sup>1</sup> · Chunghyup Mok<sup>1</sup> · Sanghoon Kim<sup>1</sup> · Seokho Moon<sup>1</sup> · Yunyoung Kyeong<sup>2</sup>  
Younggil Jin<sup>2</sup> · Woojin Lee<sup>2</sup> · Jongmyong Choi<sup>2</sup> · Seoung Bum Kim<sup>1</sup>

<sup>1</sup>Department of Industrial and Management Engineering, Korea University / <sup>2</sup>Samsung Display

Deep learning algorithms have been used in various production processes to reduce product defect rates and improve process management efficiency. In this study, we propose a method for defect detection and cause analysis in the field of display electrostatic chuck fabrication (FAB) processes. The proposed method consists of three steps. First, we construct an autoencoder model for detecting defects and obtain reconstruction errors. Second, we construct the deep neural network model based on reconstruction errors obtained from the preceding autoencoder model to predict defects. Lastly, we applied locally interpretable model-agnostic explanations (LIME) and Shapley additive explanations (SHAP) to the deep neural network model for identifying the major causes of defects. The usefulness and applicability of the proposed method was demonstrated using the actual electrostatic chuck FAB data. Our proposed model outperformed all comparative models by achieving the highest average recall score. Furthermore, lag time and electrostatic chuck driving environment were identified as major causes of the defect.

**Keywords:** Deep Learning, Explainable AI, Electrostatic Chuck FAB Process, Defect Prediction, Anomaly Detection

### 1. 서론

최근 TV, 컴퓨터, 스마트폰 등의 전자기기가 발전함에 따라 전자기기의 필수 부품인 디스플레이 기술산업 또한 그 중요성이 날이 높아지고 있다. 디스플레이 생산 공정은 크게 초미세 공정을 통해 디스플레이 기판을 생산하는 fabrication(FAB) 공정, 만들어진 디스플레이 기판을 제품의 용도에 맞는 크기로

잘라내는 셀(cell) 공정, 전자기기에 디스플레이를 부착하는 모듈(module) 공정으로 구분되며, 각 공정은 여러 생산 설비들로 구성되어 있다. 상호 유기적으로 연계된 각 공정에서 디스플레이 기판의 물리적 손상 없이 생산 설비로 이송하기 위하여 정전기력(electrostatic, ESC)으로 디스플레이 기판을 탈부착하는 정전척 장비(ESC equipment)를 사용한다. 구체적인 정전척 방식의 예시는 <Figure 1>에서 확인할 수 있다. 예로 들어 설비

이 논문은 삼성디스플레이(주) 지원을 받아 수행된 연구 결과입니다.

This research was supported by the Brain Korea 21 FOUR, Ministry of Science and ICT (MSIT) in Korea under the ITRC support program (IITP-2020-0-01749) supervised by the Information & communications Technology Planning & Evaluation (IITP), the National Research Foundation of Korea grant funded by the MSIT (NRF-2019R1A4A1024732), and IITP grant funded by the MSIT (IITP-2021-0-00034).

\* 연락저자 : 김성범 교수, 02841 서울특별시 성북구 안암로 145, 고려대학교 산업경영공학과, Tel : 02-3290-3397, Fax : 02-929-5888,

E-mail : sbkim1@korea.ac.kr

2021년 11월 8일 접수; 2021년 12월 24일 수정본 접수; 2022년 2월 3일 게재 확정.

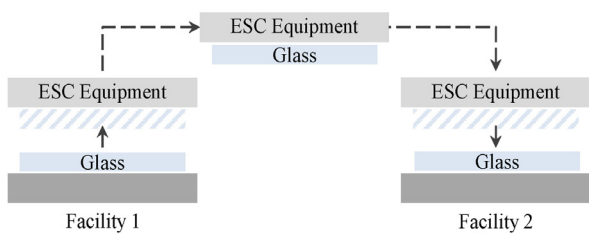


Figure 1. Overview of the Electrostatic Chuck Fabrication Process

1(facility 1)에서 설비 2(facility 2)로 디스플레이 기판(glass)을 이동할 때, 정전척 장치는 정전기력으로 설비 1의 디스플레이 기판을 장비에 부착한다. 이를 운반한 이후에는 정전기력을 감소시켜 디스플레이 기판이 정전척 장비로부터 탈착하여 설비 2에 도달하도록 작동한다.

정전척을 활용한 정확하고 신속한 이송은 디스플레이 제조 공정의 생산성과 직결되므로 정전척 장비의 고장을 미리 파악하고 선제적으로 대응하는 것이 중요하다(Sahno and Shevtshenko, 2014; Stojanovic *et al.*, 2016). 하지만, 기존에는 작업자가 정성적으로 판단하는 방식으로 정전척 장비의 고장을 파악하였으며, 이 경우 많은 시간과 비용이 요구되고 인력 가용 측면에서도 비효율적이다. 따라서, 생산 설비로부터 수집된 데이터를 활용하여 기기 고장을 사전에 탐지하고 고장 원인을 설명할 수 있는 인공지능 알고리즘의 개발은 시간과 비용의 절감 뿐만 아니라 고장에 대한 적절한 사후 조치를 시행할 수 있으므로 필수적이다(Jung and Lim, 2021).

이러한 수요에 따라 디스플레이 산업에서 인공지능을 활용한 연구들이 진행되어 왔다. Kang *et al.*(2009)은 liquid crystal display(LCD) 공정 과정에서 발생한 불량에 대한 너비, 크기, 이심률 등의 데이터에 서포트 벡터 머신(support vector machine) 알고리즘을 적용하여 다중 불량 클래스를 분류하였다. Ryu *et al.*(2017)은 디스플레이 회로 이미지 데이터 중 불량이 발생한 경우를 탐지하고자 넓은 영역의 결함 이미지를 효과적으로 학습할 수 있는 합성곱 신경망(convolutional neural networks) 구조를 활용하여 다중 클래스의 불량 회로 이미지를 분류하는 알고리즘을 제안하였다. 상기 방법론들은 디스플레이 기판의 정상과 불량을 높은 성능으로 분류하지만 충분한 불량 데이터를 필요로 한다. 하지만 실제 생산 설비에서 수집되는 데이터에는 불량 관측치의 수가 현저히 적으므로 해당 방법론들을 활용하는데 어려움이 있다.

따라서, 불량 데이터 수집이 어려운 경우 정상 데이터만을 학습하여 기기 고장을 예측하는 이상치 탐지 연구들의 필요성이 대두되고 있다. 특히 최근에는 오토인코더(autoencoder) 모델 등의 딥러닝 알고리즘으로 불량을 예측하는 방법론들이 우수한 성능을 보이고 있다(Chalopathy and Chawla, 2019; Pang *et al.*, 2021). Ku(2019)는 디스플레이 기판들 사이에 일정한 간격을 유지해주는 photo spacer(PS) 장치의 정상 이미지 데이터만을 활용하여 concentrated auto encoder(CAE) 모델을 학습하고, 이에 대한 재구축 오차를 기반으로 불량을 예측하여 높은 이

상치 탐지 성능을 보였다. 또한, Park(2019)은 디스플레이 생산 기기의 센서로부터 수집된 시계열 데이터에 순환 신경망(recurrent neural networks) 모델을 적용하여 기기의 이상 징후를 사전에 포착하였다. 이상치 탐지 알고리즘을 활용한 기존 연구들은 정상 데이터만으로 불량 탐지가 가능하다는 이점이 있지만, 불량에 대한 해석이 제공되지 않는다는 한계점이 있다(Ruff *et al.*, 2021).

따라서, 불량 발생이 희박한 실제 산업 현장에서 정상 데이터만을 활용하여 기기 결함을 높은 성능으로 예측하고, 예측 결과에 대한 해석을 제공할 수 있는 인공지능 알고리즘이 필요하다. 본 연구에서는 디스플레이 주요 생산 공정인 FAB 공정에서의 정전척 장비 불량을 오토인코더 기반 이상치 탐지 알고리즘을 활용하여 사전에 탐지하고, 오토인코더 구조에 설명 가능한 인공지능 방법론을 적용할 수 있는 reconstruction error prediction model(REPM)을 제안하여 불량에 대한 원인 해석을 도출하고자 한다.

본 논문의 기여점은 다음과 같다. (1) 디스플레이 산업에서 정전척 설비 불량 여부 탐지 및 설명 가능한 인공지능을 적용한 첫 사례이다. (2) FAB 공정에서의 정전척 장비 상태 데이터를 활용하여 불량 여부를 탐지하고 원인을 파악할 수 있는 프레임워크를 제안하였다. (3) 제안 방법론 중 오토인코더 기반 이상치 탐지 알고리즘에 설명 가능한 인공지능을 적용하기 위한 REPM 구조를 새롭게 추가하여 불량에 대한 원인 인자를 탐지하고 적용 가능성을 입증하였다.

본 논문의 구성은 다음과 같다. 제2장에서는 제안 방법론에서 활용한 관련 연구에 대하여 자세히 설명하며, 제3장에서는 제안 알고리즘을 소개한다. 제4장에서는 실험 방법 및 실험 결과를 비교하고, 제5장에서는 본 연구의 결론과 기대 효과에 대하여 다루도록 한다.

## 2. 배경 방법론

### 2.1 Autoencoder

오토인코더 모델은 불량 탐지 분야에서 주로 활용되는 딥러닝 알고리즘 중 하나로, 입력 데이터와 동일한 출력 데이터를 얻기 위하여 입력 데이터의 특성을 요약하여 저차원의 잠재

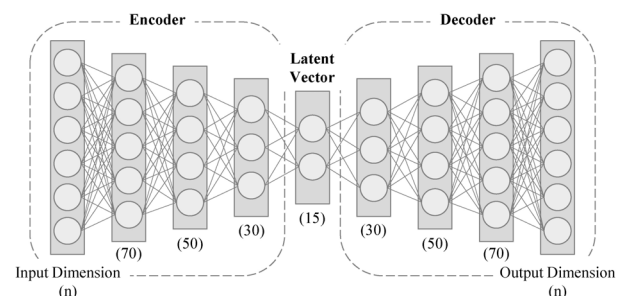


Figure 2. Overview of the Autoencoder Model

벡터(latent vector)로 축소하는 인코더(encoder)와 요약된 잠재 벡터를 다시 복원하는 디코더(decoder)의 형태로 구성된다 (Rumelhart *et al.*, 1985). 오토인코더 모델의 구체적인 구조는 <Figure 2>와 같다.

오토인코더 모델은 재구성된 데이터와 원본 데이터와의 비교를 통하여 두 데이터 사이의 차이가 작아지도록 학습한다. 원본 데이터( $x_i$ )와 재구성된 데이터( $\hat{x}_i$ ) 사이의 차이는 재구성 오차(reconstruction error)로 정의되며 식 (1)을 통해 구할 수 있다. 해당 수식은 벡터의 크기를 구하기 위한 유클리드 노름 (Euclidean norm)으로 구성되어 있다.

$$Reconstruction\ Error = \sum_{i=1}^n \|x_i - \hat{x}_i\|_2^2 \quad (1)$$

오토인코더는 데이터를 축소하고 복원하는 과정에서 데이터의 패턴을 학습한다. 이상치 탐지 알고리즘으로 활용하기 위하여 학습 단계에서는 정상 데이터만을 사용하기 때문에 다양한 정상 패턴들에 대해 학습하게 된다. 이로 인하여, 오토인코더 모델은 학습된 패턴을 기반으로 정상 데이터를 효과적으로 재구성하지만, 이전에 입력되지 않았던 불량 데이터에 대해서는 정확하게 복원하지 못하므로 원본 데이터와 재구성 데이터 사이에 큰 재구성 오차가 산출된다. 따라서, 오토인코더 모델에서는 재구성 오차가 높은 관측치를 이상치로 간주한다 (An and Cho, 2015).

## 2.2 설명 가능한 인공지능 알고리즘

설명 가능한 인공지능 알고리즘(explainable AI, XAI)은 분석에 활용한 딥러닝 모델의 출력 결과를 잘 이해하고 설명하기 위한 방법이다(Arrieta *et al.*, 2020). 이 알고리즘을 활용하면 분석한 모델에 영향을 미친 인자들을 파악한 후, 각 요인에 대한 결과를 정량적으로 산출하여 예측 결과에 대한 직관적인 설명을 제공할 수 있다.

설명 가능한 인공지능 알고리즘은 모델에 적용되는 방식에 따라, 특정 모델에만 적용 가능한 방법론(model-specific)과 다양한 모델에 적용 가능한 방법론(model-agnostic)으로 구분된다. 본 연구에서는 다양한 모델에 적용 가능한 방법론 중 정량적인 설명 결과를 각 관측치 별로 제공해주는 LIME(locally interpretable model-agnostic explanations)과 SHAP(Shapley additive explanations) 알고리즘을 활용하였다.

복잡한 딥러닝 방법론의 결과 해석을 위하여 LIME과 SHAP 알고리즘은 원인 인자를 쉽게 파악할 수 있는 대리 모델(surrogate model)을 사용한다(Adadi and Berrada, 2018). 이 때 대리 모델로는 주요 변수에 대한 해석을 제공할 수 있는 모델들이 주로 활용되며, 대표적으로 선형회귀분석과 의사결정나무 등이 있다. 대리 모델은 딥러닝 모델의 예측 결과에 대한 해석을 목표로 하므로, 실제값이 아닌 예측값에 대해 학습이 이루어진다. 최종적으로 딥러닝 모델의 예측값은 대리 모델로부

터 도출된 해석 결과를 응용하여 설명되며, 결과에 대한 원인 인자 분석이 가능해진다.

### 2.2.1 LIME

LIME 알고리즘은 관측치 단위에서 대리 모델을 구축하고 변수 별 해석을 제공하는 것을 주목적으로 두고 있다 (Ribeiro *et al.*, 2016). <Figure 3>에서 보여주듯 설명 대상 모델(original model)의 관측치들 중 설명이 필요한 대상 관측치(instance being explained,  $x$ )를 우선적으로 선정하고, 설명 대상 관측치 이외에 추가적인 생성 데이터(perturbed data,  $z$ )를 통해 해당 관측치의 예측값을 도출할 수 있는 대리 모델을 구축한다.

LIME 알고리즘은 입력값에 미비한 변화(perturbation)가 주어졌을 때 모델의 예측 결과 값이 크게 변한다면 해당 변수의 중요도가 높다는 가설을 지닌다. 따라서, 해석하고자 하는 대상 관측치를 기준으로 미비한 변화를 부여하여 추가적인 생성 데이터( $z$ )를 구축한다. 이때, 생성 데이터( $z$ )와 설명 대상 관측치( $x$ ) 간의 유사도를 커널 함수로 계산하여, 생성 데이터가 설명 대상 데이터와 최대한 유사하도록 설정한다. 커널 함수의 종류에는 지수 커널 함수(exponential kernel,  $\pi_x(z)$ ), 가우시안 커널 함수(Gaussian kernel) 등이 있으며, 대체로 정형 데이터에 적용하는 경우 식 (2)의 지수 커널 함수를 활용한다. 또한, 기존 데이터와 생성 데이터의 차원이 고차원인 경우 대리 모델로 설명이 쉽지 않을 수 있으므로, 차원이 축소된 생성 데이터( $z'$ )를 대리 모델에 적용한다.

$$\pi_x(z) = \exp\left(-\frac{D(x,z)^2}{\sigma^2}\right) \quad (2)$$

대리 모델을 학습시킬 때 사용하는 손실 함수( $L(f, g, \pi_x)$ )는 식 (3)에서 확인할 수 있듯이 기존 예측 모델에 생성 데이터를 입력시켜 얻은 결과값( $f(z)$ )과, 대리 모델에 차원이 축소된 생성 데이터( $z'$ )를 입력하여 얻은 예측값( $g(z')$ )의 차이로 구성되어 있다. 또한, 앞서 생성 데이터 구축에 활용한 지수 커널 함수( $\pi_x$ )를 손실함수의 계산된 값에 대한 가중치로 사용하여 결과값과 예측값 사이의 차이를 더욱 증대하였다. 즉, 기존 예측 모델과 새롭게 구축하려는 대리 모델이 서로 유사한 예측

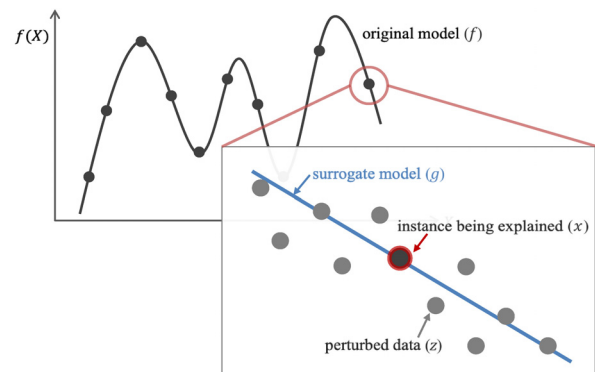


Figure 3. Overview of the LIME Algorithm

값을 갖도록 학습된다. 마지막으로, 식 (4)의 최종 손실 함수 ( $\xi(x)$ )는 대리 모델의 복잡도( $\Omega(g)$ )를 결합한 형태로 구성되며 기존 예측 모델과 유사하면서도 설명 복잡도가 낮은 대리 모델을 학습시킨다.

$$L(f, g, \pi_x) = \sum_{z, z' \in Z} \pi_x(z) (f(z) - g(z'))^2 \quad (3)$$

$$\xi(x) = \operatorname{argmin}_{g \in G} L(f, g, \pi_x) + \Omega(g) \quad (4)$$

### 2.2.2 SHAP

SHAP 알고리즘은 LIME 알고리즘과는 달리 전체 관측치를 활용하여 대리 모델을 구성하고 그 결과를 해석하며, 각 변수들이 각각의 예측값을 도출할 때에 기여하는 정도를 평가하는 방법이다(Lundberg and Lee, 2017). 이에 대한 그림은 <Figure 4>에서 확인할 수 있다. 설명 대상 모델과 가장 유사한 예측값을 도출할 수 있도록 여러 선형 함수가 결합된 형태의 대리 모델을 구축하고, 이를 해석하여 설명 결과를 제공한다.

SHAP 알고리즘에서는 특정 변수를 제거하고 모델을 재구축했을 때 예측 성능이 크게 변화한다면 해당 변수의 중요도가 높음을 가설로 두고 있다. 이 때 예측 성능의 변화 정도는 샤플리 값(Shapley value,  $\phi_j$ )으로 정의되며, 구체적인 수식은 식 (5)와 같다.

$$\phi_j = \sum_{S \subseteq F \setminus \{j\}} \frac{|S|!(|F|-|S|-1)!}{|F|!} (f_{S \cup \{j\}}(x_{S \cup \{j\}}) - f_S(x_S)) \quad (5)$$

식 (5)에서  $j$ 는 설명 대상 변수의 번호를 의미하며  $S$ 는 변수들의 전체 집합에서 설명 대상 변수를 제외한 집합으로 생성 가능한 모든 부분 집합,  $F$ 는 변수들의 전체 집합이다. 샤플리 값을 계산하는 수식은 전체 변수들 중 설명 대상 변수가 선택될 확률( $(|S|!(|F|-|S|-1)!)/(|F|!)$ )과, 설명 대상 변수가 포함되어 있을 때의 성능( $f_{S \cup \{j\}}(x_{S \cup \{j\}})$ )과 포함되어 있지 않을 때의 성능( $f_S(x_S)$ ) 차이로 구성되어 있다. 즉, 특정 변수에 대한 샤플리 값이 클수록 예측값을 도출하는 데에 많은 기여를 하는 것으로 해석되며, 이는 설명 대상이 되는 관측치에게 중요한 변수임을 의미한다.

SHAP 알고리즘에서는 식 (6)과 같이 샤플리 값의 선형 결합

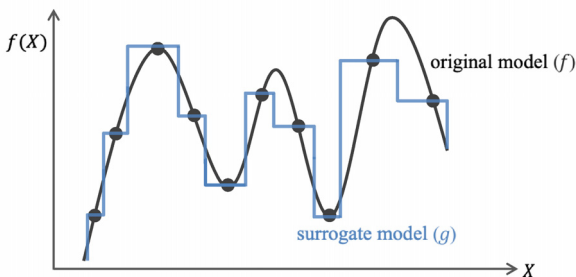


Figure 4. Overview of the SHAP Algorithm

으로 대리 모델( $g(z')$ )을 설정한다. 여기서  $z'_j$ 는  $j$ 번째 설명 대상 변수의 선정 여부를 의미하는 더미 변수로 0 또는 1의 값을 갖는다. 학습된 대리 모델을 통하여 <Figure 4>에서와 같이 각 관측치 별로 실제 예측값( $f(x)$ )과 샤플리 값들의 선형 결합( $g(z')$ )이 유사한 값을 갖게 된다. 따라서, 샤플리 값의 선형 결합인 대리 모델을 해석함으로써 해석을 필요로 하는 예측 모델( $f$ )에 대한 설명을 얻을 수 있다.

$$g(z') = \phi_0 + \sum_{j=1}^M \phi_j z'_j \quad (6)$$

$$\text{where } z'_j \in \{0, 1\} \quad \phi_j \in R$$

SHAP 알고리즘은 모든 관측치 별로 각 변수에 대한 샤플리 값을 계산해야 하기 때문에 많은 연산량을 필요로 한다. 이에 따라 더 효율적인 계산을 위하여 대리 모델의 형태를 고정한 후 연산하는 LinearSHAP, TreeSHAP, KernelSHAP, DeepSHAP과 같은 여러 변형된 알고리즘들이 제안되었다(Lundberg and Lee, 2017). 본 연구에서는 딥러닝 모델에 적용하기 적합한 것으로 알려진 DeepSHAP을 활용하여 해석을 진행하였다. DeepSHAP은 설명 가능한 인공지능의 한 종류인 Deep Learning Important Features(DeepLIFT)에서 연산 과정에 활용한 기울기(gradient)로 샤플리 값을 계산하는 방식이다(Shrikumar et al., 2017).

앞서 대표적인 설명 가능한 인공지능 알고리즘으로 LIME과 SHAP에 대해 소개하였다. 설명 가능한 인공지능 방법론들은 원인 인자 도출을 위해 관측치 단위의 해석을 제공한다는 특징이 있지만, 예측 모델의 출력값에 대해 설명을 제시한다는 점에서 오토인코더 모델에 적용하는 데에 어려움이 있다. 오토인코더 모델의 출력값은 일반적인 예측값이 아닌 재구축된 데이터이기 때문에 설명 가능한 인공지능을 적용할 경우 불량 예측에 영향을 미친 요인이 아니라, 불량 데이터를 재구축하는 데에 기여한 인자들을 설명한다는 한계점이 있다. 따라서, 오토인코더 모델에 설명 가능한 인공지능을 적용하기 위해서는 불량 결과를 예측하는 추가적인 모델이 필요하다.

Antwarg et al.(2019)는 오토인코더 모델의 결과 중 재구축 오차가 높은 변수들에 KernelSHAP 알고리즘을 적용한 후, 그 연산 결과를 샤플리 값의 부호에 따라서 정상 데이터가 되는 요인(offsetting features)과 불량 데이터가 되는 요인(contributing features)으로 나누어 해석하는 방법을 제안하였다. 하지만, 해당 연구는 오토인코더 모델에 직접 SHAP 방법론을 적용하였다는 점에서 여전히 데이터를 재구축하는 데에 기여한 인자들이 도출되는 문제점을 해결하지 못하였다. 더 나아가, 재구축 오차가 높은 특정 변수들에 대해서만 샤플리 값 연산을 진행하며, 각 변수에 대하여 반복적으로 연산을 수행해야 하기 때문에 계산 복잡도가 높다는 단점이 있다. 따라서 본 연구에서는 불량 예측이 가능한 모델을 통하여 오토인코더 기반 이상치 탐지 알고리즘에 설명 가능한 인공지능을 적용하고, 전체 관측치 및 변수에 대하여 해석하는 방안을 제안하고자 한다.

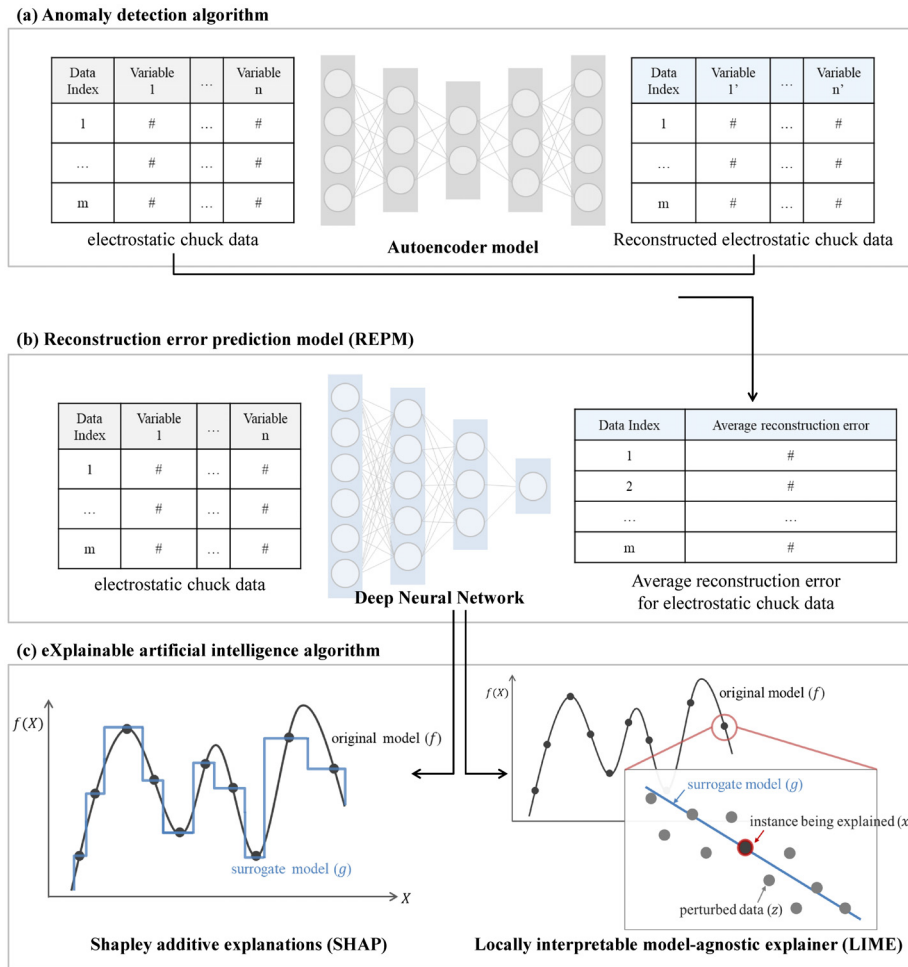


Figure 5. Overview of the Proposed Chuck Defect Prediction and Explainable Algorithm

### 3. 제안 방법론

본 연구에서는 (a) 이상치 탐지 알고리즘, (b) 불량 원인 파악을 위한 추가 제안 모델(REPM), (c) 설명 가능한 인공지능 알고리즘의 크게 세 가지 단계를 통하여 불량을 예측하고 원인을 규명하는 방법론을 제안한다. 구체적인 단계 별 구성은 <Figure 5>와 같다.

설명 가능한 인공지능 알고리즘을 적용하여 불량 원인을 파악할 수 있는 추가 모델을 구축하기 위해서는 불량 여부에 대한 예측이 가능해야 한다. 이상치 탐지 알고리즘에서 불량 예측의 기준으로 재구축 오차를 사용하였다. 따라서, 본 논문에서 제안하는 REPM은 이상치 탐지 알고리즘으로부터 도출된 재구축 절대오차의 평균값(average reconstruction error)을 예측하는 심층 신경망 회귀 모델로 구축하였다(<Figure 5(b)>). 예측값을 계산하기 위하여 식 7과 같이 오토인코더 모델의 입력 데이터와 재구축된 데이터 간의 차이를 활용하였으며, 관측치 별로 평균을 계산하였다. 해당 수식은  $m$ 개의 관측치와  $n$ 개의 변수가 주어졌을 때 한 관측치( $j$ )에 대한 각 변수( $i$ )를 기준으로 계산한다. 한 관측치를 기준으로 변수 별 벡터 간 거리

의 차이인 L1 노름(L1 norm)을 계산한 후, 연산 결과를 요약하여 평균을 내는 방식으로 구성되어 있다.

$$\text{average reconstruction error}_j = \frac{1}{n} \sum_{i=1}^n \|x_{ji} - \hat{x}_{ji}\| \quad (7)$$

$$\text{where } j \in \{1, \dots, m\}$$

REPM를 통하여 불량 관측치들은 더 높은 재구축 절대오차 평균값으로 예측된다. 즉, 학습된 REPM에 설명 가능한 인공지능 방법론을 적용하면 재구축 오차가 커지는데 기여한 변수들을 확인할 수 있으며, 이는 각 관측치가 불량으로 검출되는데 영향을 미친 주요한 원인으로 해석이 가능하다. REPM에서 활용한 심층 신경망 모델은 총 네 개의 은닉층으로 이루어져 있으며, 각각 65개, 50개, 35개, 20개의 노드를 갖도록 설정하였다.

본 연구에서 추가적으로 제안하는 REPM은 오토인코더 기반 이상치 탐지 알고리즘에 설명 가능한 인공지능을 적용하기 위한 구조로서 우수한 불량 탐지 성능을 유지함과 동시에 다양한 설명 가능한 인공지능 방법론을 적용할 수 있다. 또한, 불량 발생 원인에 대한 정량적인 지표 제공을 제공하기 때문에, 현장 엔지니어의 의사결정에도 기여할 수 있다.

## 4. 실험 방법 및 결과

### 4.1 데이터 수집 및 전처리

본 연구에서 활용한 데이터는 국내 디스플레이 기업 FAB 공정의 정전척 설비 데이터로 2020년 5월부터 2021년 2월까지의 기간 동안 총 371,124건 수집되었으며, 그 중 불량 데이터는 199건 포함되어 있다. 변수는 총 47개로 공정 시간 및 생산 경로, 전압 인가 전원의 종류, 핀의 위치, 설비 내 압력 정보, 정렬 정도, 적용 면적, 정전기의 세기 등의 정보로 이루어져 있다. 이와 같이 수집된 데이터를 훈련용 데이터, 검증용 데이터, 평가용 데이터로 분할하여 분석을 진행하였다. 훈련용 데이터와 검증용 데이터는 정상 데이터로만 구성하였으며, 모든 불량 데이터는 평가용 데이터에 포함되도록 설정하였다. 정상 데이터는 데이터의 수집 시점에 따라 64%는 훈련용 데이터, 16%는 검증용 데이터, 20%는 평가용 데이터로 분할하였다. 분할 후 데이터의 수는 훈련용 데이터 총 240,944건, 검증용 데이터 총 55,597건, 평가용 데이터 총 74,384건이다.

우선, 모델 구축 이전에 훈련용 데이터에 포함된 이상치 데이터를 제거하는 과정을 진행하였다. 이상치 데이터는 정상 데이터 중 기기의 결함 등으로 인해 잘못 수집된 데이터로, 학습에 악영향을 끼칠 수 있기 때문에 사전에 제거하여 명확한 정상 데이터만으로 모델을 학습시키고자 하였다. 반복 실험을 통하여 이상치 제거 여부가 불량 탐지 알고리즘의 성능에 유의미한 영향을 미침을 확인하였다. 이상치는 데이터의 분포에 기반하여 interquartile range(IQR)를 기준으로 제거하였다 (Hubert & Van der Veeken, 2008). IQR이란 제3 사분위수와 제1 사분위수의 차이를 의미하는 것으로서, IQR에 특정 가중치를 곱한 값을 각각 제1 사분위수에서 빼고 제3 사분위수에 더하여 계산된 범위를 정상 데이터의 범위로 인정하는 방식이다. 이 범위 밖에 존재하는 경우를 이상치로 판단하였다. 본 연구에서 활용한 연속형 변수들은 대칭 형태(symmetric), 왼쪽으로 늘어진 형태(left-skewed), 또는 오른쪽으로 늘어진 형태(right-skewed)의 세 가지로 나누어 볼 수 있었다. 따라서, 가중치 값을 대칭 형태의 경우 2, 왼쪽으로 늘어진 형태의 경우 5, 오른쪽으로 늘어진 형태의 경우 70으로 정의하여 이상치를 제거하였다. 이 외에도 범주형 변수를 원-핫 인코딩(one-hot encoding)으로 변환하였고, 전체 데이터에 표준 스케일러(standard scaler)를 적용하여 분석을 위한 전처리 과정을 수행하였다.

### 4.2 평가 지표

본 연구에서는 이상치 탐지 알고리즘의 성능을 평가하기 위하여 재현율(recall)과 geometric mean(GM) 점수, 그리고 area under the receiver operating characteristics(AUROC) 점수를 사용하였다. 재현율은 실제 불량 데이터 중 불량으로 올바르게 예측한 데이터의 비율을 의미하며, 구체적인 수식은 식 (8)로 표현된다.

$$recall = \frac{TP}{TP+FN} \quad (8)$$

$$specificity = \frac{TN}{TN+FP} \quad (9)$$

$$geometric\ mean(GM) = \sqrt{recall \times specificity} \quad (10)$$

식 (8)과 식 (9)에서 TP(true positive)는 실제 불량 데이터 중 불량으로 예측한 개수, FP(false positive)는 실제 정상 데이터 중 불량으로 잘못 맞춘 개수에 해당된다. 또한, FN(false negative)은 실제 불량 데이터 중 정상으로 잘못 예측한 개수, TN(true negative)은 실제 정상 데이터 중 정상으로 올바르게 맞춘 개수를 의미한다. 식 (9)의 특이도(specificity)는 실제 정상 데이터 중 정상으로 올바르게 예측한 데이터의 비율이다. GM 점수는 재현율과 특이도의 기하 평균이며, 식 (10)과 같이 계산된다. 재현율과 GM 점수는 모델이 예측한 불량 확률이 주어진 임계값을 초과했는지에 기반해서 계산한 값이다. 이에 반해, AUROC 점수는 모델이 예측한 불량 확률에 대하여 모든 임계값에 대한 분류 성능을 요약하여 하나의 지표로 표현한 것이다. 이는 임계값 별로 불량을 불량으로 올바르게 예측한 비율과 정상 중 불량으로 잘못 예측한 비율을 그래프 형태로 표현한 후 곡선 아래의 넓이로 계산한다.

### 4.3 실험 결과

오토인코더 기반 모델에는 다양한 알고리즘들이 존재한다. 본 연구에서는 FAB 공정 정전척 설비 데이터에 대한 이상치 탐지에 적절한 알고리즘을 선정하기 위하여 오토인코더(autoencoder), 변이형 오토인코더(variational autoencoder), 그리고 적대적 오토인코더(adversarial autoencoder) 모델의 예측 성능을 비교하였다. 먼저, 변이형 오토인코더는 데이터에 대한 잠재 벡터를 계산할 때 정규 분포와 유사해지도록 제약식을 추가하는 방법론이다(Kingma and Welling, 2013). 입력 데이터로부터 축소된 데이터에서 평균과 표준편차를 계산한 후, 정규 분포로부터 임의로 추출된 값을 이용하여 잠재 벡터를 생성한다. 이 후 디코더 단계에서는 정규분포로 근사시킨 잠재 벡터를 활용하여 재구성 데이터로 복원한다. 다음으로, 적대적 오토인코더는 잠재 벡터와 임의의 정규분포로부터 얻은 샘플 데이터를 구별하는 식별자를 추가적으로 학습하는 알고리즘이다(Makhzani et al., 2015). 이는 기존의 오토인코더 구조를 생성 모델의 한 형태로 변형한 것으로, 식별자를 활용하여 입력 데이터를 효과적으로 요약한 잠재 벡터를 구축하고자 제안되었다.

비교 실험으로 활용한 모델들의 구조는 격자 탐색(grid search)을 통해 가장 좋은 성능을 보인 구조로 설정하였다. 우선, 오토인코더 모델의 인코더와 디코더 부분 모두 3개의 은닉층으로 구성된 심층 신경망으로 구축하였다. 인코더 부분의 경우 각각 70개, 50개, 30개의 노드를 가진 3개의 은닉층으로 구성되어 있으며, 최종 15차원의 잠재 벡터로 요약하였다. 디

코더 부분은 각각 30개, 50개, 70개의 노드를 가진 3개의 은닉층으로 구성하였으며, 입력된 데이터와 같은 차원의 재구축 데이터를 출력하도록 설정하였다. 다음으로, 변이형 오토인코더의 인코더 부분은 30개의 노드를 가진 은닉층 1개를 포함한 인공 신경망 구조로 구성되어 있으며, 15차원의 잠재 벡터로 요약하였다. 디코더 부분 역시 30개의 노드를 가진 은닉층 1개의 인공 신경망으로 구축하였다. 마지막으로, 적대적 오토인코더 모델의 인코더 및 디코더 부분은 모두 2개의 은닉층으로

구성된 심층 신경망에 해당한다. 인코더 부분으로 50개의 노드를 가진 2개의 은닉층을 설정한 후 15차원의 잠재 벡터로 요약하였으며, 디코더 부분은 50개의 노드를 가진 2개의 은닉층으로 구축하였다. 적대적 오토인코더의 식별자 부분은 15차원의 잠재 벡터를 구분하는 심층 신경망 구조로, 8개의 노드를 가진 2개의 은닉층으로 구성되어 있다.

오토인코더 모델과 변이형 오토인코더, 적대적 오토인코더 모델의 비교 실험 결과는 10회 반복 실험을 통하여 평균 재현율, GM 평균 점수, AUROC 평균 점수로 각각 <Table 1>, <Table 2>, <Table 3>에 정리하였으며, 표준편차는 괄호 안에 표기하였다. 특히, 구체적인 비교를 위하여 재현율과 GM 점수는 임계값 별로 결과를 산출하였다. 이 때, 임계값은 현업의 의견에 따라 0.91부터 0.99까지 변화시키며 반복 실험을 진행하였다.

세 가지 모델에 대한 비교 실험 결과를 살펴볼 때 오토인코더 모델이 모든 성능 지표에서 가장 우수한 성능을 보였다. 특히, 불량에 대한 재현율 기준으로는 임계값 전체 평균 0.9524로 가장 높은 값을 확인하였다. 이는 오토인코더 모델이 FAB 공정 정전척 설비 데이터의 불량을 가장 잘 검출함을 의미한다. 더 나아가, GM 점수는 전체 평균 0.8400, AUROC 점수는 0.9629로 우수한 성능을 보였다. 이에 따라 본 연구에서 제안하는 오토인코더 기반 이상치 탐지 알고리즘은 FAB 공정 정전척 설비 데이터의 불량 관측치 뿐만 아니라 정상 데이터도 효과적으로 분류하는 최적 모델임을 알 수 있다.

다음으로, 최적 모델인 오토인코더 모델의 재구축 오차를 기반으로 REPM을 구축하였으며, REPM에 설명 가능한 인공 지능 알고리즘인 LIME과 SHAP을 각각 적용하여 그 결과를 확인하였다. LIME과 SHAP 알고리즘은 각 관측치에 대한 변수 별 중요도 점수(feature importance measure)를 도출한다. 요약된 변수 별 중요도를 산출하기 위해서는 일반적으로 변수 중요도 평균점수(average feature importance measure)를 활용하며, 이는 식 (11)과 같이 관측치 별 변수 중요도 점수에 절댓값을 취한 후 평균을 계산하는 방식이다(Lundberg and Lee, 2017). 이 때  $k$ 는 각 변수 번호,  $l$ 는 각 관측치 번호,  $n$ 은 전체 관측치의 개수이다.

$$\begin{aligned} \text{average feature importance measure}_k & \quad (11) \\ & = \sum_l |feature\ importance\ measure_{l,k}|/n \end{aligned}$$

실제 불량 데이터 중 불량으로 올바르게 예측된 관측치들에 대한 변수 중요도 평균점수 결과는 <Figure 6>으로 정리되며, 그래프(a)는 LIME 알고리즘, 그래프(b)는 SHAP 알고리즘의 변수 중요도 결과이다. 두 그래프의 x축은 변수 별 중요도 평균점수, y축은 상위 20개 변수들의 이름이다. 상위 20개 변수들 중 LIME과 SHAP 알고리즘에서 공통적으로 주요하게 뽑힌 변수들은 \*로 표시하였다. 또한, <Figure 6>의 상위 10개 변수들에 대한 10개 관측치 별 중요도 점수를 <Appendix>의 <Table 4>와 <Table 5>에 정리하였다.

**Table 1.** Average Recall for Autoencoder, Variational Autoencoder, and Adversarial Autoencoder

Threshold	Autoencoder	Variational Autoencoder	Adversarial Autoencoder
0.91	<b>0.9764 (0.0048)</b>	<b>0.8090 (0.0000)</b>	<b>0.9040 (0.0208)</b>
0.92	0.9754 (0.0037)	0.8050 (0.0021)	0.8854 (0.0183)
0.93	0.9739 (0.0046)	0.7970 (0.0035)	0.8678 (0.0168)
0.94	0.9704 (0.0055)	0.7849 (0.0021)	0.8497 (0.0216)
0.95	0.9623 (0.0093)	0.7789 (0.0000)	0.8385 (0.0223)
0.96	0.9548 (0.0063)	0.7487 (0.0000)	0.7985 (0.0090)
0.97	0.9462 (0.0086)	0.7171 (0.0041)	0.7653 (0.0067)
0.98	0.9276 (0.0076)	0.6683 (0.0000)	0.7095 (0.0094)
0.99	0.8844 (0.0085)	0.6015 (0.0024)	0.6543 (0.0116)
Total Average	<b>0.9524 (0.0049)</b>	0.7456 (0.0004)	0.8081 (0.0103)

**Table 2.** Average Geometric Mean Score for Autoencoder, Variational Autoencoder, and Adversarial Autoencoder

Threshold	Autoencoder	Variational Autoencoder	Adversarial Autoencoder
0.91	0.7433 (0.0417)	0.8600 (0.0000)	0.8099 (0.0197)
0.92	0.7671 (0.0459)	<b>0.8607 (0.0012)</b>	0.8236 (0.0166)
0.93	0.7924 (0.0484)	0.8599 (0.0019)	0.8356 (0.0119)
0.94	0.8179 (0.0484)	0.8574 (0.0012)	0.8462 (0.0113)
0.95	0.8421 (0.0469)	0.8590 (0.0000)	<b>0.8580 (0.0099)</b>
0.96	0.8681 (0.0411)	0.8465 (0.0000)	0.8533 (0.0044)
0.97	0.8948 (0.0283)	0.8324 (0.0024)	0.8484 (0.0052)
0.98	0.9156 (0.0136)	0.8075 (0.0000)	0.8256 (0.0059)
0.99	<b>0.9183 (0.0044)</b>	0.7697 (0.0016)	0.7992 (0.0074)
Total Average	<b>0.8400 (0.0347)</b>	0.8392 (0.0002)	0.8333 (0.0072)

**Table 3.** Average AUROC score for Autoencoder, Variational Autoencoder, and Adversarial Autoencoder

	Autoencoder	Variational Autoencoder	Adversarial Autoencoder
AUROC	<b>0.9629 (0.0035)</b>	0.9430 (0.0011)	0.9270 (0.0081)

LIME과 SHAP 알고리즘의 결과를 요약하기 위하여 각각 10회 반복 실험을 수행하여 결과를 도출하였다. 이 때, SHAP 알고리즘은 예측 모델 REPM의 결과가 고정적이기 때문에 결과의 변동이 없는 반면, LIME 알고리즘의 경우 시행할 때마다 미비하게 상이한 변수 별 중요도 점수를 확인할 수 있었다. 이는 LIME 알고리즘 내에서 설명 대상 관측치와 유사하도록 생성된 데이터(perturbed data,  $z$ )가 확률에 기반한 무작위성을 띄

기 때문이다. 이에 따라 LIME 알고리즘은 반복 실험을 통하여 도출된 결과값들의 상위 20개 변수들에 대한 변동 정도를 <Figure 7>에 박스 플랏(box plot)의 형태로 요약하였다. 10회 반복 실험을 진행하더라도 도출되는 변수 중요도 점수에는 큰 차이가 없음을 확인할 수 있다.

변수에 대한 보안 문제로 인해 자세한 설명은 포함하지 못했으나 LIME과 SHAP 알고리즘 모두에서 중요했던 변수들은

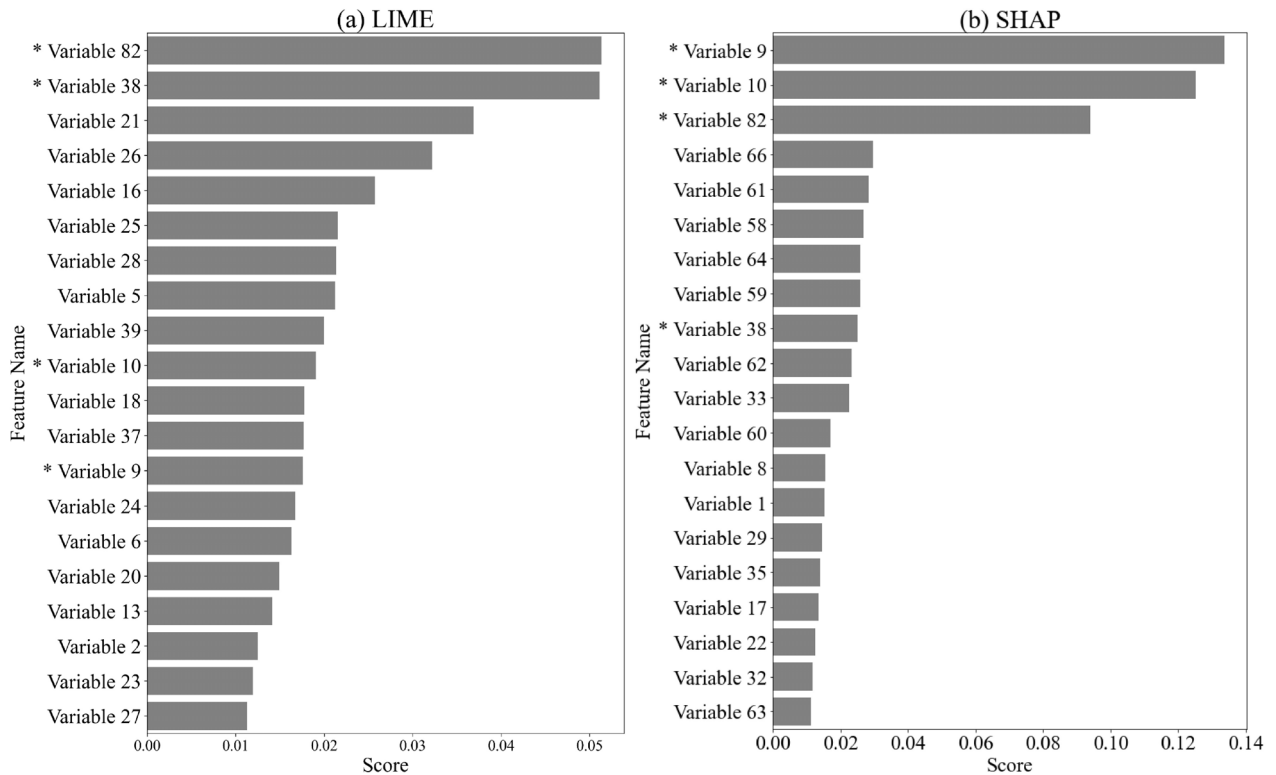


Figure 6. Results of LIME and SHAP Algorithm for Defect data

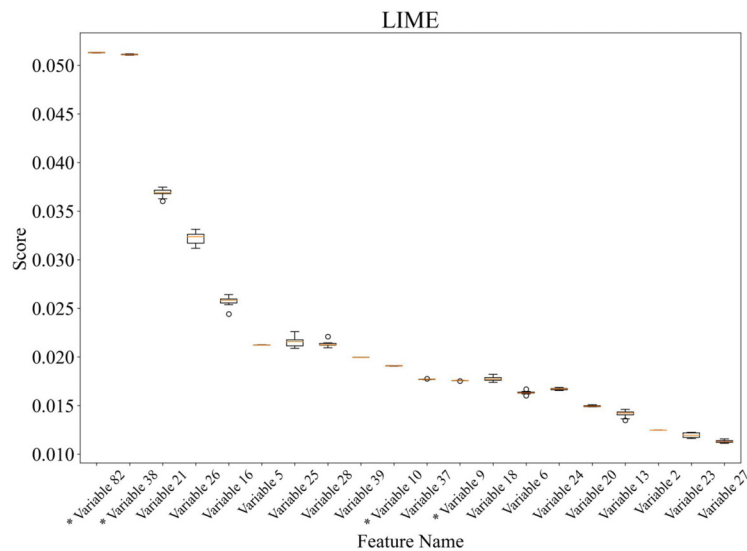


Figure 7. Box plot Result of LIME Algorithm after 10 Replicates



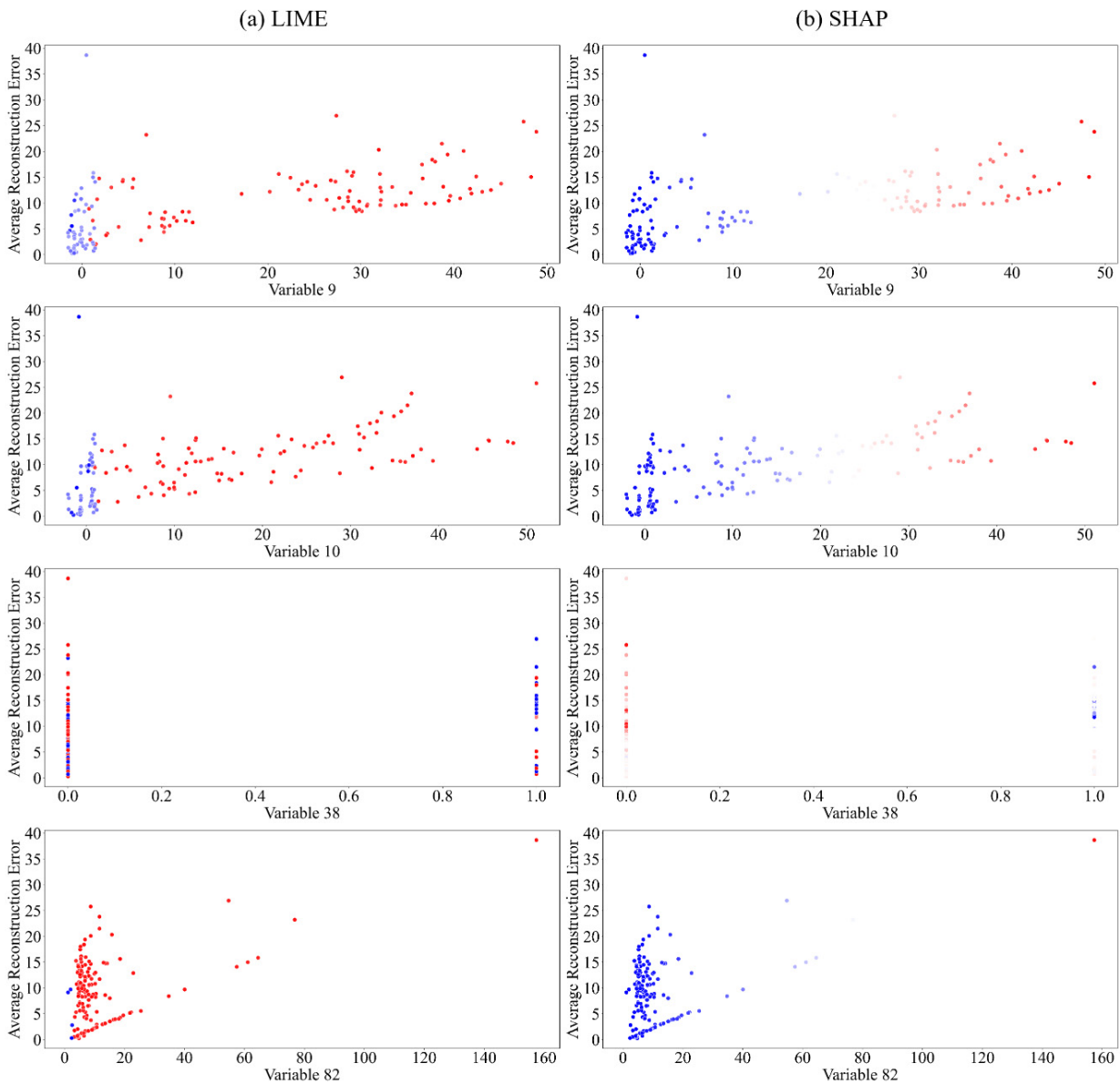


Figure 8. Scatter Plots of Important Features and Average Reconstruction Error

지체 시간, 설비, 정전척 구동 환경, 적용 면적 등이다. 해당 변수들은 다른 변수들 대비 높은 중요도 점수를 획득하였으며, 두 가지 설명 가능한 인공지능 알고리즘 모두에서 상위권에 위치하는 일관된 결과가 도출되었다. 구체적으로 각 변수들과 재구축 절대오차 평균값 간의 상관관계를 살펴보고자 산점도 그림을 그려 해석하였다. 그 결과는 <Figure 8>과 같으며, x축은 주요 변수들, y축은 재구축 절대오차 평균값에 해당한다. 또한, 각 점들의 색상은 LIME 및 SHAP 알고리즘에서 도출된 변수 별 중요도 점수의 크기로, 붉은색일수록 큰 변수 중요도 점수를 갖는 관측치임을 의미한다. 해당 산점도 그림을 통하여 지체 시간이 길어질수록, 적용 면적이 넓어질수록 불량률이 높아짐을 검증하였다. 이에 따라 지체 시간이 짧도록, 그리고 적용 면적이 작도록 공정 상에서 관리하는 것이 불량을 방지하는 데에 도움이 될 것으로 보인다. 더 나아가, 현업 엔지니어

로부터도 각 요인들이 현업에서 실제로 주요한 인자로 관리된다는 평가를 받아 정성적으로도 해석력의 타당성이 입증되었다.

### 5. 결론

최근 급속도로 발전하는 디스플레이 산업의 생산 공정 고도화를 위하여 제품 불량 탐지와 원인 분석 연구를 수행하였다. 본 연구에서는 디스플레이 주요 생산 공정인 FAB 공정의 정전척 설비에 오토인코더 기반 이상치 탐지 알고리즘을 적용하여 제조 효율을 높이고자 하였다. 특히, 오토인코더 기반 이상치 탐지 알고리즘의 결과에 설명 가능한 인공지능을 적용하기 위하여 심층 신경망 기반 REPM을 제안하였으며 불량 원인 변수

선정 및 현장 엔지니어의 검증으로 제안 모델의 효용성을 입증하였다.

본 연구는 이상 탐지 알고리즘과 설명 가능한 인공지능을 디스플레이 FAB 공정의 정전척 설비에 적용 가능하도록 고안한 REPM 방법을 제안하였는데 의의가 있다. 실제 생산 공정에 제안 방법론을 적용한다면 기기의 고장을 사전에 예측하고 대응함으로써 제품의 완결성을 높이고, 기기 가동 중단에 따른 비용도 절감할 수 있다. 또한, 도출된 정량적인 불량 원인 인자는 현장 엔지니어의 유지 및 보수 작업의 의사결정에 기여할 수 있다. 더 나아가, 디스플레이 산업의 전체 생산 공정 효율화를 증진시킬 수 있을 것이다. 본 연구에서는 디스플레이 FAB 공정 정전척 설비 데이터에 대해 실험을 진행하였지만, 추후 다양한 산업의 제조 공정 데이터에 적용하여 불량 탐지부터 원인 파악까지 수행한다면 더욱 높은 활용도를 가질 것이라 기대한다.

## 참고문헌

- Adadi, A. and Berrada, M. (2018), Peeking inside the black-box: a survey on explainable artificial intelligence (XAI), *IEEE Access*, 6, 52138-52160.
- An, J. and Cho, S. (2015), Variational autoencoder based anomaly detection using reconstruction probability, *Special lecture on Industrial Engineering*, 2(1), 1-18.
- Antwarg, L., Miller, R. M., Shapira, B., and Rokach, L. (2019), Explaining anomalies detected by autoencoders using SHAP, *ArXiv Preprint ArXiv:1903.02407*.
- Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., and Benjamins, R. (2020), Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82-115.
- Chalapathy, R. and Chawla, S. (2019), Deep learning for anomaly detection: A survey, *ArXiv Preprint ArXiv:1901.03407*.
- Hubert, M. and Van der Veeken, S. (2008), Outlier detection for skewed data, *Journal of Chemometrics*, 22(3-4), 235-246.
- Jung, S. and Lim, H. (2021), AI/BIG DATA-based Smart Factory Technology Status Analysis for Effective Display Manufacturing, *Journal of the Korea Institute of Information AndCommunication Engineering*, 25(3), 471-477.
- Kang, S. B., Lee, J. H., Song, K. Y., and Pahk, H. J. (2009), Automatic defect classification of TFT-LCD panels using machine learning, *2009 IEEE International Symposium on Industrial Electronics*, 2175-2177.
- Kingma, D. P. and Welling, M. (2013). Auto-encoding variational bayes, *ArXiv Preprint ArXiv:1312.6114*.
- Ku, D. (2019), Defect inspection in display panel using concentrated auto encoder, *International Journal of Optics*, 2019.
- Lundberg, S. M. and Lee, S.-I. (2017), A unified approach to interpreting model predictions, *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 4768-4777.
- Makhzani, A., Shlens, J., Jaitly, N., Goodfellow, I., and Frey, B. (2015), Adversarial autoencoders, *ArXiv Preprint ArXiv:1511.05644*.
- Pang, G., Shen, C., Cao, L., and Van Hengel, A. D. (2021), Deep learning for anomaly detection: A review, *ACM Computing Surveys (CSUR)*, 54(2), 1-38.
- Park, H. (2019), *Design and Implementation of the Anomaly Detection for Flat Panel Display*, Master's thesis, Hanyang University.
- Ribeiro, M. T., Singh, S., and Guestrin, C. (2016), "Why should I trust you?" Explaining the predictions of any classifier, *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135-1144.
- Ruff, L., Kauffmann, J. R., Vandermeulen, R. A., Montavon, G., Samek, W., Kloft, M., Dietterich, T. G., and Müller, K.-R. (2021), A unifying review of deep and shallow anomaly detection, *Proceedings of the IEEE*, 756-795.
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1985), *Learning internal representations by error propagation*, University of California, San Diego, Institute for Cognitive Science.
- Ryu, J.-H., Heo, M.-O., and Zhang, B.-T. (2017), CNN-based Classification Methods for Filtering of Defect Pixel Circuit Images, *Journal of Computing Science and Engineering*, 907-909.
- Sahno, J. and Shevtshenko, E. (2014), Quality improvement methodologies for continuous improvement of production processes and product quality and their evolution, *9th International DAAAM Baltic Conference "Industrial Engineering"*, 181-186.
- Shrikumar, A., Greenside, P., and Kundaje, A. (2017), Learning important features through propagating activation differences, *2017 International Conference on Machine Learning*, 3145-3153.
- Stojanovic, L., Dinic, M., Stojanovic, N., and Stojadinovic, A. (2016), Big-data-driven anomaly detection in industry (4.0): An approach and a case study, *2016 IEEE International Conference on Big Data (Big Data)*, 1647-1652.

## 저자소개

**김서연** : 한양대학교 산업공학과에서 2020년 학사 학위를 취득하고 고려대학교 산업경영공학과에서 2022년 석사학위를 취득하였다. 연구분야는 인공지능, 머신러닝, Explainable Artificial Intelligence이다.

**이지윤** : 명지대학교 산업공학과에서 2017년 학사 학위를 취득하고 고려대학교 산업경영공학과에서 박사과정에 재학 중이다. 연구 분야는 Machine Learning Algorithms, Explainable AI for Time-Series Data이다.

**목충협** : 고려대학교 산업경영공학과에서 2019년 학사 학위를 취득하고 고려대학교 산업경영공학과에서 박사과정에 재학 중이다. 연구분야는 Anomaly Detection, Self-Supervised Learning, Multi-task Learning이다.

**김상훈** : 고려대학교 산업경영공학과에서 2019년 학사 학위를 취득하고 고려대학교 산업경영공학과에서 박사과정에 재학 중이다. 연구분야는 Machine Learning Algorithms, Open-Set Recognition Algorithms, Lable Noise Learning이다.

**문석호** : 고려대학교 산업경영공학과에서 2019년 학사 학위를 취득하고 고려대학교 산업경영공학과에서 박사과정에 재학 중이다. 연구분야는 Anomaly Detection, Time series representation, Self-semi Supervised Learning 이다.

**경윤영** : 건국대학교 산업공학과에서 2012년 학사 학위를 취득하고 현재 성균관대학교 데이터사이언스융합학과에서 석사과정 재학 중이며 삼성디스플레이 AI팀에 재직하고 있다. 주요 연구 분야는 제조 분야 AI 모델링 및 알고리즘 개발이다.

**진영길** : 서울대학교 원자핵공학과에서 2010년 학사, 2012년 석사, 2018년 박사 학위를 취득하고 현재 삼성디스플레이 AI팀에 재직하고 있다. 주요 연구 분야는 제조 분야 AI 모델링 및 알고리즘 개발이다.

**이우진** : 경희대학교 화학과에서 2009년 학사, 서울대학교 재료공학부에서 2011년 석사, 2015년 박사 학위를 취득하고 현재 삼성디스플레이에 재직하고 있다. 주요 연구 분야는 Encapsulation

for organic light-emitting devices, Machine Health Monitoring 이다.

**최중명** : 아주대학교 산업공학과에서 2001년 학사, Iowa State University, Industrial Manufacturing & Systems Engineering에서 2005년 석사, 2009년 박사 학위를 취득하였고 Post-Doc 연구원 후 2011년부터 삼성디스플레이 MES(Manufacturing Execution System)팀에 재직하고 있다. 주요 업무는 삼성디스플레이 Smart Factory 구축을 위한 제조 시스템 지능화(System Intelligence) Application 개발을 담당하고 있고, 주요 연구 분야는 Engineering Automation, 제조 라인 자율 조정 및 제어를 위한 제조 빅데이터 기반 모델링 및 알고리즘 개발이다.

**김성범** : 고려대학교 산업경영공학부 교수로 2009년부터 재직하고 있으며, 인공지능공학연구소 소장 및 기업산학협력센터 센터장을 맡고 있다. 미국 텍사스주립대학교 산업공학과 교수를 역임하였으며, 한양대학교 산업공학과에서 학사학위를 미국 Georgia Institute of Technology에서 산업공학 석사 및 박사학위를 취득하였다. 연구 분야는 인공지능, 머신러닝, 최적화이다.

## &lt;Appendix&gt;

**Table 4.** Explanation Score Results of LIME for 10 Defect Samples

	Variable 82	Variable 32	Variable 21	Variable 26	Variable 16	Variable 5	Variable 25	Variable 28	Variable 39	Variable 10
Sample 1	0.0521	0.0518	0.0047	0.0092	0.0294	0.0228	0.0000	0.0361	0.0186	-0.0199
Sample 2	0.0529	0.0530	0.0095	0.0732	0.0000	0.0193	0.0000	0.0000	0.0195	-0.0197
Sample 3	0.0548	-0.0516	0.0207	0.0207	0.0005	0.0211	0.0207	0.0000	0.0206	-0.0064
Sample 4	0.0529	-0.0510	0.0357	0.0165	0.0941	0.0231	0.0941	0.0403	0.0186	-0.0101
Sample 5	0.0529	-0.0523	0.0199	0.0199	0.0000	0.0213	0.0199	0.0199	0.0202	0.0249
Sample 6	0.0528	-0.0535	0.0178	0.0313	0.0000	0.0199	0.0424	0.0424	0.0176	0.0247
Sample 7	0.0515	-0.0509	0.0264	0.0058	0.0014	0.0180	0.0264	0.0264	0.0197	0.0265
Sample 8	0.0535	-0.0519	0.0461	0.0116	0.0000	0.0219	0.0461	0.0059	0.0210	0.0252
Sample 9	0.0528	-0.0530	0.0218	0.0490	0.0257	0.0200	0.0055	0.0490	0.0202	0.0222
Sample 10	0.0534	0.0530	0.0335	0.0000	0.0000	-0.0228	0.0602	0.0602	-0.0213	-0.0084

**Table 5.** Explanation Score Results of SHAP for 10 Defect Samples

	Variable 9	Variable 10	Variable 82	Variable 66	Variable 61	Variable 58	Variable 64	Variable 59	Variable 38	Variable 62
Sample 1	-0.0092	-0.112	0.2464	-0.0277	-0.0250	-0.0322	-0.0121	-0.0252	0.0195	-0.0220
Sample 2	0.0017	-0.0174	1.6746	-0.0075	-0.1122	-0.0224	-0.0179	-0.0240	0.0244	-0.0595
Sample 3	0.0003	0.0009	0.0335	-0.0167	-0.0254	-0.0114	-0.0121	-0.0247	0.0152	-0.0208
Sample 4	-0.0010	-0.0005	0.4050	-0.0306	-0.0336	-0.0279	-0.0154	-0.0261	0.0146	-0.0258
Sample 5	0.0114	0.2623	0.0846	-0.0263	-0.0581	-0.0260	-0.0140	-0.0225	0.0139	-0.0160
Sample 6	0.0147	0.2243	0.0205	-0.0663	-0.0110	-0.0209	0.5799	-0.0234	0.0111	-0.0038
Sample 7	-0.0048	0.1359	0.0515	-0.0673	-0.0183	-0.0291	0.5727	-0.0258	0.0170	-0.0106
Sample 8	-0.0059	0.2568	0.0580	-0.0498	-0.0186	-0.0192	0.5666	-0.0242	0.0082	-0.0026
Sample 9	0.0636	0.1066	0.8028	-0.0243	-0.0239	-0.0313	-0.0116	-0.0291	0.0160	-0.0578
Sample 10	-0.0010	-0.0026	0.1718	-0.0238	-0.0267	-0.0171	-0.0113	-0.0265	0.0214	-0.0177