

# Explainable AI Application for Machine Predictive Maintenance

Kang Min Cheon\* · Jaekyung Yang\*\*†

\*LG Energy Solution

\*\*Department of Industrial and Information Systems Engineering, Jeonbuk National University

## 설명 가능한 AI를 적용한 기계 예지 정비 방법

천강민\* · 양재경\*\*†

\*LG에너지솔루션

\*\*전북대학교 산업정보시스템공학과

Predictive maintenance has been one of important applications of data science technology that creates a predictive model by collecting numerous data related to management targeted equipment. It does not predict equipment failure with just one or two signs, but quantifies and models numerous symptoms and historical data of actual failure. Statistical methods were used a lot in the past as this predictive maintenance method, but recently, many machine learning-based methods have been proposed. Such proposed machine learning-based methods are preferable in that they show more accurate prediction performance. However, with the exception of some learning models such as decision tree-based models, it is very difficult to explicitly know the structure of learning models (Black-Box Model) and to explain to what extent certain attributes (features or variables) of the learning model affected the prediction results. To overcome this problem, a recently proposed study is an explainable artificial intelligence (AI). It is a methodology that makes it easy for users to understand and trust the results of machine learning-based learning models. In this paper, we propose an explainable AI method to further enhance the explanatory power of the existing learning model by targeting the previously proposed predictive model [5] that learned data from a core facility (Hyper Compressor) of a domestic chemical plant that produces polyethylene. The ensemble prediction model, which is a black box model, was converted to a white box model using the Explainable AI. The proposed methodology explains the direction of control for the major features in the failure prediction results through the Explainable AI. Through this methodology, it is possible to flexibly replace the timing of maintenance of the machine and supply and demand of parts, and to improve the efficiency of the facility operation through proper pre-control.

**Keywords :** Prediction Model, Explainable AI, Predictive Maintenance

### 1. 서 론

과거 설비 전문 기술자가 조그만 망치를 들고 공장을 점검하던 시절이 있었다. 이 설비 전문 기술자는 소리만

듣고도 설비의 상태를 알 수 있었다. 현재 대부분 공장 및 현장에서는 이런 모습이 보이지 않는다. 신입 관리자의 자질을 탓하기도 하지만, 그보다는 날로 거대해지고 복잡해지는 생산 현장이 과거처럼 장인 한 사람의 탁월한 역량에 의존하는 방식으로는 감당하기 어려워졌기 때문일 것이다. 이처럼 현실적으로 확보하기 어려운 베테랑 설비 전문가를 대신할 수 있는 것은 예지정비(Predictive Maintenance)기술이다. 예지정비는 장비가 고장 나면 수리하는 사후적인 활동이 아니다.

Received 13 December 2021; Finally Revised 20 December 2021;  
Accepted 21 December 2021

† Corresponding Author : jkyang@jbnu.ac.kr

예지정비는 관리 대상 장비와 관련된 수많은 데이터를 모아 예측 모형을 만드는 데이터 과학 기술이다. 단지 한 두 가지의 징조로 장비의 장애를 예측하는 것이 아니라 수많은 증상과 실제 장애의 이력 데이터를 수치화하고 모델화한다.

이 예지정비 방법으로 이전에는 통계적 방법이 많이 사용되었지만 최근에는 기계학습 기반의 방법들이 많이 제시되고 있다. 그런 제안된 기계학습 기반의 방법들이 보다 정확한 예측 성능을 보인다는 점에서 바람직하다고 할 수 있다. 다만 의사결정나무 기반 모델 등 몇가지 학습모델을 제외하고 학습모델의 어떤 속성(Attribute, Feature, 또는 Variable)이 예측 결과에 어느 정도의 영향을 미쳤는지를 설명하기 매우 어려운 구조(Black-Box Model)를 가지고 있다. 이를 극복하고자 최근에 제안되고 있는 연구가 설명 가능한 AI(Explainable Artificial Intelligence)이다. 기계학습 기반의 학습모델의 결과를 사용자가 이해하기 쉽도록 하고 이를 신뢰할 수 있도록 하는 방법론이라고 할 수 있다.

본 논문에서는 폴리에틸렌을 생산하는 국내 한 화학공장의 핵심설비를 대상으로 데이터를 기반으로 생성된 학습 모델[5]을 대상으로 하여, 기존의 학습모델의 설명력을 더 높이고자 설명 가능한 AI 방법을 적용하는 연구를 하였다. 이상한 고장 예측 모델을 설명 가능한 AI를 통해 공정을 통제할 수 있도록 설비 제어 설정 방안을 제시하였다.

## 2. 기존 연구

설비관리란 생산성과 수익성을 향상시키기 위하여 설비의 계획, 구축, 유지, 개선함으로써 설비의 기능을 최대한 활용하려고 조치하는 활동이다. 설비를 관리하는 목적은 신뢰성을 확보하고, 보전성을 향상시키며, 경제성을 추구하고, 가용성을 증대시키기 위한 것이다. 설비는 사용기간의 경과와 더불어 마모, 오손, 변형, 변질 등이 진행되어 노후 열화하여 간다. 당초 계획된 기능을 유지하려면 고장을 점검하고 정비하여 부품을 교체나 수리를 하여야 한다. 이를 보통 유지, 보수, 보전 등의 용어로 사용된다. 전통적으로 유지보수 방법은 계획보전(Planned Maintenance)과 비계획 사후보전(Breakdown Maintenance, BM)으로 구분된다. 계획보전은 예방보전(Preventive Maintenance, PM)과 개량보전(Corrective Maintenance, CM)으로 구분되고, 예방보전은 다시 시간기준보전(Time Based Maintenance, TBM)과 상태기준보전(Condition Based Maintenance, CBM)으로 분류된다[11]. 최근에는 머신러닝과 알고리즘을 통해 데이터를 분석해 장애 탐지와 수명 예측을 가능하게 하는 PdM(Predictive Maintenance)으로 발전하고 있다[2]. 설비의 고장 예측은 이상 탐지 방법이 사용되기도 한다. 하지만 단순히 이상을 탐지한다고 해서 설비의 고장을 예측

한다고는 할 수 없다. 고장 예측 대상으로는 설비 또는 부품, 센서 등 예측하고자 하는 대상을 선정한다. 디지털 센서 상태 데이터를 LSTM(Long Short Term Memory)을 통해 분석하여 고장을 진단하고 예측한 모델 구축 사례가 있다[12]. 회전 기기의 중요한 부품 고장 예측 사례로는 DNN(Deep Neural Network)과 머신러닝 알고리즘을 활용한 베어링 고장 예측 사례가 있다[13]. 회전기기의 이상 진단 사례로는 회전기기의 신호처리 데이터를 정상그룹과 비정상 그룹 선정 후 마할라노비스 거리(Mahalanobis distance) 계산하고 타구치(Taguchi)법을 통해 주요 속성치를 선정한 사례가 있다[16].

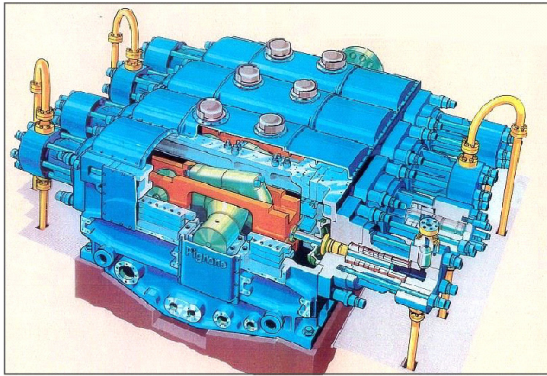
최근의 설명가능한 AI 연구로는 고장예지(PHM: Prognostics and Health Management)에 적용된 사례는 국내에서는 찾아볼 수 없고, 국외에서 최근 새로운 연구가 시작 단계에 있음을 알 수 있다. 그 외 분야에서의 연구로는 신용 정보 데이터를 활용한 인공지능 기반의 신용평가 알고리즘에서 데이터의 다양한 속성들 중 어떤 속성이 결과 도출에 큰 영향을 끼쳤는지 도출하기 위해 설명가능한 AI를 활용하였다[6]. 교수자를 위한 학습 지원 시스템은 인공지능 모델인 DKT와 XGBoost를 이용해 학습자의 지식 상태를 모델링하고, 그 결과를 설명 가능한 AI 기법인 LRP와 SHAP(SHapley Additive exPlanation)으로 분석해 학습자의 지식 상태를 해석 가능한 형태로 교수자에게 제공한 연구도 제안되었다[9]. 통합된 속성 기여도 분석 기법 개발을 위하여 인공지능의 입력 속성 각각에 기여도를 부여함으로써 판단을 설명하고자 설명 가능한 AI를 적용 방향을 제시했다[10].

정찬일은 머신러닝 모델 중 LSTM 알고리즘을 이용한 프로세스 예측 모델을 개발하여 프로세스의 수행결과를 예측하였고 다시 이 예측모델에 설명 가능한 AI를 적용하여 분석함으로써 최종적으로 프로세스 수행결과에 가장 큰 영향을 미치는 원인을 탐색하는 작업을 자동화하였다[8].

## 3. 연구 방법

### 3.1 연구 방법 절차

본 논문에서는 단순히 설비 상태 이상을 예측하는 것을 넘어 고장의 원인을 설명 가능한 AI를 통해 해석하고 설비 고장이 발생하지 않도록 제어하는데 이용하고자 한다. 본 연구 대상 공정 설비인 아래 <Figure 1>의 초고압왕복동압축기(Hyper Compressor)는 해당 공정에서 연속 생산 공정의 핵심설비이기 때문에 단순 이상이 발생한다고 설비를 점검할 수 없는 상황이다.



<Figure 1> Hyper Compressor

본 연구의 제안 방법론은 아래 <Table 2>에 설명되어있다. 우선 폴리에틸렌 생산 공정에 대한 현황 분석을 하였다. 공정과 설비에 대한 이해를 높이기 위해 설비 담당자와 인터뷰를 진행하였으며, 데이터 정합성을 위해 생산 및 설비 담당자들과 수많은 피드백 과정을 거쳤다.

데이터는 대상 기업의 공정 정보시스템 중 RTDB(Real Time DataBase)와 CMS(Condition Monitoring System), LIMS(Laboratory Information Management System), MES를 통해 수집하였다. 아래 <Table 1>에 중요 속성(Attribute)과 예시 데이터를 제시하였다.

<Table 1> Raw Data Set Sample

Ethylene Pressure	Cylinder Temp.	Bearing Temp.	Plunger Temp.	Flux	Class
254.6	61.5	63.4	54.4	339.3	0
254.6	61.6	63.4	54.4	340.0	0
254.7	61.9	63.4	54.5	332.8	0
254.8	62.0	63.5	54.5	334.7	0
254.0	62.4	63.5	54.6	336.3	0
254.2	62.5	63.6	54.6	334.7	0
254.8	62.6	63.6	54.6	330.0	0
255.8	62.8	63.6	54.7	329.4	1
255.8	62.9	63.6	54.7	328.0	1
255.6	63.0	63.6	54.8	328.3	1
:	:	:	:	:	:
255.7	63.7	63.8	55.1	325.2	0

활용 데이터는 2011년부터 2019년까지 약 9년간의 데이터가 활용되었으며, 모델의 정확도 향상을 위해 Primary Compressor, Intercooler, Reactor, Pump 등 Hyper Compressor와 연계되는 설비의 센서 및 공정 속성(약400EA)들을 데이터 셋에 포함시켰다.

각 속성은 초고압왕복동압축기의 부위별(실린더/베어링/플런저 등) 압력, 온도, 유량(Flux)이며 Class는 고장 유무이다.

<Table 2> Explainable AI Application Processes for Machine Predictive Maintenance

Order	Procedures
1	Environmental analysis
2	Data preprocessing for about 400 attributes
3	Configuring training sets based on failure history
4	Composing training set (normal data 70% vs abnormal data 30%)
5	Learning the ensemble model with a set of dependent and independent variables
6	Predicting risk scores & feature (attribute) importance evaluation
7	Learning model evaluation
8	Explainable AI (SHAP Tree Explainer)generation
9	Extracting SHAP value and evaluating contribution of attributes

수집된 데이터는 전처리 과정을 거쳐 학습데이터 셋을 구성한다. 구성된 학습 데이터(Training set)셋을 학습시켜 이상 예측 모델을 생성한다. 이상 예측 모델은 지도학습기반의 XGBoost(Extreme Gradient Boosting) 앙상블 모델을 적용하였다. XGBoost 앙상블 모델은 CART(Classification and Regression Tree)의 부스팅 버전으로 최근 그 정확도가 인정되어 많이 활용되고 있는 알고리즘이다. 이상(Abnormal) 예측 모델에 대한 평가 및 XGBoost 앙상블 모델 예측 정확도에 대한 비교를 하였으며 그 모델의 우수성 및 효율성을 확인하였다. 학습모델과 그 평가부분에 대한 자세한 설명은 이전 논문[5]에서 확인할 수 있다. 또한 설명가능한 AI를 통해 앙상블 모델을 해석해 주요 속성의 제어 방향성을 제시한다.

### 3.2 고장 위험 예측 방법

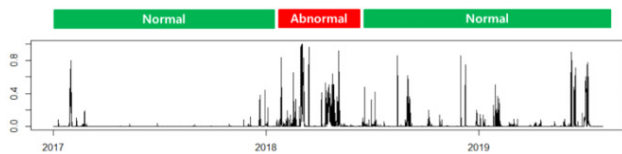
본 연구에서는 분류 기반의 XGBoost 학습 모델이 사용되었다. XGBoost 알고리즘은 앞에서 설명했듯이 좋은 성능을 보여주고 있는 의사결정나무 기법으로[1, 7], 과적합을 방지와 병렬화 장점을 가지고 있다[17]. XGBoost는 다른 Gradient Boosted Machine과 비교했을 때 과적합을 방지하기 위해 더 많은 정규화 항을 사용한다는 특징을 가지고 있어서, 이를 통해 더 좋은 성능을 발휘한다[3]. XGBoost 알고리즘을 자세히 설명하면 다음과 같다. 먼저 CART 모델을 입력 데이터로 학습시키고 학습된 CART 모델을 개선하기위해 CART 알고리즘의 예러에 대해 반복 학습한다. t단계에서의 예측을 위한 함수는 식 (1)과 같다[14].

$$f_t^{(t)} = \sum_{k=1}^t f_k(x_i) = f_t^{(t-1)} + f_t(x_i) \quad (1)$$

$x_i$ 는 input 데이터이고  $f_k(x_i)$ 는 t단계에서의 예측 정도다.

$f_i^{(t)}$ 와  $f_i^{(t-1)}$ 는 t-1 단계에서의 예측 정도다.

XGBoost 학습모델은 정상, 비정상에 대한 시간 구간 설정은 학습모델의 성능에 큰 영향을 준다. 이는 단순 설비의 이상 예측이 아닌 고장 시점 예측과도 연결된다. 또한 설비의 과거 고장 이력 및 상태에 대해서 전체 기간 데이터를 수집하는 것은 불가능하기에 기록이 남아있는 고장 이력을 중심으로 학습 데이터 구간을 설정하였다. 설비데이터와 설비 정상, 비정상(고장 포함)을 포함한 학습데이터를 이용해 학습한 학습 모델을 통해 고압 압축 설비의 위험상황 탐지 및 예측을 하였다. 설비 고장 위험 정도 예측에서 중요한 점은 설비의 이상 전 전조 위험신호의 여부이다. 따라서 학습 모델의 위험신호에 대한 예측 결과에서 전조 현상이 잘 보여지고 있는가에 대한 확인이 필요하다. 설비의 위험 예측 모델을 학습한 후, 테스트 데이터를 이용해 모델을 검증하였다. 검증 방법으로는 테스트 데이터를 학습데이터와 마찬가지로 정상 및 비정상 구간으로 나누고, 그 실제 결과와 XGBoost 모델의 예측 결과 비교를 통해 예측 정확도를 확인하였다. 학습데이터와 테스트데이터의 비율은 7 대 3이다. <Figure 2>은 설비 고장 위험 예측 모델의 정상과 비정상 구간을 분류한 예시이다. 예측 정도가 0에 근접한 경우에는 설비가 정상 상태에 가깝고 예측 정도가 1에 가까운 경우는 설비가 이상 상태라고 판단할 수 있다.



<Figure 2> Example of the Facility Risk Prediction

### 3.3 설명 가능한 AI 적용 절차

XGBoost 는 Tree 기반 모델이기에 입력 속성(센서 등)을 통해 설비 상태를 예측할 수 있다. 이러한 응용문제를 해결하기 위해서는 예측에 대한 정확한 해석이 가능해야 한다. 모델이 해석 가능하다는 것은 모델의 속성이 결과 예측에 어떻게 사용되는지 이해할 수 있음을 의미한다 [15]. Scott M. Lundberg는 게임이론의 Shapley Value에 의해 정의된 SHAP(SHapley Additive exPlanation)을 다루기 쉽게 계산할 수 있는 새로운 로컬 설명 방법인 Tree Explainer를 제안했다[4]. Shapley value는 게임이론 기반으로 하나의 속성 중요도를 알기 위해 다수 속성들의 조합을 구성하고 해당 속성의 유무에 따른 평균적인 변화를 통해 얻어낸 값이다. <Figure 3>에 보듯이 Shapley Value는 모든 가능한 조합에 대해서 하나의 속성 기여도를 종합적으로 합한 값이다. Shapley value를 계산하기 위한 함수

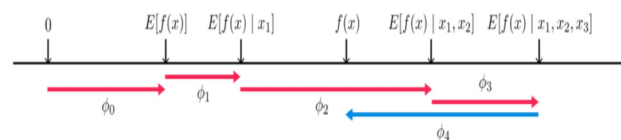
$\phi(v)$ 는 다음과 같다.

$$\phi(v) = \sum_{s \subset A, t \in A} \frac{(|S|-1)! \times (n-|S|)!}{n!} \times (v(S) - v(S-i)) \quad (2)$$

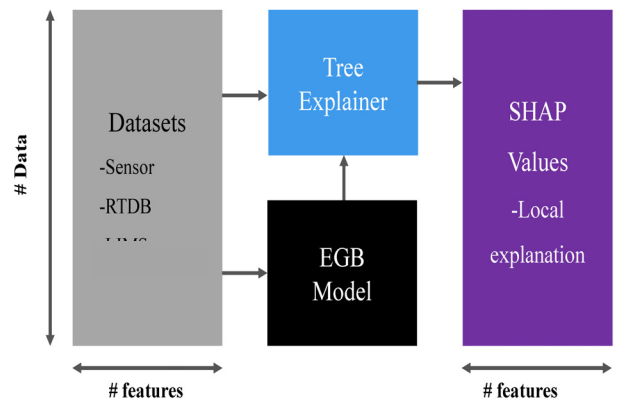
$n$ 은 전체 플레이어 수이고,  $|S|$ 는 집합  $S$ 의 개수이다.  $v(s)$ 는 플레이어  $i$ 가 협력했을 때 속성 함수 값이고,  $v(S-i)$ 는 플레이어  $i$ 를 제외한 나머지 플레이어들이 협력했을 때 속성함수 값이다. 따라서  $v(S) - v(S-i)$ 는 플레이어  $i$ 가 협력했을 때 기여하는 크기이고  $\frac{(|S|-1) \times (n-|S|)!}{n!}$ 는 각 협력하는 조합의 확률에 해당한다. 따라서 Shapley value는 전체 협력에 각 플레이어  $i$ 가 기여할 것으로 예상되는 기대값이다. 또한 SHAP은 예측에 대한 각 속성의 기여도를 계산하여 관측치  $x$ 의 예측값을 설명한다. SHAP 방법은 연합 게임 이론(Coalitional Game Theory)을 사용하여 Shapley value를 계산하고 관측치의 속성값은 연합에서 플레이어로서 역할을 한다. SHAP은 다음과 같은 수식으로 표현할 수 있다.

$$g(z') = \phi_0 + \sum_{j=1}^M \phi_j z'_j \quad (3)$$

$g$ 는 설명모델,  $z' \in \{0, 1\}^M$ 은 연합 벡터(Coalition Vector),  $M$ 은 최대 연합사이즈,  $\phi_j \in R$ 은 속성  $j$ 의 속성기여도 Shapley value이다[18].



<Figure3> SHAP Values Explain the Output of a Function  $f$  as a Sum of the Effects of Each Feature Being Introduced Into a Conditional Expectation

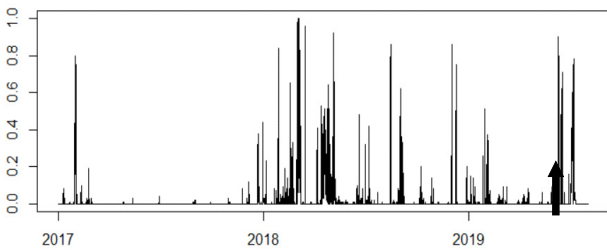


<Figure 4> Tree Model Explainer Process

<Figure 4>는 모델 Explainer 절차를 설명하고 있다. XGBoost 모델(그림에서는 EGB 모델로 표현)을 통해 추출된 주요 속성으로는 방향성을 확인할 수 없기 때문에 XGBoost 모델의 Explainer를 적용해서 각 속성의 SHAP value를 추출해야 한다. Explainer를 통해 모델에 영향을 주는 속성에 대해서 확인할 수 있으며 주요 속성의 제어 값에 따른 모델의 영향도 확인이 가능하다.

### 4. 실험 결과

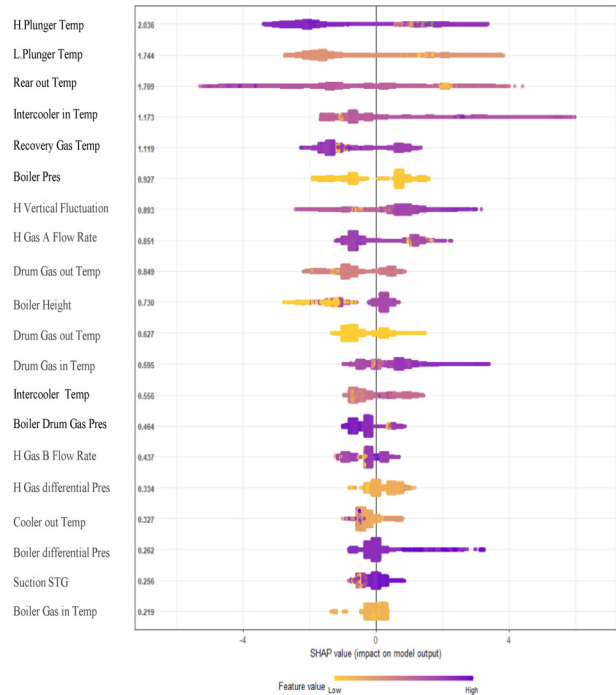
화학 공정의 고압압축 설비의 고장 위험을 예측하고 선택된 주요 속성을 추출하였으며, 이 속성을 기반으로 설비 고장 원인 파악에 이용하였다. 또한 설비 위험 예측 모델의 위험도가 상승하게 될 경우, 위험도 상승 속성에 대한 제어가 동반되어야 한다. 제시된 모델로는 주요원인을 알 수 있지만 주요 원인 속성의 제어 방향성에 대해서는 위험도로 설명하기에 한계가 있었다. 따라서 본 논문에서는 설비 위험 정도에 영향을 주는 주요 속성들의 실제 수치 별 방향성을 확인하고 앙상블 모델에 대한 설명 방법을 적용하였다. 먼저 아래 <Figure 5>같이 테스트 셋(2017년~2019년)에 대한 설비 위험 정도 결과를 모니터링 중 2019년도 6월 초에 설비 위험 정도(Threshold = 0.9 이상)가 상승했다고 가정한다.



<Figure 5> Risk Prediction Score Results for Test set

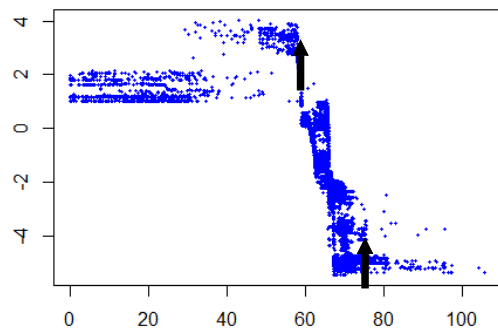
설비 위험 정도가 임계치를 넘어섰기 때문에 설비 고장 위험이 높아졌다고 판단 후 설비 점검 또는 공정 및 생산 조건을 제어하기 위하여 <Figure 6>과 같이 위험 예측 모델의 주요 원인 속성에 대한 SHAP값의 요약 결과를 확인할 수 있다.

XGBoost 모델의 Feature Importance 분석을 통해서 H.plunger 온도, 보일러압력, L.plunger 온도 등이 설비 고장에 주요 원인이라는 것을 확인할 수 있지만 해당 속성 제어를 위한 방향성은 확인하기 어렵다. 따라서 제안된 모델(XGBoost) Explainer를 통해 주요 속성들의 실제 값에 따른 모델 영향도(Impact on log-odds)를 확인할 수 있다.



<Figure 6> Risk Prediction Score Results for Test Set

<Figure 7>은 L.plunger 온도의 실제 값에 따른 모델 영향도를 확인할 수 있는 그래프이다. 그래프에서 L.plunger 값(Value) 확인 결과 “60” 전 후와 “70” 전 후로 값이 주로 분포되어 있는 것을 확인할 수 있다.

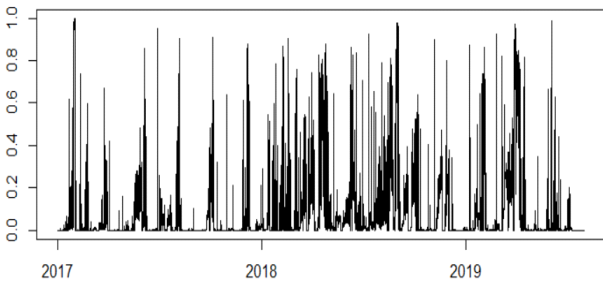


<Figure 7> Impact Log-odds According to L.plunger Temperature Values

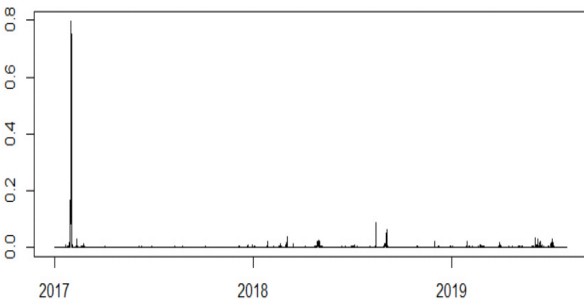
즉 “60” 전 후의 값은 Log-odds(Y축) 값이 양수(“0”이상)이고 “70” 전 후의 값은 Log-odds(Y축) 값이 음수인 것을 확인할 수 있다. 따라서 후단압축기출구온도 값을 “60”과 “70” 두 개 기준으로 값을 대체(Imputation) 후 예측 시뮬레이션을 해보았다.

<Figure 8>는 L.plunger 온도 값이 “60”일 경우 예측 모델 시뮬레이션 결과이다. 이는 원래 위험 예측 정도와 비교하여 위험 정도가 전반적으로 높아지는 것을 확인할 수

있었다. <Figure 9>의 L.plunger 온도 값이 “70”일 경우 예측 모델 시뮬레이션 결과이다. L.plunger 온도 값이 “60”일 경우와는 반대로 위험 정도가 대폭 낮아진 것을 확인할 수 있었다. 이는 Log-odds 값이 양수이며 값이 클수록 (1<2<3<4) 예측 모델의 위험 정도가 높아지고 Log-odds 값이 음수이며 값이 작을수록 예측 모델의 정도가 낮아지는 패턴을 보여준다.

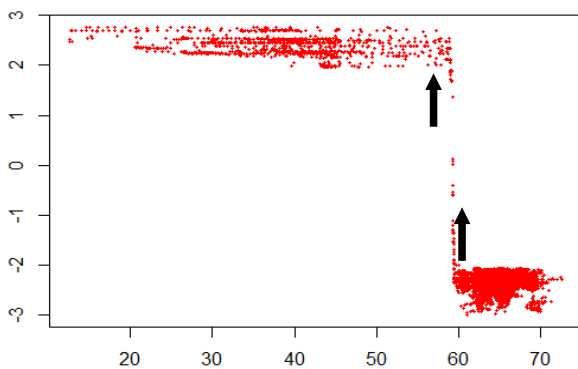


<Figure 8> Simulation Results for Predictive Models with a L.plunger Temperature Value of 60



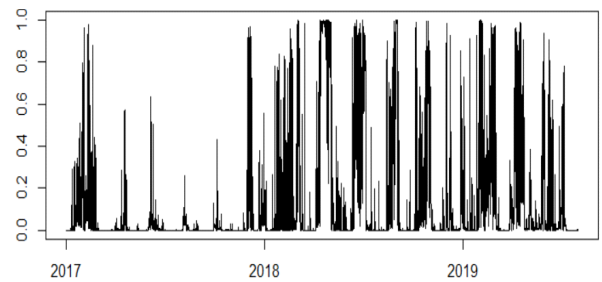
<Figure 9> Simulation Results for Predictive Models with L.plunger Temperature Value of 70

아래 <Figure 10>은 H.plunger 온도의 실제 값에 따른 모델 영향도를 확인할 수 있는 그래프이다.

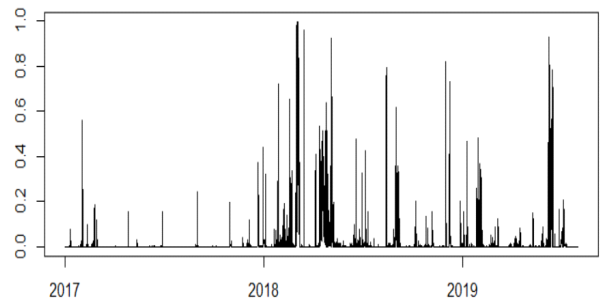


<Figure 10> Impact Log-odds according to Feature H.plunger Temperature Values

<Figure 10>의 그래프에서 H.plunger 온도 값(Value) 확인 결과 “58” 전 후와 “73” 전 후로 값이 주로 분포 되어 있는 것을 확인할 수 있다. 즉 “58” 전 후의 값은 Log-odds 값이 양수(0이상)이고 “73” 전 후의 값은 Log-odds 값이 음수 인 것을 확인할 수 있다. 따라서 H.plunger 온도 값을 “58”과 “73” 두 개 기준으로 값을 대체(Imputation) 후 예측 시뮬레이션을 해보았다.



<Figure 11> Simulation Results for Predictive Models with a H.plunger Temperature Value of 58



<Figure 12> Simulation Results for Predictive Models with a H.plunger Temperature Value of 73

위험도가 전반적으로 높아지는 것을 확인할 수 있었다. 위 <Figure 11>은 H.plunger 온도 값이 “58”일 경우 예측 모델 시뮬레이션 결과이고 <Figure 12>은 H.plunger 온도 값이 “73”일 경우 예측 모델 시뮬레이션 결과이다.

본 논문에서는 위험도 관리 방향성을 확인하기 위해 설비 위험도에 영향을 주는 주요 속성 중 L.plunger 온도, H.plunger 온도의 2가지 요인의 값에 따른 모델의 Log-odd 결과 값을 확인했다. 이를 통해 요인의 값 변동을 통해 설비 고장 예측 정도가 변화하는 것을 시뮬레이션을 통해서 확인할 수 있었다. 즉 앙상블 모델 Explainer를 통해 주요 속성의 제어 방향성에 대해서 설정이 가능한 것을 확인할 수 있었다. 주요 속성을 선정하고 일부 제어를 통해 설비 고장을 제어 가능하다는 부분에 초점을 두고 싶다. 또한 주요 속성 제어 시 공정 속성에 따른 제어 가능 여부도 향후 확인이 필요하다.

## 5. 결 론

본 논문에서는 화학공장의 특수 설비인 고압압축기의 과거 고장 이력을 수집하여 고장 위험 및 고장 상황을 설명하기 위해 설명가능한 AI 기법인 설명가능한 XGBoost 모델을 제시하였다. XGBoost 앙상블 예측모델은 설비 상태에 대한 객관적인 정보가 없거나 부정확한 과거 이력 정보들에 취약할 수 있다. 따라서 이에 대한 신뢰도를 높이기 위해서 설명 가능한 AI를 적용하였다. 설비 위험도에 영향을 주는 주요 속성들의 실제 수치 별 경향성을 확인할 수 있었다. 또한 주요 속성에 대해 선별이 필요한 경우 SHAP value를 통해 예측 모델 속성들의 영향도를 이용할 수 있다. 이는 주요 속성 별로 (+)기여도와 (-)기여도로 구성이 되어있으며 변동이 큰 속성일수록 주요 속성으로서 의미가 있다고 판단되며 특성 값이 낮은 속성인 경우 후보 속성에서 제외시켜 예측 모델의 주요 속성을 선별 할 수 있음을 보여주었다. 이는 설비 고장 위험의 원인을 설명할 수 있기에 이 속성값을 적절히 제어함으로써 설비 고장 위험을 낮추는데 사용할 수 있음을 보여주었다.

## References

- [1] Adam-Bourdarios, C., Cowan G., Germain, C., Guyon, I., Kgl, B., Rousseau, D., The Higgs boson machine learning challenge, *NIPS 2014 Workshop on High-energy Physics and Machine Learning*, 2015.
- [2] Ahn, D., Smart Factory and Data Analysis, ETNews, 2019.06.20.
- [3] Babajide Mustapha, I. and Saeed, F. Bioactive molecule prediction using extreme gradient boosting, *Molecules*, 2016, Vol. 21, No. 8, p. 983.
- [4] Baehrens, D., Schroeter, T., Harmeling, S., Kawanabe, M., Hansen, K., and Muller, K., How to explain individual classification decisions. In: *Journal of Machine Learning Research*, 2010, Vol. 11, pp. 1803-1831.
- [5] Cheon, K. and Yang, J. An Ensemble Model for Machine Failure Prediction, *Journal of Society of Korea Industrial and Systems Engineering*, 2020, Vol. 43, No. 1, pp 123-131.
- [6] Chun, Y., Kim, S., Lee, J., and Woo, J., Study on credit rating model using explainable AI, *Journal of the Korean Data & Information Science Society*, 2021, Vol. 32, No. 2, pp. 283-295.
- [7] Heckert, N.A., Filliben, CM., Croarkin, B., Hembree, William, F., Guthrie, P., Tobias, and Prinz, J., e-Handbook of Statistical Method-NIST/SEMATECH, 2002, [Online]. Available: <http://www.itl.nist.gov/div898/handbook/>.
- [8] Jung, C. and Lee, H., Explainable artificial intelligence based process mining analysis automation, *Journal of the Institute of Electronics and Information Engineers*, 2019, Vol. 56, No. 11, pp. 45-51.
- [9] Kim, S., Kim, W., Jang, Y., Kim, H., Development of Explainable AI-Based Learning Support System, *The Journal of Korean Association of Computer Education*, 2021, Vol. 24, No. 1, pp. 107-115.
- [10] Ko, G., Lim, G., and Cho, H., Survey on feature attribution methods in explainable AI, *Journal of KIISE*, 2020, Vol. 47, No. 12, pp. 1181-1191.
- [11] Korea Electric Power Co., Research on facility investment and failure evaluation system establishment to prevent failure of substation facilities, 2010, KEPCO.
- [12] Lee, G.H., An RNN-based Fault Detection Scheme for Digital Sensor, *The Journal of The Institute of Internet, Broadcasting and Communication*, 2019, Vol. 19, No. 1, pp. 29-35.
- [13] Lee, J.H., Fault diagnosis of bearings using machine learning algorithm, *Journal of the Korean Society of Marine Engineering*, 2019, Vol. 43, No.6, pp.455-462.
- [14] Lim, J., Introduction of conformity verification methodology for credit rating system, in *Risk Review of Financial Supervisory Service*, 2005.
- [15] Lundberg, S.M. and Lee, S.I., A Unified Approach to Interpreting Model Predictions. In: *Advances in Neural Information Processing Systems*, 2017, Vol. 30, pp. 4768-4777.
- [16] Park, S.G., A Fault Diagnosis on the Rotating Machinery Using MTS, *Transactions of the Korean Society for Noise and Vibration Engineering*, 2008, Vol. 18, No. 6, pp. 619-623.
- [17] Phoboo, Abha Eli. Machine learning wins the Higgs challenge, 2014, No. BULNA-2014-265.
- [18] Shapley, L.S., A value for n-person games. In: *Contributions to the Theory of Games*, 2, 1953, Princeton University Press, pp. 307-317.

## ORCID

Kang Min Cheon | <http://orcid.org/0000-0001-8877-2870>  
 Jaekyung Yang | <http://orcid.org/0000-0002-4904-1351>