



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

2024년 2월
석사학위 논문

인공지능을 활용한 비정상 상황에서
잔여 트립 시간 예측 시스템 개발

조선대학교 대학원

원자력공학과

오상원

인공지능을 활용한 비정상 상황에서
잔여 트립 시간 예측 시스템 개발

Development of a remaining trip time prediction
system in abnormal states using artificial intelligence

2024년 2월 23일

조선대학교 대학원

원자력공학과

오상원

인공지능을 활용한 비정상 상황에서
잔여 트립 시간 예측 시스템 개발

지도교수 나 만 균

이 논문을 공학 석사학위신청 논문으로 제출함

2023년 10월

조선대학교 대학원

원자력공학과

오상원

오상원의 석사학위논문을 인준함

위원장 김진원 (인)

위원 나만균 (인)

위원 김종현 (인)

2023년 12월

조선대학교 대학원

목차

표 목차	iii
그림 목차	iv
ABSTRACT	v
제 1 장 서론	1
제 2 장 인공지능 방법론	4
제 1 절 AE	4
제 2 절 LightGBM	5
제 3 절 SHAP	7
제 3 장 데이터 수집 및 전처리	9
제 1 절 데이터 수집	9
제 2 절 데이터 전처리	12
제 4 장 잔여 트립 시간 예측 시스템	13
제 1 절 잔여 트립 시간 예측 알고리즘	13
제 2 절 훈련 상태 진단 기능	15
제 3 절 비정상 시나리오 진단 기능	17
제 4 절 잔여 트립 시간 예측 기능	19
제 5 절 잔여 트립 시간 예측 결과 해석 기능	22
제 5 장 잔여 트립 시간 예측 시스템 인터페이스를 활용한 Case Study	26
제 1 절 잔여 트립 시간 예측 시스템 인터페이스	26
제 2 절 Case Study	29
제 6 장 결론	37

참고문헌 39

표 목차

표 1. 수집된 비정상 시나리오 정보 목록	11
표 2. AE 하이퍼파라미터 최적화 결과	17
표 3. 잔여 트립 시간 예측 모델 최적화 결과	21
표 4. SHAP 결과에 표시되는 변수 정보	25

그림 목차

그림 1. 잔여 트립 시간 예측 개념도	3
그림 2. AE 구조 개요	5
그림 3. 인공지능 방법들에 대한 설명 가능성 및 성능 비교	6
그림 4. Level-wise와 Leaf-wise 트리 분할 방법의 구조	6
그림 5. SHAP 적용 결과 예시	8
그림 6. CNS 인터페이스 개요	9
그림 7. 수집된 데이터의 형태	10
그림 8. 잔여 트립 시간 예측 알고리즘 개요	14
그림 9. 시그마 규칙에 따른 신뢰구간	16
그림 10. 비정상 시나리오 진단 기능 테스트 결과	18
그림 11. 전체 시나리오 모델을 활용한 잔여 트립 시간 예측 결과	21
그림 12. 단일 시나리오 모델을 활용한 잔여 트립 시간 예측 결과	22
그림 13. 가압기 PORV 고장 ‘열림’ 시나리오에서 잔여 트립 시간 예측 결과	23
그림 14. 가압기 PORV 고장 ‘열림’ 시나리오 초반부 예측값에 SHAP을 적용한 결과	25
그림 15. 가압기 PORV 고장 ‘열림’ 시나리오 중반부 예측값에 SHAP을 적용한 결과	25
그림 16. 가압기 PORV 고장 ‘열림’ 시나리오 후반부 예측값에 SHAP을 적용한 결과	25
그림 17. 잔여 트립 시간 예측 시스템 인터페이스 개요	28
그림 18. 가압기 압력 채널 고장 ‘고’ 시나리오에서 실제 압력과 계측 압력	29
그림 19. 실험 설정 흐름도	30
그림 20. 잔여 트립 시간 예측 시스템 결과 - 비정상 시나리오 주입	32
그림 21. 잔여 트립 시간 예측 시스템 결과 - 완화 조치 1 (가압기 히터 동작)	33
그림 22. 잔여 트립 시간 예측 시스템 결과 - 완화 조치 1 이후 회복 실패	34
그림 23. 잔여 트립 시간 예측 시스템 결과 - 인적오류 (가압기 스프레이 밸브 개방)	35
그림 24. 잔여 트립 시간 예측 시스템 결과 - 완화 조치 2 (가압기 스프레이 밸브 개도 조절)	36

ABSTRACT

Development of a remaining trip time prediction system in abnormal states using artificial intelligence

Sang Won Oh

Advisor : Prof. Man Gyun Na, Ph.D.

Department of Nuclear Engineering

Graduate School of Chosun University

When an abnormal state occurs in the Nuclear Power Plants (NPPs), operators perform mitigation actions to prevent the reactor trip and restore normal conditions. However, in abnormal states, the operator's tasks progress urgently checking many monitoring variables within a limited time. Consequently, operators may feel psychological pressure, which can lead to human errors. If mitigation actions fail due to human errors by the operator, the status of the NPPs may worsen and a reactor trip may occur. Furthermore, sudden shutdown by the reactor trip can cause the failure of numerous facilities in the NPPs. In addition, once the NPPs have been stopped, it takes a long time to restart, potentially leading to significant economic losses.

Accordingly, in this thesis, a remaining trip time prediction system was developed as part of an operator support system with the goal of assisting operator decision-making in abnormal states, aiming to prevent potential human errors. The remaining trip time prediction system is implemented through an algorithm that uses artificial intelligence (AI) and explainable AI (XAI) methods such as autoencoder, light gradient boosting machine, and shapley additive explanation. The

remaining trip time prediction algorithm provides diagnosis information about the abnormal state that has occurred and prediction information about the remaining time until the reactor trip. In addition, the XAI was used to provide the rationales for the remaining trip time prediction result and to improve the reliability of the AI.

An interface was developed to efficiently provide the results of the remaining trip time prediction algorithm. The interface can provide results of the remaining trip time prediction algorithm in real time regarding the current status of the NPPs. Therefore, the interface indicates the change in the results of the remaining trip time prediction algorithm according to the operator's mitigation actions. As a result, the remaining trip time prediction system provides diagnosis and prediction information and is expected to be helpful to the operator's tasks in abnormal states.

제 1 장 서 론

원자력 발전소에서 운전원은 발생 가능한 수많은 이벤트에 대응할 수 있도록 훈련한다. 그러나 계측기 오류, 기기 상태 이상, 배관 누설 등 다양한 원인으로 인한 실제 이벤트 상태가 발생하면 다음과 같은 이유로 운전원은 인적오류를 발생시킬 수 있다. 먼저, 업무 과중에 의한 운전원의 실수가 발생할 수 있다. 운전원은 이벤트 상황에서 많은 모니터링 변수를 감시한다. 모니터링 변수는 이벤트 상황에서 급격하게 변동하며 이에 따라 운전원은 잘못된 판단을 할 수 있다. 이러한 잘못된 판단은 원전의 상태와 직결되기 때문에 운전원은 심리적으로 막중한 부담감을 느낄 수 있다. 또한 주 제어실에는 5명의 운전원이 있으며, 모니터링과 제어장치 조작 시 다른 운전원으로부터 정보를 공유하고 수집하며 운영을 계획한다. 이러한 운전원의 소통 오류로 인해 잘못된 정보를 인식할 수 있다. 이러한 인적오류는 원전의 상태를 악화시킬 수 있으며, 복구가 불가능해지면 원자로를 정지시켜 비상 상태를 초래할 수 있다. 정상적인 원자로 정지는 오랜 시간에 걸쳐 계획된 절차를 통해 수행된다. 본 연구에서는 계획되지 않은 원자로 정지를 트립(Reactor Trip)으로 정의한다. 원자로 트립으로 인한 갑작스러운 가동 중단은 원전 내 수많은 시설의 고장을 초래할 수 있다. 또한, 한 번 정지된 원전은 재가동하는 데 오랜 시간이 소요되며 막대한 경제적 손실을 초래할 수 있다.

따라서 원자력 분야에서는 원전의 무결성을 유지하고 운전원의 잠재적인 인적 오류를 방지하는 것을 목표로 인공지능을 활용한 운전 지원 시스템에 대한 연구가 활발하게 수행되고 있다. 운전 지원 시스템에 관해 주로 수행되는 연구로는 원전의 이상 탐지, 진단, 예측 등이 있다. 관련 연구로는 Yoo 등[1]은 Support Vector Classification을 활용한 중대사고 상황에서 사건 식별, Cascaded Fuzzy Neural Network를 활용한 격납건물 내의 수소 농도 예측, 냉각재 상실 사고에서 파단 크기 예측 등을 수행하며 중대사고 상황에서 운전원의 의사결정 지원을 목적으로 스마트 자원 시스템을 구현하였다. Kim 등[2]은 Long Short-Term Memory와 Variational Autoencoder 방법을 이용하여 높은 신뢰성의 비정상 상황 진단을 목표로 알고리즘을 개발하였다. Park 등[3]은 인공지능을 활용하여 비정상 상황을 진단하며, 인공지능의 신뢰성 문제를 해결하기 위해 설명 가능한 인공지능(XAI; eXplainable Artificial Intelligence)을 활용하는 신뢰할 수 있는 지능형 진단 시스템을 제안하였다. Kim 등[4]은 원자력 발전소의 비상 상황에

서 안전 변수들의 장기간 예측을 수행하고 불확실성 정보를 제공하는 알고리즘을 제안하였다. Jo 등[5]은 원자로 트립을 유발하는 변수(트립변수)의 추세를 예측하며 트립이 발생하는 시기를 예측하는 모델을 개발하였다.

관련 연구 중 예측 분야에는 가압기 압력, 가압기 수위, 증기발생기 수위 등과 같은 주요 매개변수의 추세를 예측하는 것이 주요 접근법으로 활용되었다. 예측된 주요 매개변수에서 관찰된 추세는 운전원이 작업을 효과적으로 계획하고 사전 조치에 대한 지원을 제공하여 운전 안전을 향상시킬 수 있다. 또한, 트립변수의 추세를 예측하고, 그 변수의 추세가 트립 설정치에 도달하는 시간을 계산할 수 있으며 운전원은 원전의 상태를 이해하고 적절한 조치를 취할 수 있다.

그러나 트립변수의 추세를 예측하여 트립 시점을 예측하는 데에는 한계가 존재한다. 첫 번째로 다양한 트립 변수의 모든 추세를 예측해야 하며, 높은 성능이 보장되어야 한다. 두 번째로 예측하는 시간 내에서만 트립 발생 결과를 얻을 수 있다. 트립 변수의 예측 결과가 설정치에 도달하지 못하면 예상 트립 시간을 확인하기에 어려움이 있다. 예를 들어 100초 이후를 예측하는 기능은 100초 이내에 트립이 발생하는 경우에만 트립 시점을 판단할 수 있다. 세 번째로 인공지능 예측 결과에 대한 근거를 확인할 수 없다. 예측 모델은 구조가 복잡한 딥러닝 기법이 주로 활용되며 이에 따라 인공지능의 판단에 대한 근거를 확인할 수 없는 블랙박스 특성이 있다. 원자력 발전소에서는 의사 결정 책임이 여전히 운전원에게 있으며, 인공지능의 블랙박스 특성은 운전원의 판단을 더욱 어렵게 할 수 있다.

따라서 본 연구에서는 위와 같은 한계점을 극복하기 위해 그림 1과 같이 잔여 수명 예측의 개념을 활용하여 잔여 트립 시간 예측 시스템을 개발한다. 잔여 트립 시간 예측 시스템은 높은 성능을 목표로 인공지능 방법들을 활용한 알고리즘을 통해 구현되며, 알고리즘에는 인공지능의 블랙박스 특성을 해소하기 위해 XAI를 활용한 기능이 포함된다. 또한 알고리즘의 결과를 효율적으로 제공하기 위해 잔여 트립 시간 예측 시스템의 인터페이스를 설계한다.

잔여 트립 시간 예측 시스템을 통해 트립이 발생하는 시기를 제공할 수 있으며, 원자력 발전소의 상태에 따라 잔여 트립 시간 예측값이 변동된다. 따라서 잔여 트립 시간 예측 시스템의 결과는 원자력 발전소의 상태를 나타내는 지표로 활용될 수 있으며, XAI를 통해 도출한 인공지능의 예측 결과에 대한 근거를 해석하여 원자력 발전소 상

태에 대한 원인을 파악할 수 있을 것으로 기대된다.

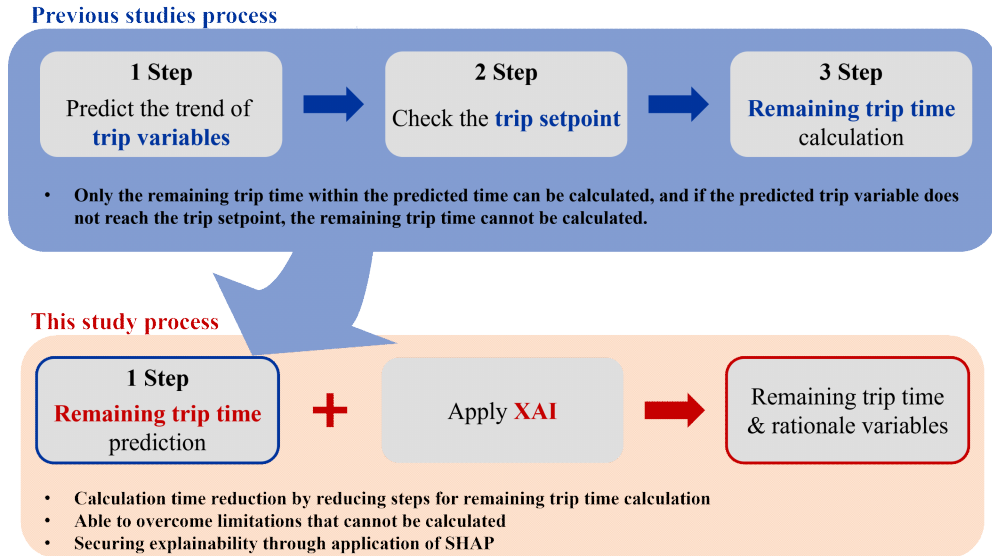


그림 1. 잔여 트립 시간 예측 개념도

제 2 장 인공지능 방법론

인공지능 방법들은 매우 다양하며 각각의 특성이 다르다. 본 연구에서는 잔여 트립 시간 예측 시스템의 알고리즘을 구성하는 기능들의 목적을 고려하여 인공지능 방법을 선택하였다. 알고리즘에는 Autoencoder (AE), Light Gradient Boosting Machine (LightGBM) 인공지능 방법이 활용되며, XAI 방법으로 Shapely Additive Explanation (SHAP) 방법이 활용된다. 잔여 트립 시간 예측 시스템의 알고리즘은 4장에서 설명되며, 해당 장에서는 알고리즘에 활용되는 인공지능 및 XAI 방법에 대해 설명하고자 한다.

제 1 절 AE

AE는 1980년대에 처음 소개되어 많은 분야에서 널리 활용되고 있는 비지도 학습 방법이다[6]. AE는 입력된 데이터와 동일한 데이터가 재구성되도록 학습하는 방법이며, 그림 2과 같이 입력층, 인코더, 잠재 벡터, 디코더, 출력층으로 구성된다. 또한 입력 데이터의 차원을 축소하고 원래 차원으로 복원하기 위해 대칭적인 구조로 되어 있는 것이 특징이다. 이러한 차원 축소 및 복원은 데이터가 재구성될 때 입력 데이터가 복사되는 것을 방지하기 위함이다. 따라서 재구성된 데이터는 아래 식 (1)과 같이 입력 데이터와 차이가 발생하게 되며, 이를 재구성 오차라고 한다. AE는 이러한 재구성 오차가 작아지는 방향으로 학습된다. 따라서 AE는 학습된 데이터가 입력되었을 때는 재구성 오차가 작게 복원되며, 반대로 학습되지 않은 데이터가 입력되었을 때는 제대로 축소 및 복원이 되지 않아 높은 재구성 오차가 계산된다. 이러한 특징을 통해 이상 탐지 및 Unknown 데이터 탐지 분야에서 주로 활용된다. 이상 탐지 분야에서는 정상 데이터만을 학습시키며, 이에 따라 정상이 아닌 데이터 즉, 이상 데이터가 입력되었을 때 탐지가 가능하다. 또한 수집된 데이터의 학습을 통해 수집되지 않은 데이터 즉, Unknown 데이터의 탐지가 가능하다.

$$\text{Reconstruction Error}(x, x') = \|x - x'\|^2 \quad (1)$$

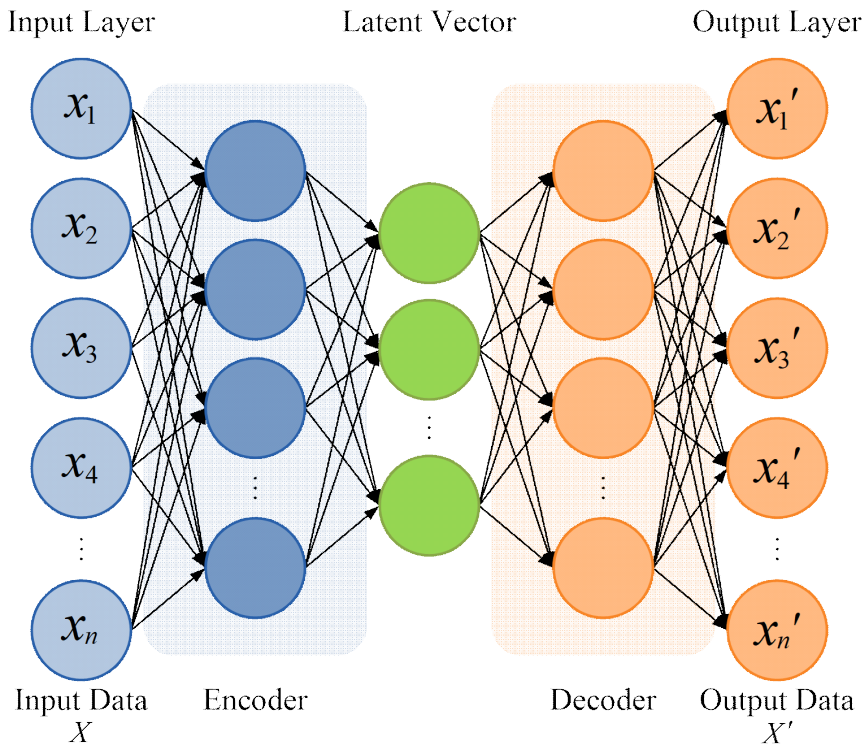


그림 2. AE 구조 개요[6]

제 2 절 LightGBM

LightGBM은 정확도 및 해석 가능성이 높아 다양한 분야에서 좋은 성능을 보이는 GBM 프레임워크를 사용하는 트리 기반 학습 알고리즘이다[7]. 트리 기반의 모델은 그림 3과 같이 높은 설명 가능성을 보이나 일반적으로 딥러닝 방법에 비해 성능이 낮다고 알려져 있다[8]. 기존의 GBM에서는 가능한 모든 트리 분할 지점의 정보 이득을 추정하기 위해 모든 기능에 대한 인스턴스가 계산된다. 이는 결과적으로 학습 시간이 오래 걸리며 계산 비용이 많이 발생하는 문제가 있다. LightGBM은 이러한 문제를 개선하기 위해 Gradient-Based One Side Sampling (GOSS) 와 Exclusive Feature Bundling (EFB) 두 기술을 활용한다. GOSS는 기울기가 큰 인스턴스를 고정으로 선택

하고 기울기가 작은 인스턴스를 무작위로 제외한다. 다시 말해 GOSS는 기울기를 기반으로 데이터를 샘플링하여 학습 효율성을 높이는 기능이다. EFB는 상호 배타적인 특징을 그룹화하여 특징 크기를 줄이는 기술이다. 따라서 EFB는 모델의 복잡성을 줄이고 학습 시간을 단축할 수 있다. 또한 LightGBM은 기존의 Level-wise 트리 분할 방법이 아닌 Leaf-wise 트리 분할 방법을 사용한다. Leaf-wise 트리 분할 방법은 트리의 균형을 고려하지 않으며, 기울기가 가장 큰 Leaf를 중심으로 연속적으로 분할되는 방법이다. 이는 기존의 Leaf-wise 트리 분할 방법에 비해 예측 오차의 손실을 최소화하고 계산 속도를 높일 수 있는 장점이 있다. 그림 4는 Level-wise와 Leaf-wise 트리 분할 방법의 구조를 보여준다.

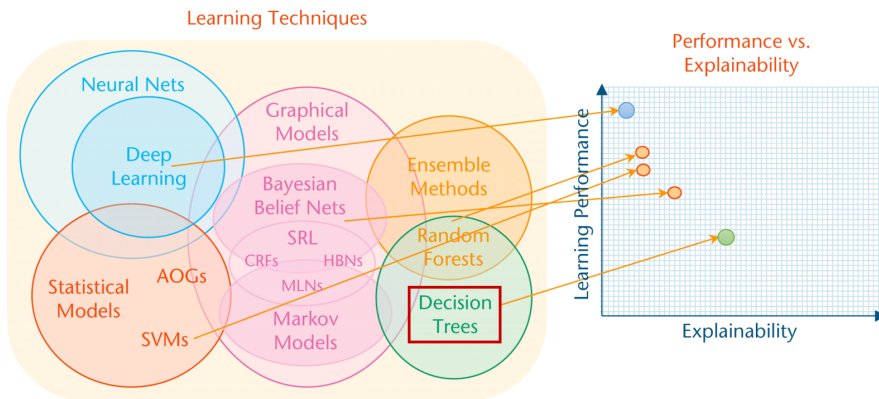


그림 3. 인공지능 방법들에 대한 설명 가능성 및 성능 비교[8]

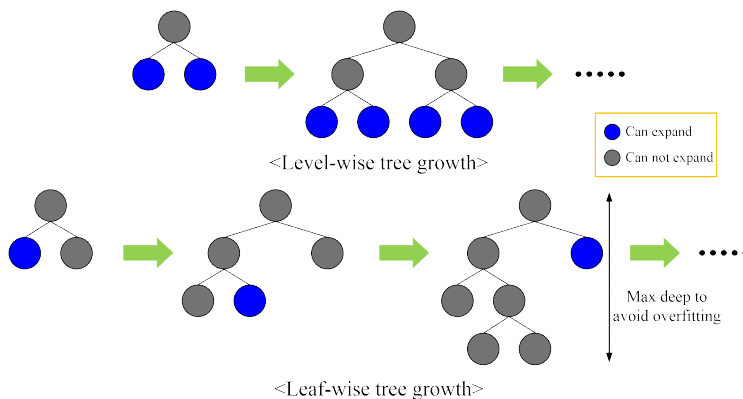


그림 4. Level-wise와 Leaf-wise 트리 분할 방법의 구조[9]

제 3 절 SHAP

SHAP은 Lloyd Shapley가 만든 이론 위에 변수 사이의 독립성을 근거로 덧셈이 가능하도록 활용도를 넓힌 방법으로 논문을 통해 공개되었다[10]. 다시 말해 SHAP은 모델 출력의 원인에 대한 Shapley 값을 계산하는 방법이다. Shapley 값은 게임 이론을 바탕으로 모델 출력에 대한 모든 변수의 조합을 통해 계산된 각 변수의 기여도이다. 이는 모델 결과에 대해 각 변수가 얼마나 기여했는지 수치상으로 계산할 수 있다. Shapley 값의 계산식은 식 (2)와 같이 표현되며, 각 변수의 기여도가 제외되었을 때 전체 결과에 미치는 변화를 계산한다.

$$\phi_i(f) = \sum_{P \in N \setminus \{i\}} \frac{|P|!(n - |P| - 1)!}{n!} (f(P \cup \{i\}) - \nu(P)) \quad (2)$$

이때, ϕ_i 는 i 변수에 대한 Shapley 값이며, n 은 전체 변수 개수, P 는 전체 그룹에서 i 변수를 제외한 모든 집합, $|P|$ 는 P 의 원소 개수, $f(P)$ 는 i 변수를 제외하고 나머지 부분 집합이 결과에 공헌한 기여도, $f(P \cup \{i\})$ 는 i 변수를 포함한 전체 기여도를 의미한다. 즉, i 변수가 기여하는 정도는 전체 기여도에서 i 변수가 제외된 기여도의 합을 뺀 값이다.

Shapley 값은 모든 순열 조합에 대해 계산되기 때문에 시간이 오래 걸리는 단점이 있다. 이러한 단점을 개선하기 위해 각 모델의 특성에 맞게 KernelSHAP, DeepSHAP, TreeSHAP이 개발되었다. 본 연구에서는 2절의 LightGBM 모델에 SHAP을 적용하며, 이에 따라 TreeSHAP이 활용되었다. TreeSHAP은 조건부 평균을 사용하는 트리 기반 인공지능에 특화된 방법으로 빠른 계산 속도를 나타낸다. 그림 5는 SHAP을 적용한 간단한 예시를 보여준다. SHAP 결과에는 모델의 예측 결과($f(x)$), 예측값의 평균(base value), Shapley 값 (ϕ_i)이 나타난다. 화살표는 Base Value를 기준으로 밀어서 늘리거나 줄이려는 성질을 나타낸다. 또한 화살표의 길이는 Shapley 값의 크기로 기여 정도를 보여준다. 이러한 결과를 통해 SHAP은 예측 결과에 대한 부정적 기여도와 긍정적 기여도를 보여 대조 설명이 가능하다는 특징이 있다. 본 논문에서는 긍정적 기여도를 Contribution, 반대로 부정적 기여도를 Offsetting으로 명칭한다.

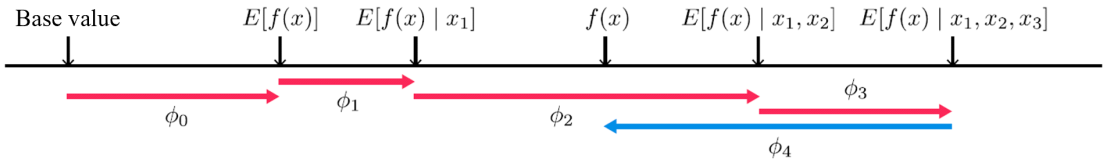


그림 5. SHAP 적용 결과 예시[10]

제 3 장 비정상 시나리오 데이터

본 연구의 잔여 트립 시간 예측 시스템은 다양한 인공지능 방법들이 활용된다. 인공지능 모델의 학습을 위해 원자력 발전소의 운전 데이터가 필요하다. 이에 해당 장에서는 인공지능 훈련 및 테스트를 위한 정상 데이터와 비정상 시나리오의 데이터를 수집하고 수집된 데이터를 전처리하는 과정에 대해 다루고자 한다.

제 1 절 데이터 수집

실제 원자력 발전소에서 정상 상태 및 다양한 비정상 상황에 대한 운전 데이터의 수집은 정보 보안 및 원자력 발전소의 건전성 유지 등에 의해 제한된다. 이에 본 연구에서는 원자력 발전소의 운전 데이터를 대처하기 위해 웨스팅하우스 993 MWe 3 Loop 가압 경수로를 기반한 고리 원자력 발전소 3, 4호기의 설계를 참고하여 교육을 목적으로 한국원자력연구원에서 개발된 Compact Nuclear Simulator (CNS)를 활용하여 정상 상태 및 비정상 시나리오 데이터를 수집하였다[11]. 그림 6은 CNS 인터페이스를 나타낸다. CNS는 간단한 출력 운전을 비롯한 오작동 주입을 통해 다양한 비정상 시나리오를 모사할 수 있으며 밸브, 펌프, 히터 등을 통해 제어할 수 있다.

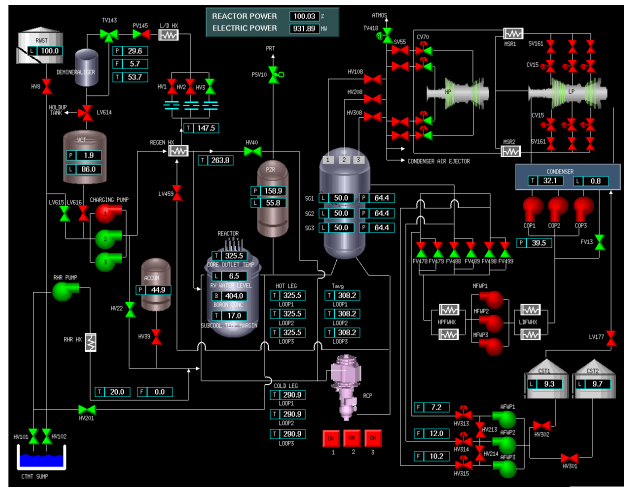


그림 6. CNS 인터페이스 개요

본 연구에서는 정상 상태 및 계측기 오류, 기기 상태 이상, 배관 누설에 대한 비정상 시나리오 데이터를 수집하였다. 수집된 비정상 데이터는 정상 상태에서 시작하여 30초에 Malfunction이 주입된다. 또한 잔여 트립 시간 예측 모델의 학습에 있어서 트립이 발생하는 시점에 대한 정보는 필수적이다. 이에 따라 비정상 시나리오 데이터를 수집함에 있어서 운전원의 조치는 포함되지 않는다. 운전원의 조치에 의해 트립이 발생되지 않으면 인공지능 학습이 불가능하기 때문이다. 또한 데이터는 트립이 발생한 시점까지 데이터를 수집하였다. 그림 7은 수집된 데이터의 형태를 간략하게 보여준다.

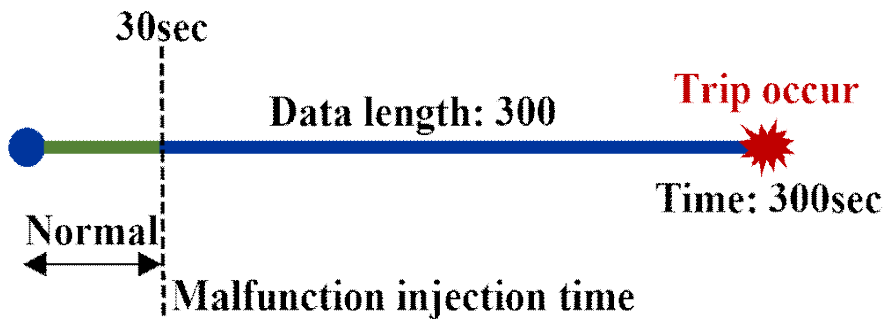


그림 7. 수집된 비정상 데이터의 형태

수집된 비정상 시나리오 데이터 중 트립이 발생하지 않는 시나리오에 관한 데이터는 제외되었다. 운전원의 조치가 수행되지 않았음에도 불구하고 트립이 발생되지 않은 이유는 CNS의 자동 조치로 인해 원자로의 상태가 완화되기 때문이다. 예를 들어 가압기 압력 채널 고장 ‘저’ 시나리오의 경우 압력 채널의 고장으로 인해 저압력을 지시하게 되며 실제 압력이 증가하는 시나리오다. 그러나 실제 압력이 증가함에 따라 Pressurizer Power-Operated Relief Valve (PORV)가 자동으로 개방되며 실제 압력이 회복된다. 또한 트립이 발생하는 시나리오는 자동 조치에 의한 복구가 실패되는 시나리오이며, 대부분 1,800초 이내에 트립이 발생된다. 따라서 본 연구에서는 1,800초를 기준으로 트립이 발생하는 시나리오를 구분하였다. 추가적으로 잔여 트립 시간 예측 시스템에 구성되는 인공지능의 훈련 및 테스트를 위해 수집된 데이터는 케이스에 따라 8대2의 비율로 훈련 데이터와 테스트 데이터를 분할하여 활용된다. 최종적으로 본 연구에 활용되는 비정상 시나리오 정보는 표 1에 나타낸다.

표 1. 수집된 비정상 시나리오 정보 목록

NO.	비정상 시나리오 번호	비정상 시나리오 이름	데이터 개수 (훈련, 테스트)	전체 데이터 행렬 (행, 열)
0	Normal	정상 상태	20(15, 5)	(8604, 2222)
<자동로직 및 계측기 오류>				
1	Ab21-01	가압기 압력 채널 고장 '고'	18(13, 5)	(4698, 2222)
2	Ab20-04	가압기 수위 채널 고장 '저'	15(12, 3)	(7954, 2222)
3	Ab15-08	증기발생기 수위 채널 고장 '고'	40(35, 5)	(11384, 2222)
<기기 상태 이상>				
4	Ab21-12	가압기 PORV 고장 '열림'	52(45, 7)	(13573, 2222)
5	Ab19-02	가압기 안전밸브 고장	51(45, 6)	(17370, 2222)
6	Ab21-11	가압기 살수밸브 고장 '열림'	50(45, 5)	(31391, 2222)
<배관 누설>				
7	Ab23-01	*RCS → *CCW 누설	30(25, 5)	(2900, 2222)
8	Ab23-06	증기발생기 전열관 누설	36(30, 6)	(2738, 2222)

*RCS: Reactor Coolant System

*CCW: Component Cooling Water

제 2 절 데이터 전처리

일반적으로 인공지능 학습 시 편향으로 인한 성능저하를 방지하기 위해 전처리가 수행된다. 전처리에는 변수 선택 및 정규화가 수행된다. 수집된 CNS 데이터는 2,222개의 변수로 구성되어 있다. 인공지능의 학습 과정에서 높은 성능을 위해 충분히 많은 수의 변수를 활용하는 것이 중요하다. 그러나 불필요한 변수를 무분별하게 사용하게 되면 인공지능 학습 과정에서 각 변수의 영향이 제대로 반영되지 않을 수 있다. 다시 말해 불필요한 변수의 존재는 인공지능 모델의 불확실성을 증가시켜 인공지능의 전반적인 성능을 저하시킬 수 있다[12]. 이에 따라 본 논문에서는 비정상 운전 절차서(Abnormal Operating Procedure)를 참조하고, 비정상 시나리오의 증상 및 변수의 상관관계 분석을 통해 152개의 변수를 선정하였다.

또한 수집된 CNS의 변수는 각각 다른 범위에 위치된다. 예를 들어, 밸브 개도 상태 변수는 0(닫힘)에서 1(100% 열림)까지 인 반면에 가압기 압력 변수는 약 15,000,000 Pa이므로 각각의 변수 크기에 상당한 차이가 있다. 이러한 큰 차이의 데이터 분포에서 인공지능은 더 큰 값의 변수에 치중되어 학습될 수 있다[13]. 이를 방지하기 위해 수집된 데이터의 최솟값과 최댓값을 기준으로 정규화를 수행하는 Min-Max 정규화 방법을 사용하였다. Min-Max 정규화 방법은 식 (3)과 같이 표현된다. 이때, X_{\min} 은 수집된 데이터의 152개의 변수 각각의 최솟값이며, X_{\max} 는 최댓값이다. 정규화를 통해 데이터의 범위를 0과 1 사이로 변환한다.

$$X_{scale} = (X - X_{min}) / (X_{max} - X_{min}) \quad (3)$$

제 4 장 잔여 트립 시간 예측 시스템

본 연구에서 잔여 트립 시간 예측 시스템은 크게 3장 2절에서의 데이터 전처리를 포함하여, 혼란 상태 진단 기능, 비정상 시나리오 진단 기능, 잔여 트립 시간 예측 기능, 잔여 트립 시간 예측 결과 해석 기능으로 구성된 알고리즘을 기반으로 개발되었다. 해당 장에서는 개발된 잔여 트립 시간 예측 알고리즘을 설명하며, 개발된 각 기능의 결과를 보여주고자 한다.

제 1 절 잔여 트립 시간 예측 알고리즘

잔여 트립 시간 예측 알고리즘은 그림 8과 같이 1) 데이터 전처리 기능, 2) 혼란 상태 진단 기능, 3) 비정상 시나리오 진단 기능, 4) 잔여 트립 시간 예측 기능, 5) 잔여 트립 시간 예측 결과 해석 기능으로 구성되며, 순차적으로 작동한다.

혼란 상태 진단 기능, 비정상 시나리오 진단 기능, 잔여 트립 시간 예측 기능은 데이터 전처리 기능을 통해 변환한 데이터가 활용된다. 데이터 전처리 기능은 3장 2절에서 설명된다. 혼란 상태 진단 기능은 입력된 데이터가 기존에 수집된 정상 및 비정상 시나리오와 다른 분포의 데이터를 탐지하여 Unknown 데이터를 판단하는 기능이다. 이는 이후에 수행되는 비정상 시나리오 진단 기능의 오진단을 방지하기 위해 활용되며, 2절에서 자세히 설명된다. 비정상 시나리오 진단 기능은 원전의 상태에 대한 진단 정보를 제공하기 위한 기능이며, 3절에서 자세히 설명된다. 잔여 트립 시간 예측 기능은 트립까지 남은 시간에 대한 예측 정보를 제공하기 위한 기능이다. 앞서 비정상 시나리오 진단 기능의 결과를 기반으로 예측 모델이 활용되며, 4절에서 자세히 설명된다. 잔여 트립 시간 예측 결과 해석 기능은 XAI 활용하여 잔여 트립 시간 예측 결과에 대한 변수들의 기여도를 해석하여, 인공지능이 예측 결과를 도출함에 있어서 중점적으로 보는 변수들의 정보를 확인한다. 이는 인공지능의 신뢰성 향상과 예측 결과에 대한 추가적인 정보제공을 위한 기능이며, 5절에서 자세히 설명된다.

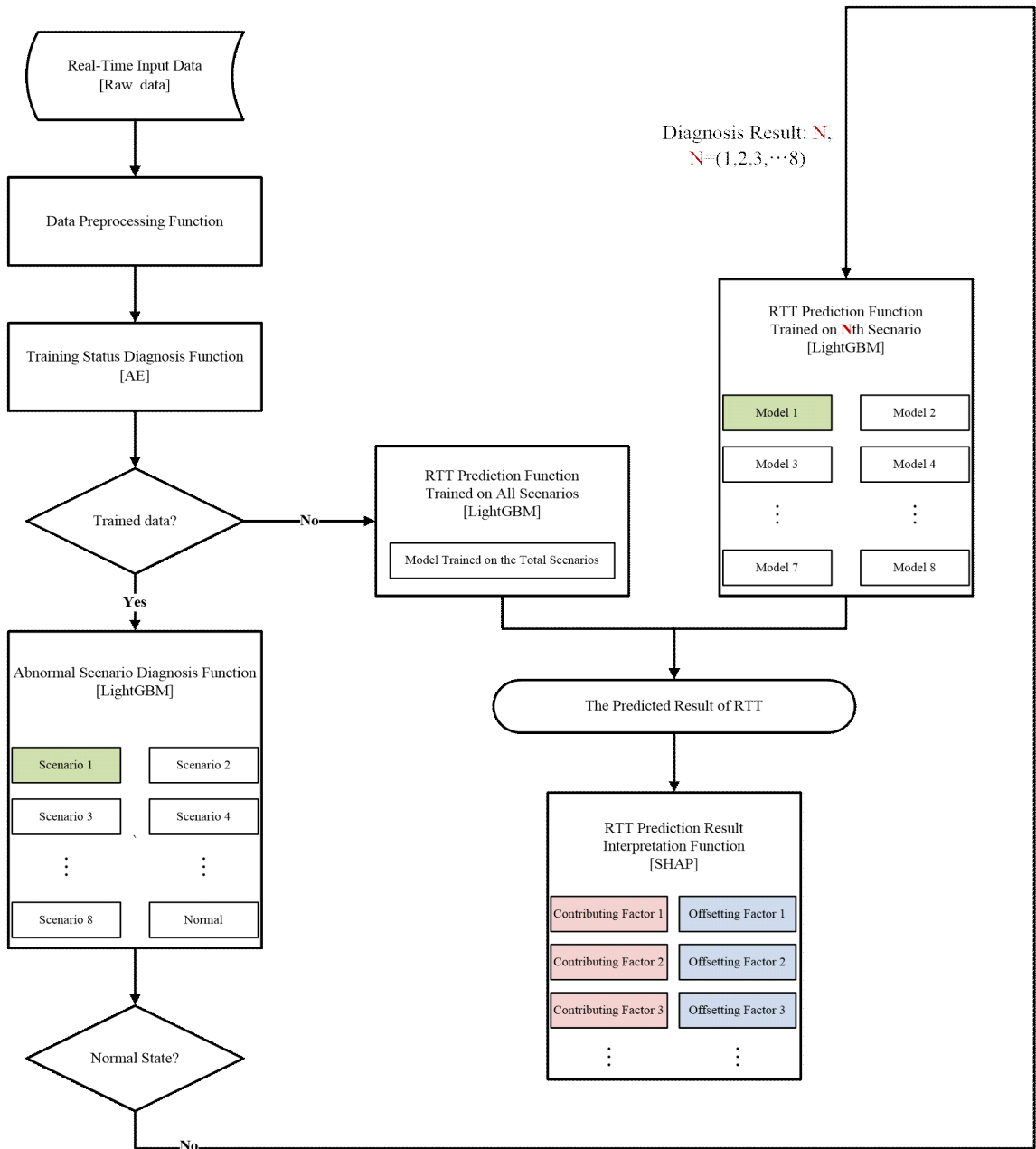


그림 8. 잔여 트립 시간 예측 알고리즘 개요

제 2 절 훈련 상태 진단 기능

훈련 상태 진단 기능은 입력된 데이터가 이전에 학습된 분포의 데이터인지에 대한 여부를 판단하기 위해 수행된다. 이는 그림 8에서 훈련 상태 진단 기능 이후에 수행되는 비정상 시나리오 진단 기능에서 오진단을 방지하기 위함이다. 비정상 시나리오 진단 기능은 지도학습 기반의 인공지능 방법이 활용된다. 지도학습은 인공지능이 훈련 데이터를 기반으로 정답을 맞히는 학습 방법이다. 정답은 사전에 사용자가 훈련 데이터에 라벨링(Labeling) 작업을 통해 설정된다. 이에 따라 인공지능은 사전에 설정된 정답(Label)에 대한 데이터만 학습된다. 따라서 모르는 정답에 있어서는 ‘모른다’를 제공할 수 없으며 학습된 정답 내에서 가장 가능성이 높은 정답을 제공한다[14]. 다시 말해 모르는 시나리오의 데이터가 입력되어도 학습된 시나리오 중 가장 비슷한 시나리오로 진단하게 되며 잘못된 진단 정보를 제공하게 된다. 이러한 잘못된 진단 정보는 운전원의 판단을 더욱 어렵게 만들며 수행되는 알고리즘의 결과에 심각한 영향을 줄 수 있다. 훈련 상태 진단 기능은 학습되지 않은 데이터(Untrained data)를 탐지하는 기능이다. 따라서 학습 상태 진단 기능에서 Untrained data로 진단된 경우에는 비정상 시나리오 진단 기능은 동작하지 않으며, 운전원의 판단으로 진단이 수행된다.

훈련 상태 진단 기능은 비지도 학습 기반의 인공지능 방법인 AE를 사용한다. AE의 훈련 데이터로 수집된 8가지 비정상 시나리오 데이터와 정상 상태 데이터를 사용한다. AE는 수집된 모든 데이터의 분포를 학습하게 되며 Untrained data를 입력받게 되면 큰 재구성 오차를 나타낸다. 따라서 재구성 오차가 설정된 임계값(Threshold)을 넘게 되면 Untrained data를 진단할 수 있다. 여기서 재구성 오차에 대한 Threshold 값은 훈련 데이터의 재구성 오차의 평균과 편차를 활용하여 시그마 규칙을 통해 설정되었다. 시그마 규칙에는 그림 9와 같이 다양한 신뢰도 구간의 규칙이 있다. 본 연구에서는 식 (4)의 2-시그마 규칙을 활용하여 Threshold 값을 설정하였다. 2-시그마 규칙은 95.45%의 신뢰구간을 제공한다. 즉, 데이터의 약 4.55%가 학습된 범위에 속하는 데이터임에도 불구하고 Threshold 값을 초과하여 학습되지 않은 데이터로 진단될 수 있다. 그러나 이 접근 방식은 훈련되지 않은 데이터에 대해 더욱 민감하게 반응할 수 있다.

$$Threshold = \mu \pm 2\sigma \quad (4)$$

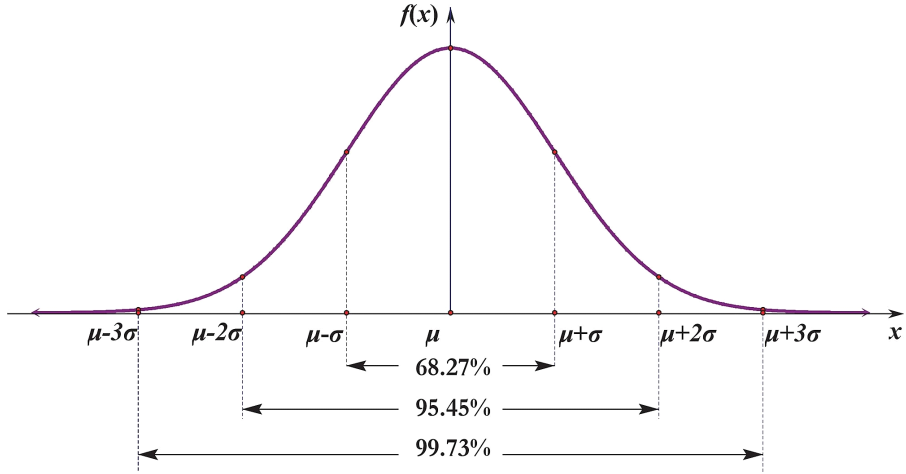


그림 9. 시그마 규칙에 따른 신뢰구간

훈련 상태 진단 모델의 최적의 성능을 위해 하이퍼파라미터 튜닝을 통한 최적화가 수행되었다. 하이퍼파라미터 최적화에는 AE의 Layer 수, Batch Size, Hidden Layer에 대한 활성화 함수의 조절이 포함된다. Layer 수는 AE의 구조적 특성으로 인해 대칭적이어야 한다. 따라서 3, 5, 7개의 Layer가 사용되었다. 또한, Layer 당 노드의 개수는 비례적으로 줄어들며, 152개의 변수(차원)를 사용하기 때문에 2배수로 줄어든다. 예를 들어, 5개 Layer의 경우 노드는 76, 38, 19, 38, 76으로 구성되었다. 최적화에 대한 결과를 비교하기 위해 테스트 데이터에 대한 성능 평가 지표로 평균 제곱 오차(MSE; Mean Squared Error)를 사용했다. 식 (5)는 MSE를 계산하는 식을 나타낸다. 이때, X_i 는 입력 데이터, X'_i 는 재구성된 데이터(AE 출력값), N 은 입력 데이터의 전체 행 길이를 의미한다. 표 2는 각 하이퍼파라미터 정보에 대한 결과를 나타내며, MSE가 가장 작은 하이퍼파라미터 조합이 최적의 모델로 되었다.

$$MSE = \frac{1}{N} \sum_{i=1}^N (X_i - X'_i)^2 \quad (5)$$

표 2. AE 하이퍼파라미터 최적화 결과

No.	레이어 수	Batch size	활성화 함수	테스트 데이터 MSE
1	3	32	Relu	0.0077
2	3	64	Relu	0.0073
3	3	128	Relu	0.0080
4	3	32	Sigmoid	0.0092
5	3	64	Sigmoid	0.0076
6	3	128	Sigmoid	0.0086
7	5	32	Relu	0.0088
8	5	64	Relu	0.009
9	5	128	Relu	0.009
10	5	32	Sigmoid	0.0094
11	5	64	Sigmoid	0.0104
12	5	128	Sigmoid	0.0107
13	7	32	Relu	0.0104
14	7	64	Relu	0.0107
15	7	128	Relu	0.0114
16	7	32	Sigmoid	0.0173
17	7	64	Sigmoid	0.0120
18	7	128	Sigmoid	0.0135

제 3 절 비정상 시나리오 진단 기능

비정상 시나리오 기능은 그림 8과 같이 훈련 상태 진단 기능 결과가 훈련된 상태 조건일 경우에만 동작한다. 데이터 전처리를 통해 정규화된 데이터를 입력으로 받아 현재 원전의 상태에 대한 진단 정보를 출력으로 제공한다. 출력값은 0부터 8까지 숫자로 나타내며 이는 표 1의 시나리오 번호와 동일하다. 이러한 출력값은 4절의 잔여 트립 시간 예측 기능에 연결되어 활용되기 때문에 높은 정확도를 요구한다.

해당 기능의 진단 모델은 LightGBM을 활용하여 구현하였다. 최적의 진단 모델을 위해 LightGBM의 하이퍼파라미터를 경험적으로 조정하여 최적화하였다. 최적의 모델을 선택하기 위해 성능 평가 지표로 테스트 데이터에 대한 정확도를 나타내었다. 정확도

는 실제 정답과 인공지능 예측값이 일치하는 정도의 비율로 식 (6)을 통해 계산된다. 그림 10은 테스트 데이터에 대한 진단 결과를 Confusion Matrix로 나타낸 결과이다. Confusion Matrix에서 각 축의 숫자는 각 시나리오의 라벨 번호에 해당하며, x축은 진단 모델의 결과를 나타내고, y축은 실제 정답을 나타낸다. 진단 결과는 간단한 최적화 작업을 통해 모든 시나리오에서 100% 정확도를 보였다. 이는 특징이 분명한 원전의 정상 상태 및 8가지 비정상 시나리오를 분류하는 비교적 간단한 문제이기 때문으로 판단된다.

$$Accuracy\ score = \frac{number\ of\ correctly\ prediction\ data}{number\ of\ total\ data} \quad (6)$$

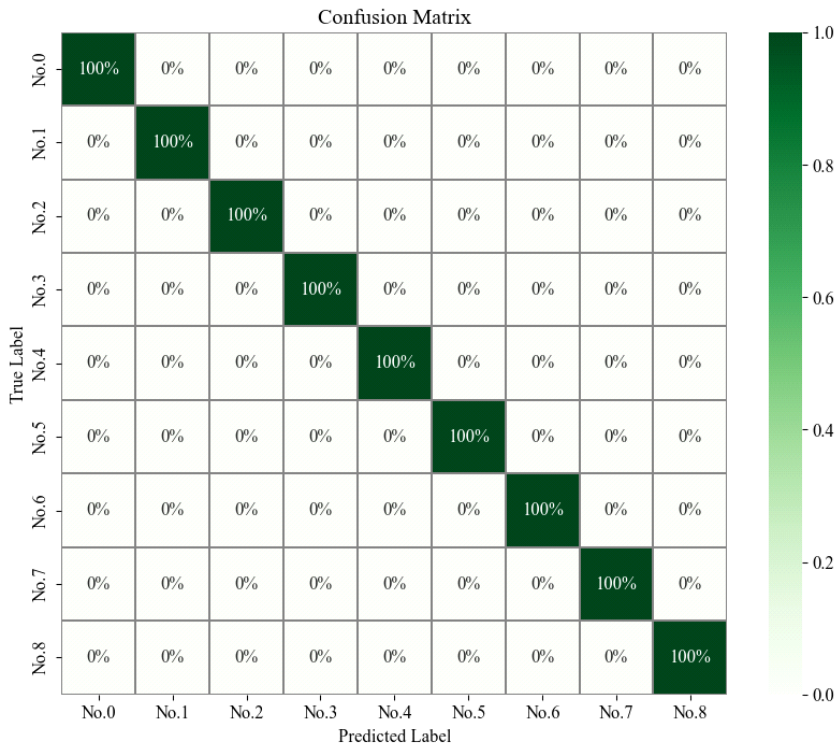


그림 10. 비정상 시나리오 진단 기능 테스트 결과

제 4 절 잔여 트립 시간 예측 기능

본 연구에서는 잔여 트립 시간 예측 성능뿐만 아니라 인공지능의 신뢰성 또한 중요하게 간주한다. 따라서 잔여 트립 시간 예측 기능은 LightGBM이 활용된다. LightGBM은 잔여 트립 시간 예측 결과 해석 기능의 XAI 방법을 효과적으로 활용할 수 있으며, 많은 분야에서 높은 예측 성능을 보이기 때문이다. 또한 그림 8과 같이 잔여 트립 시간 예측 모델은 단일 시나리오 모델과 전체 시나리오 모델로 구분하여 구현된다. 단일 시나리오 모델은 수집된 비정상 시나리오 각각에 대한 예측 모델이 개발되며, 전체 시나리오 모델은 수집된 전체 비정상 시나리오 데이터가 한 번에 학습된다. 단일 시나리오 모델의 경우 비정상 시나리오 각각에 대해 잔여 트립 시간이 학습되며, 이는 적은 데이터 분포를 학습하며 더 높은 성능을 기대할 수 있다. 이러한 단일 시나리오 모델을 활용하기 위해서는 원전의 현재 상태에 대한 정보가 필요하다. 이에 따라 비정상 시나리오 진단 모델의 결과를 활용한다. 비정상 시나리오 진단 기능을 통해 제공된 현재 원전의 상태에 따라 단일 시나리오 모델이 선택되어 활용된다. 또한 원자력 발전소에서 발생할 수 있는 비정상 상황은 매우 다양하며, 이에 따라 모르는 비정상 상태가 발생할 수 있다[15]. 모르는 비정상 상태가 발생할 경우 훈련 상태 진단 기능을 통해 Untrained Data를 탐지하며, 전체 시나리오 모델을 통해 잔여 트립 시간이 예측된다. 전체 시나리오 모델은 다양한 데이터 분포를 학습하며, 더 넓은 적용 범위를 가지고 있다. 이러한 이유로 본 연구에서는 단일 시나리오 모델에 더불어 전체 시나리오 모델 또한 개발되었다.

잔여 트립 시간 예측 모델은 수집된 시나리오별로 개별적으로 개발된다. 개발되는 각각의 모델마다 최적화 작업이 수행된다. 최적화 작업에는 LightGBM의 하이퍼파라미터 중 학습률(Learning Rate), 최대 깊이(Max_depth), 잎의 개수(Num_leaves)의 조정을 통해 수행된다. 최적화된 모델의 성능은 일반적으로 회귀모델 평가에 많이 활용되는 평가지표인 평균 제곱근 오차(RMSE; Root Mean Squared Error)를 통해 평가된다. RMSE는 식 (7)을 통해 계산되며, 실제값과 예측값의 차이를 나타내기 때문에 RMSE 값이 작을수록 높은 성능을 의미한다. 최적화된 모델의 하이퍼파라미터 정보는 표 3에 나타난다. 최적화 결과 단일 시나리오 예측 모델이 전체 시나리오 학습 모델보다 전체적으로 높은 예측 성능을 나타내었다. 추가적으로 단일 시나리오 예측 모델과

전체 시나리오 예측 모델의 성능 비교를 위해 가압기 압력 채널 고장 ‘고’ 시나리오의 잔여 트립 시간 예측을 수행하였다. 그림 11은 전체 시나리오 학습 모델의 예측 결과를 보여준다. 그림에서 파란 선은 예측된 결과를 보여주며, 빨간색 선은 실제 잔여 트립 시간을 나타낸다. 전체 시나리오 학습 모델에서는 초반 부분에서 예측 정확도가 떨어지며 시간이 지날수록 안정적인 예측 성능을 보인다. 그림 12는 해당 시나리오에 대한 단일 시나리오 예측 모델의 결과를 나타낸다. 단일 시나리오 예측 모델 결과 전체적으로 5% 오차 공간인 주황색 영역에서 안정적으로 예측이 수행됨을 보인다. 또한 초반 부분에 큰 오차를 보이는 전체 시나리오 예측 모델에 비해 높은 성능을 나타낸다. 전체 시나리오 예측 모델은 성능이 낮은 성능을 보임에도 불구하고 다양한 데이터 분포가 학습되기 때문에 학습되지 않은 데이터에 대한 예측에 활용될 수 있다. 학습되지 않은 데이터에 대한 예측 성능은 5장에서 실험을 통해 확인된다.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_{actual}^i - y_{predicted}^i)^2} \quad (7)$$

여기서, y_{actual}^i 와 $y_{predicted}^i$ 는 각각 실제값과 예측값을 의미하며, N 은 입력 데이터의 전체 행 길이를 의미한다.

표 3. 잔여 트립 시간 예측 모델 최적화 결과

No.	비정상 시나리오 번호	Learning Rate	Max_depth	Num_leaves	테스트 데이터 RMSE
1	Ab21-01	0.01	5	25	0.0025
2	Ab20-04	0.02	7	49	0.0077
3	Ab15-08	0.01	5	25	0.0056
4	Ab21-12	0.02	4	14	0.2369
5	Ab19-02	0.01	6	36	0.0194
6	Ab21-11	0.03	7	49	0.2658
7	Ab23-01	0.01	4	16	0.0856
8	Ab23-06	0.02	7	49	0.0328
All	All	0.01	5	25	0.3049

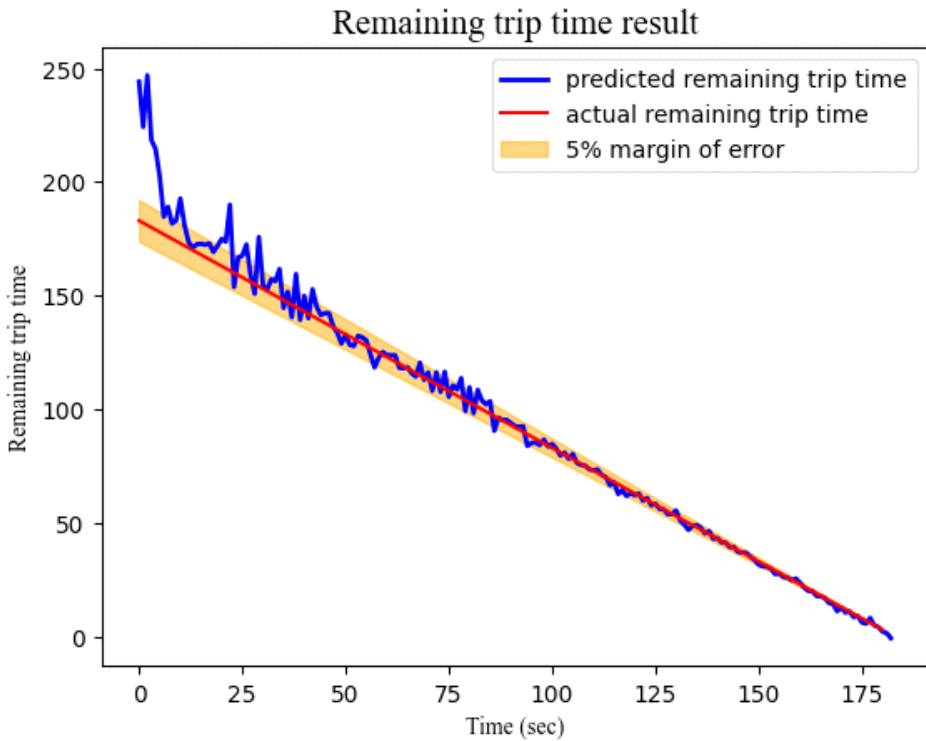


그림 11. 전체 시나리오 모델을 활용한 잔여 트립 시간 예측 결과

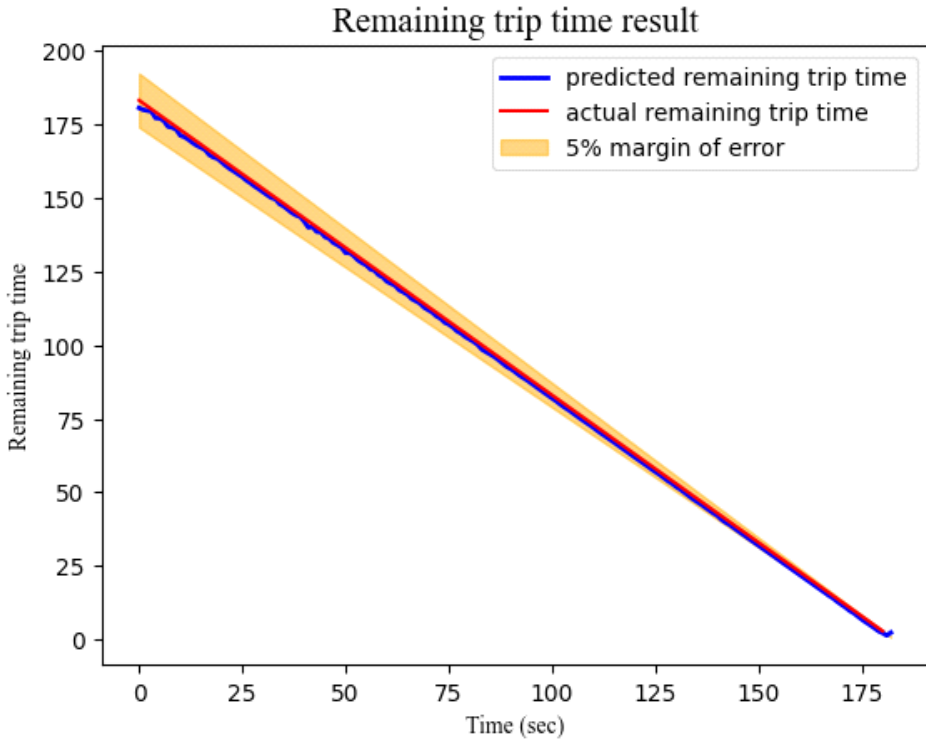


그림 12. 단일 시나리오 모델을 활용한 잔여 트립 시간 예측 결과

제 5 절 잔여 트립 시간 예측 결과 해석 기능

잔여 트립 시간 예측 결과 해석 기능은 XAI 방법인 SHAP을 활용하여 구현되었다. SHAP은 잔여 트립 시간 예측 모델에서 출력된 결과에 영향을 준 변수들에 대한 정보를 제공한다. SHAP의 출력 정보는 예측 결과에 대한 주요 영향 변수 정보와 각 변수들에 대한 기여도를 제공하게 된다.

해당 절에서는 비정상 시나리오 중 가압기 PORV 고장 ‘열림’ 시나리오의 잔여 트립 시간 예측을 수행하며, 예측 결과에 대한 SHAP의 도출 인자들을 확인한다. 우선 SHAP의 도출 인자에 대한 타당성을 확인하기 위해 해당 시나리오의 증상을 알 필요가 있다. 가압기 PORV 고장 ‘열림’ 시나리오는 다음과 같은 증상이 있다. 1) PORV 개방으로 인해 가압기 압력 및 온도가 낮아진다. 2) 낮아진 압력을 회복하기 위해 가압기 히터가 동작한다. 3) 가압기 배압 탱크(PRT)의 수위가 증가한다. 4) 압력 강하로

인해 Hot leg의 온도가 감소한다. 이러한 증상과 SHAP의 결과를 비교하며, 인공지능 예측 결과의 주요 영향 인자들에 대한 타당성을 검토한다.

그림 13은 해당 시나리오에 대한 잔여 트립 시간 예측 결과를 보여준다. 그림에서 빨간색 선은 실제 트립까지 남은 시간이며, 파란색 선은 예측된 잔여 트립 시간에 대한 결과이다. 주황색 영역은 실제 트립까지 남은 시간의 5% 오차 범위를 나타내며, 오차 범위 내에서 예측이 수행된 것을 보인다.

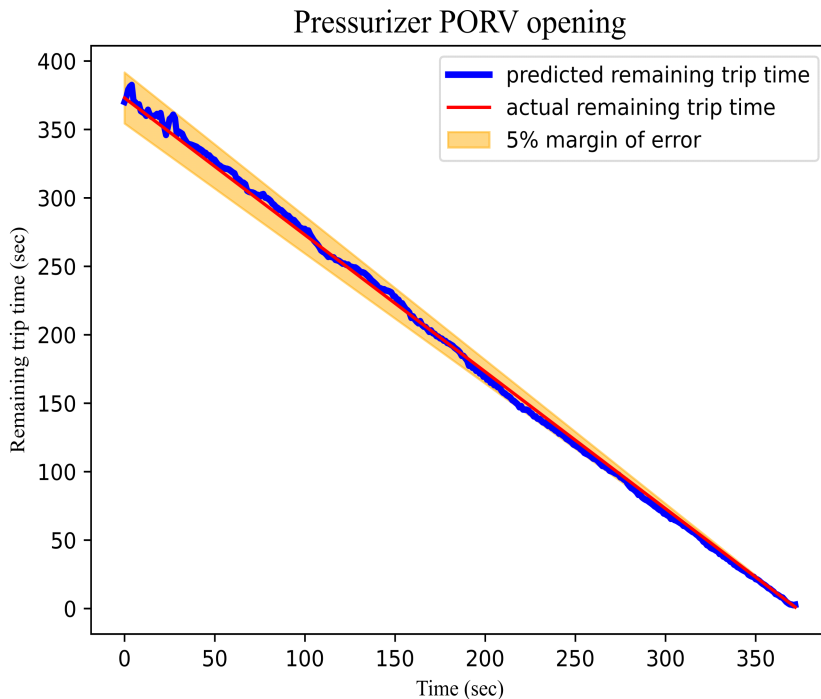


그림 13. 가압기 PORV 고장 ‘열림’ 시나리오에서 잔여 트립 시간 예측 결과

SHAP은 잔여 트립 시간 예측 모델 출력값에 적용되어 결과를 도출한다. 해당 절에서는 예측 모델의 초반, 중반, 후반부(20초, 170초, 350초)의 출력값에 적용하여 결과를 확인한다. SHAP을 적용한 결과는 그림 14~16에 나타내며, 표 4는 SHAP 결과에서 주목할 만한 변수의 정보를 보여준다. 그림 14는 20초 시점에서의 잔여 트립 시간 예측 결과에 대한 주요 영향 인자를 보여준다. 20초 시점에서는 PORV 개도 상태(BPORV), 가압기 온도(UPRZ)와 같은 변수가 잔여 트립 시간 예측값을 낮추는 데 기여한다. 반

대로 Over Temperature Delta-T (UOVER), 가압기 압력(PPRZ) 등은 예측 결과를 높이는 데 기여한다. SHAP 결과를 통해 시나리오 초반에는 PORV의 개방 상태가 원전의 상태를 악화시키는 원인임을 유추할 수 있다. 또한 그림 15는 170초 시점에 대한 SHAP 결과이며, 그림을 통해 Offsetting 변수인 가압기 온도 변수의 기여도가 커지는 것을 보인다. 운전원의 조치가 수행되지 않음에 따라 원전의 상태가 악화되며, 가압기 온도가 정상치보다 많이 낮아지기 때문이다. 마지막으로 그림 16은 350초 시점의 결과를 보여준다. 해당 시나리오에는 376초에 트립이 발생되며, 350초는 원자로 트립 직전을 의미한다. 트립 발생 직전은 초반과 달리 UOVER과 PPRZ 변수가 Offsetting 변수로 변경되었음을 나타내었다. 이러한 결과들을 통해 PORV의 개방 상태가 비정상 상황을 유발한 원인임을 유추할 수 있으며, 시간이 지남에 따라 가압기 압력과 온도가 떨어지는 등 원전 상태의 악화에 대한 인공지능의 근거를 확인할 수 있다.

잔여 트립 시간 예측 결과 해석 기능에서 도출한 결과는 인공지능이 예측 결과를 도출하는 데 영향을 주는 주요 변수들에 대한 정보이다. 해당 절에서는 가압기 PORV 고장 ‘열림’ 시나리오에서 잔여 트립 시간 예측 결과에 대한 영향 인자들이 해당 시나리오의 주요 증상과 일부 일치하는 것을 확인하였다. 따라서 인공지능이 예측 결과를 도출함에 있어 주로 보는 증상들이 원전 상태에 대한 주요 정보라고 판단된다.

표 4. SHAP 결과에 표시되는 변수 정보

변수 이름	변수 설명
BPORV	PORV open state
UPRZ	Pressurizer temperature
UOVER	Over temperature delta temperature
PPRZ	Pressurizer pressure (actual)
ZINST65	Pressurizer pressure (channel)
BHTV	High pressure turbine control valve position
CIODMPC	Average iodine concentration

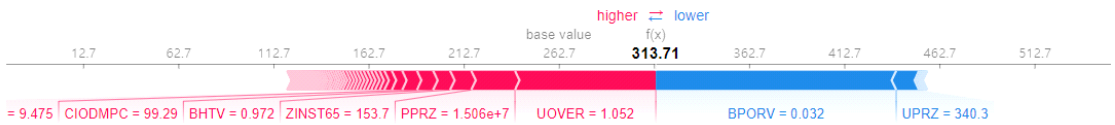


그림 14. 가압기 PORV 고장 ‘열림’ 시나리오 초반부 예측값에 SHAP을 적용한 결과

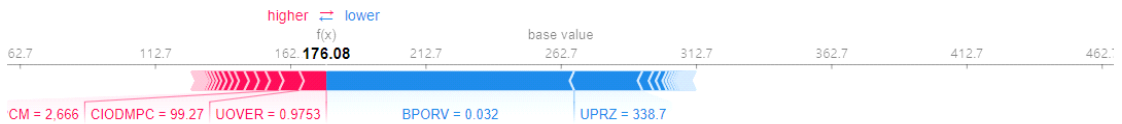


그림 15. 가압기 PORV 고장 ‘열림’ 시나리오 중반부 예측값에 SHAP을 적용한 결과

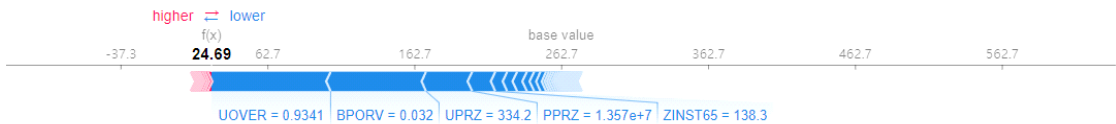


그림 16. 가압기 PORV 고장 ‘열림’ 시나리오 후반부 예측값에 SHAP을 적용한 결과

제 5 장 잔여 트립 시간 예측 시스템 인터페이스를 활용한 Case Study

잔여 트립 시간 예측 알고리즘의 각 기능의 결과를 한 번에 확인하기 어렵다는 문제가 있다. 따라서 본 연구에서는 알고리즘의 각 기능의 결과들을 효율적으로 제공하기 위해 인터페이스를 개발하였다. 해당 장에서는 개발된 잔여 트립 시간 예측 시스템의 인터페이스를 설명하고, 인터페이스를 통해 잔여 트립 시간 예측 시스템의 실험 결과를 보여주고자 한다.

제 1 절 잔여 트립 시간 예측 시스템 인터페이스

잔여 트립 시간 예측 시스템의 인터페이스는 그림 17과 같이 설계되었다. 가장 상단에는 운전 시간과 원자로 출력을 보여준다. 이때, 원자로 트립이 발생한 정보를 표현하기 위해 트립이 발생할 경우 원자로 출력 버튼이 빨간색으로 점등되도록 구현하였다.

그림 17의 파란색 박스는 훈련 상태 진단 기능에 대한 인터페이스를 보여준다. 훈련 상태 진단 기능의 출력값은 AE의 재구성 오차이며, 출력값에 따라 Trained 상태와 Untrained 상태를 나타낸다. 훈련 상태 진단 기능의 출력을 시각화하기 위해 Trained와 Untrained 버튼을 구현하였으며, 입력 데이터의 재구성 오차를 그래프로 나타낸다. 입력된 데이터의 재구성 오차가 파란색 점선으로 표시되는 Threshold 아래에 있을 경우 그래프는 초록색으로 표시되며, Trained 버튼에 초록색으로 알람 된다. 반대로 재구성 오차가 Threshold 위에 있을 경우 그래프가 빨간색으로 표시되며, Untrained 버튼에 빨간색으로 알람 된다.

그림 17의 초록색 박스는 비정상 시나리오 진단 기능의 인터페이스를 나타낸다. 비정상 시나리오 진단 기능의 인터페이스는 Normal, Abnormal, 진단 결과 버튼으로 구성된다. 정상 상태로 진단되면 Normal 버튼이 초록색으로 활성화된다. 비정상이 발생하면 Abnormal 버튼이 빨간색으로 알람 되며, 아래 버튼에 진단 결과를 나타내도록 구현하였다. 또한 훈련 상태 진단 기능에서 Normal 상태는 훈련되었으므로 Untrained 상태일 경우에도 Abnormal 버튼이 활성화된다.

그림 17의 빨간색 박스는 잔여 트립 시간 예측 기능의 인터페이스를 나타낸다. 잔여 트립 시간 예측 기능의 결과는 버튼과 그래프로 시각화한다. 버튼에는 잔여 트립 시간 예측 결과에 대한 정확한 정보를 제공하기 위해 구현되었다. 또한 본 논문에서는 예측 결과의 정도를 용이하게 파악할 수 있도록 관심, 주의, 경고, 긴급으로 나누어 색으로 표현하였다. 잔여 트립 시간 예측 결과가 900초 이상일 경우에는 관심으로 흰색으로 표현되고, 900초 이하일 경우에는 주의로 노란색, 600초 이하일 경우 경고로 주황색, 120초 이하일 경우 긴급으로 빨간색으로 표시된다. 그래프는 예측 결과의 추세를 파악하기 위해 구현된다. 그래프는 실시간으로 표시되며 잔여 트립 시간 예측 결과의 추세를 파악할 수 있다. 이는 조치가 수행되었을 때 예측 결과의 증가와 감소를 확인하기 위함이며, 예측 결과의 추세에 따라 원전 상태의 완화 및 악화를 파악할 수 있다.

그림 17의 갈색 박스는 잔여 트립 시간 예측 결과 해석 기능의 인터페이스를 나타낸다. 잔여 트립 시간 예측 결과 해석 기능 인터페이스는 예측 모델의 주요 영향 변수들에 대한 정보를 버튼에 나타나도록 구현되었다. 인터페이스에서는 Contribution과 Offsetting 변수들의 기여도 크기를 기준으로 상위 4개의 변수들에 대한 정보를 나타낸다. 잔여 트립 시간 예측 결과 해석 기능 인터페이스에서 빨간색 버튼은 Contribution 변수를 나타내며, 파란색 버튼은 Offsetting 변수를 나타낸다. 또한 버튼을 클릭하면 해당하는 변수의 추세를 시각화하기 위해 그래프로 나타내도록 구현하였다. 변수들의 추세를 나타내는 이유는 SHAP의 결과에서 제공하는 정보의 한계가 있기 때문이다. SHAP을 통해 예측 결과에 대하여 변수들이 영향을 주는 정도와 변수들의 현재 값을 확인할 수 있다. 그러나 변수들의 현재 값에 대한 정보만으로는 정상 상태에 비해 얼마나 차이가 나는지 확인하기 어렵다. 이를 해결하기 위해 변수들의 추세를 그래프로 나타내는 방법을 활용하며, 이를 통해 원전의 상태에 따른 주요 변수들의 변화 정도를 확인할 수 있다.

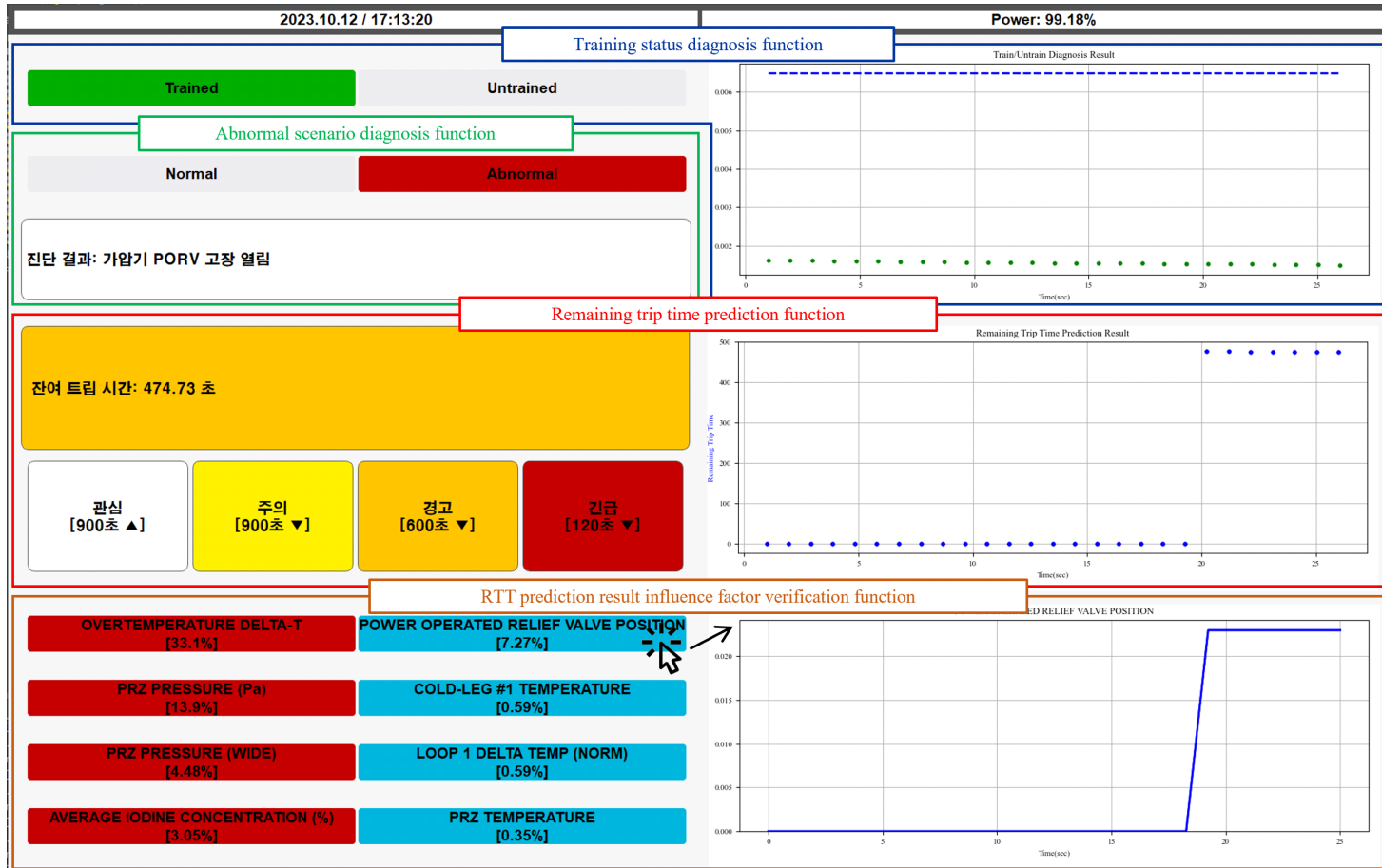


그림 17. 잔여 트립 시간 예측 시스템 인터페이스 개요

제 2 절 Case Study

잔여 트립 시간 예측 시스템은 비정상 상황에서 조치가 수행되었을 때를 가정하여 실험되며 인터페이스를 통해 결과를 확인한다. 인터페이스는 CNS와 연결되어 실시간으로 시스템의 결과를 확인할 수 있다. 해당 장에서의 실험은 가압기 압력 채널 고장 ‘고’ 비정상 시나리오 상황에서 완화 조치를 수행했을 때와 인적오류를 가정하여 잘못된 조치를 수행했을 때의 결과를 확인한다. 가압기 압력 채널 고장 ‘고’ 시나리오는 그림 18과 같이 실제 압력은 정상이나, 압력 채널의 고장으로 고압력을 지시한다. 이에 따라 자동 조치로 가압기 히터의 동작 정지와 가압기 스프레이 밸브가 개방되며 실제 압력이 트립 설정치 이하까지 감소하여 트립이 발생하는 시나리오이다. 해당 실험은 정상 상태에서 시작되며 30초에 비정상 사고가 주입된다. 또한 첫 번째 완화 조치로 가압기 히터를 동작시킨다. 완화 조치 이후 인적오류를 가정하여 스프레이 밸브를 더 개방하여 압력을 낮추는 작업이 수행된다. 마지막으로, 두 번째 완화 조치로 개방된 스프레이 밸브를 닫는다. 해당 실험의 순서는 그림 19에 나타난다.

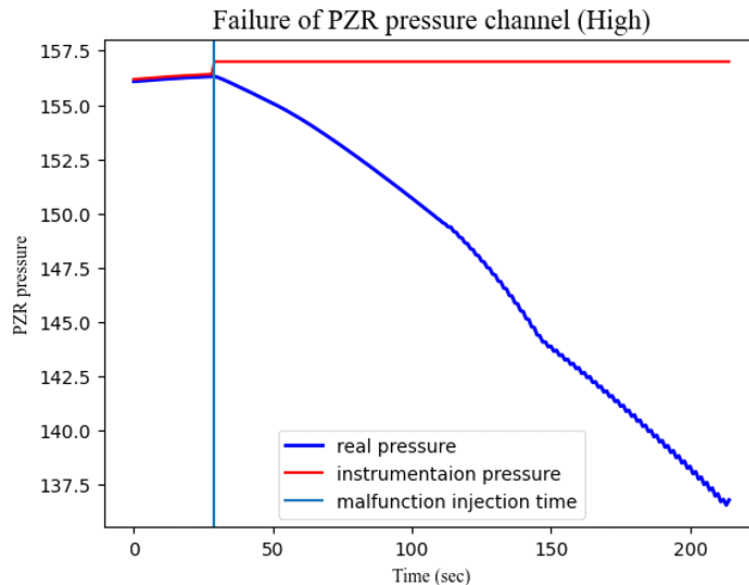


그림 18. 가압기 압력 채널 고장 ‘고’ 시나리오에서 실제 압력과 계측 압력



그림 19. 실험 설정 흐름도

그림 20은 비정상 사고가 주입된 이후 잔여 트립 시간 예측 시스템의 결과를 보여준다. 고장이 주입된 이후 비정상 시나리오 진단 기능 인터페이스에서 Abnormal 버튼에 빨간색으로 표시되며, 진단 결과를 나타낸다. 해당 실험에서 잔여 트립 시간 예측 결과는 트립까지 약 180초 남아있음을 나타내며, 주황색 배경을 통해 위험 상황임을 보여준다. 잔여 트립 시간 예측 결과 해석 기능 인터페이스에는 예측 결과에 대한 주요 영향 변수를 나타낸다. 주요 변수들 중 주목할 만한 변수로 가압기 압력이 고착되어 있는 것을 보여준다.

그림 21은 첫 번째 완화 조치를 수행하며, 완화 조치로 인해 변동된 잔여 트립 시간 예측 시스템의 결과를 보여준다. 해당 시나리오는 실제 압력이 감소하는 시나리오이며 압력 회복을 위해 가압기 히터를 동작하는 완화 조치가 수행된다. 완화 조치가 수행됨에 따라 재구성 오차가 크게 증가하며 Untrained 버튼이 알람 된 것을 보인다. 본 연구에는 조치가 포함되지 않은 데이터만 학습되었으며, 조치가 수행된 데이터는 학습되지 않은 데이터이기 때문이다. 또한 완화 조치가 수행됨에 따라 잔여 트립 시간이 증가하는 추세로 예측된다. 이러한 예측 결과에 대해 인공지능은 가압기 온도와 압력에 주목하는 것을 보인다. 가압기 압력은 고장 주입 이후 급격하게 감소하는 추세를 보였으며, 가압기 히터가 동작함에 따라 압력이 회복되고 있는 것을 나타낸다. 그러나 그림 22에서 잔여 트립 시간이 다시 감소하는 추세를 나타낸다. 이는 가압기 압력이 일정 수준에서 유지될 뿐 회복에 실패함에 따라 제어봉이 삽입되기 때문이다.

그림 23은 추가적인 조치로 가압기 스프레이 개도를 조절한 이후 잔여 트립 시간 예측 시스템의 결과를 보여준다. 기존의 완화 조치는 절차에 따라 압력 회복을 위해 가압기 스프레이 밸브를 닫아야 한다. 그러나 본 실험에서는 운전원의 인적오류를 가정하여 반대로 스프레이 밸브를 더욱 개방한다. 인적오류가 수행됨에 따라 잔여 트립 시간의 예측 결과가 급격하게 감소하는 추세를 보인다. 또한 잔여 트립 시간 예측 결과 해석 기능에서 가압기 스프레이 유량을 Offsetting 변수로 나타낸다. 잔여 트립 시간 예측 시스템의 결과를 통해 운전원의 조치가 잘못되었음을 인지할 수 있다.

그림 24는 인적오류 이후 개방된 스프레이 밸브를 닫는 완화 조치가 수행된 결과를 보여준다. 완화 조치 이후 잔여 트립 시간 예측 결과는 다시 상승하는 추세를 보인다. 또한 기존에 Offsetting 변수였던 가압기 스프레이 유량 변수가 Contribution 변수로 변한 것을 볼 수 있다. 이를 통해 가압기 스프레이 밸브의 조작이 잔여 트립 시간 예측 결과에 긍정적인 영향을 준 것을 알 수 있다.

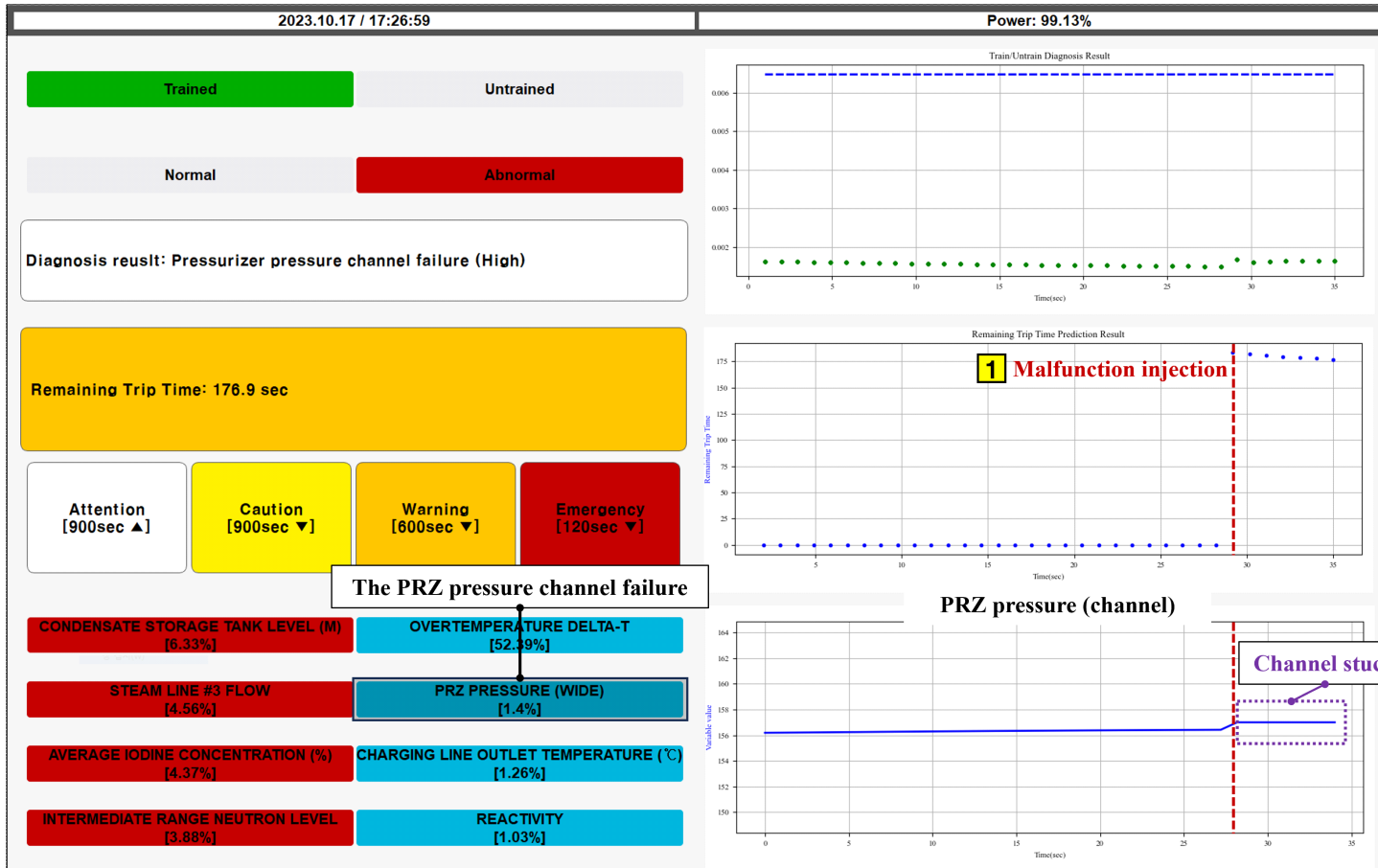


그림 20. 잔여 트립 시간 예측 시스템 결과 - 비정상 시나리오 주입

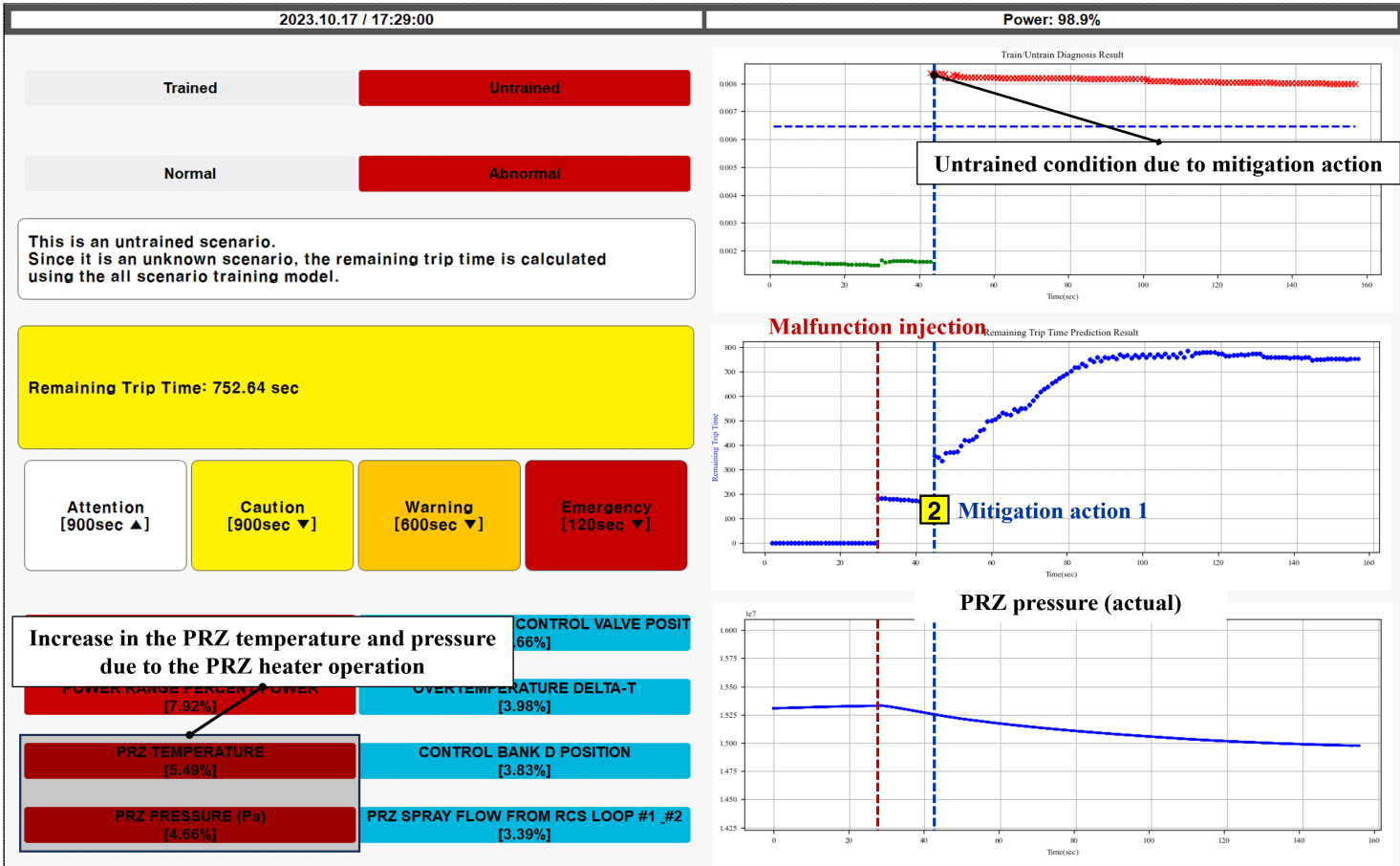


그림 21. 잔여 트립 시간 예측 시스템 결과 - 완화 조치 1 (가압기 히터 동작)

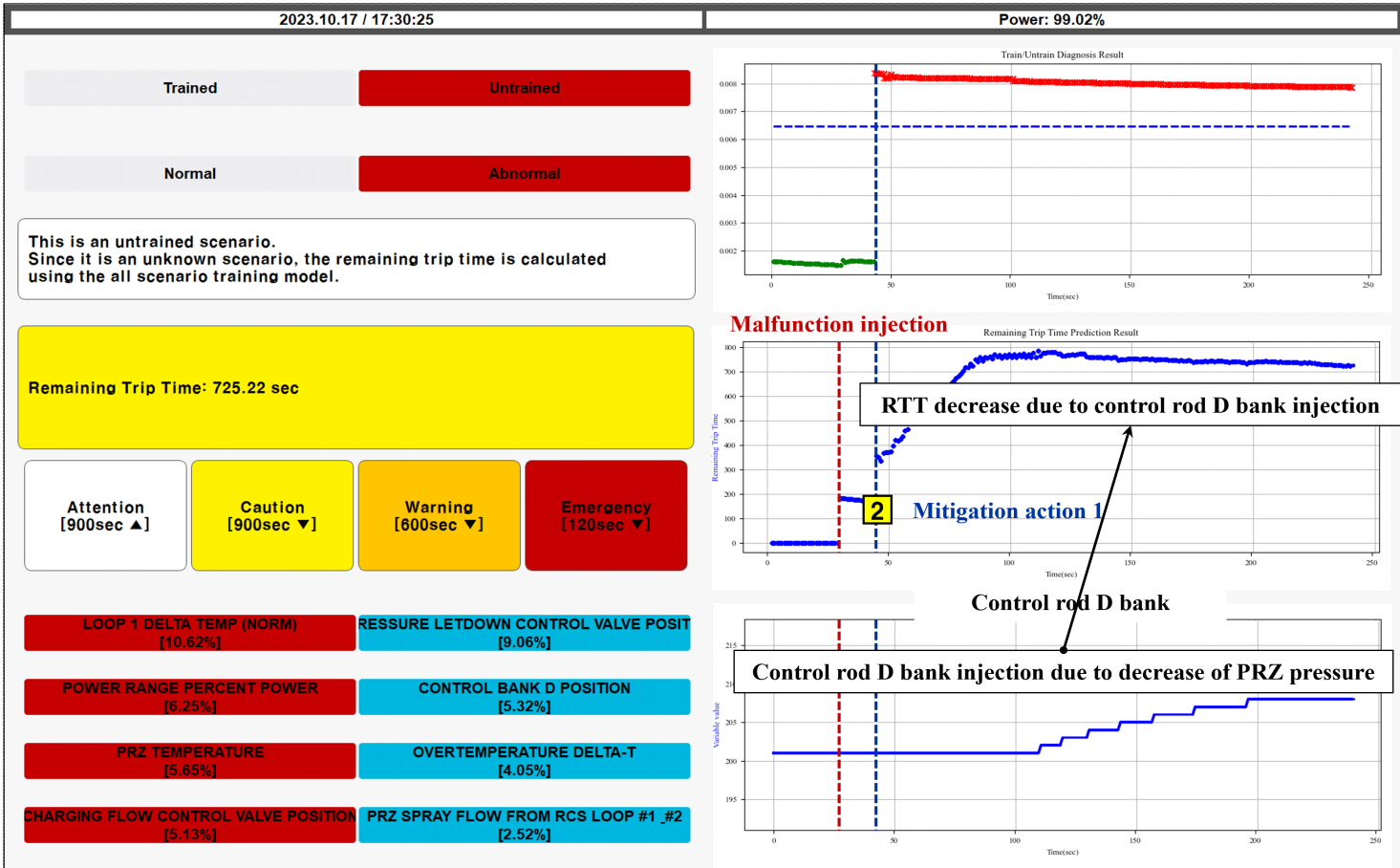


그림 22. 잔여 트립 시간 예측 시스템 결과 - 완화 조치 1 이후 회복 실패

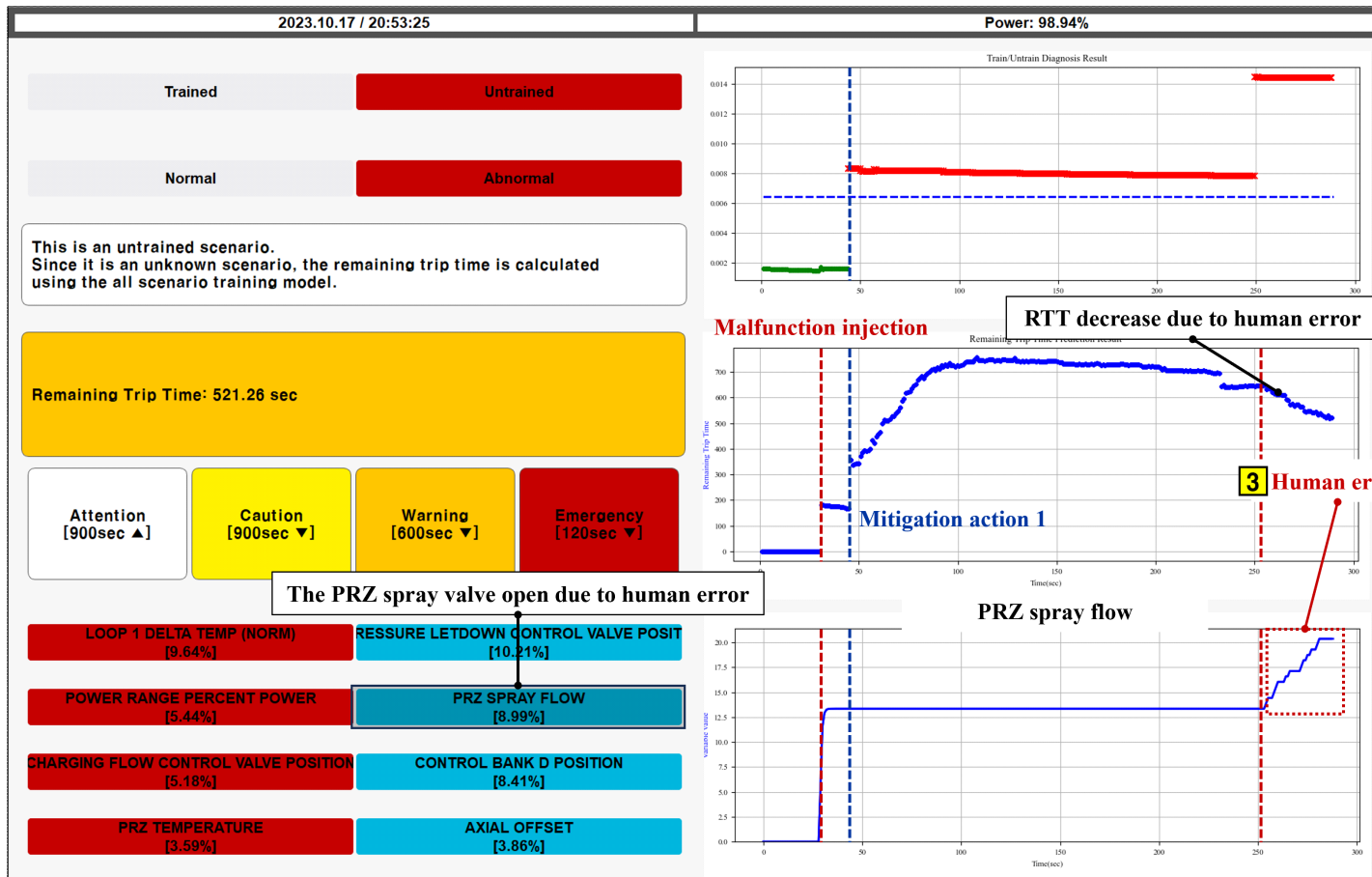


그림 23. 잔여 트립 시간 예측 시스템 결과 - 인적오류 (가압기 스프레이 밸브 개방)

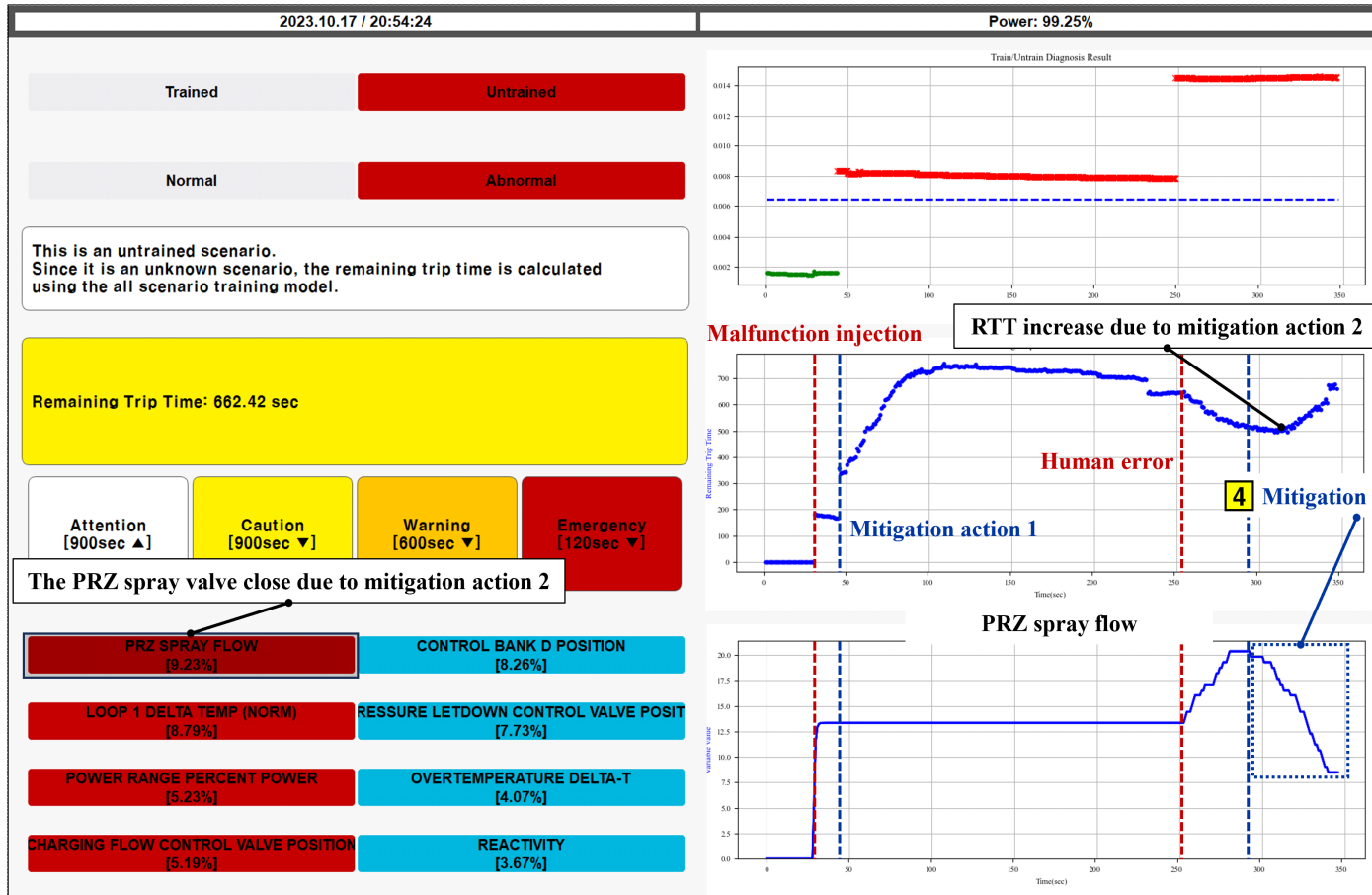


그림 24. 잔여 트립 시간 예측 시스템 결과 - 완화 조치 2 (가압기 스프레이 밸브 개도 조절)

제 6 장 결 론

원자력 발전소에서는 많은 원인에 의해 비정상 상황이 발생할 수 있다. 이러한 비정상 상황 속에서 운전원은 원자로 트립을 방지하며, 건전성을 회복하기 위한 작업을 수행한다. 이러한 작업은 제한된 시간 안에 급박하게 수행되며, 운전원의 업무에 부담이 가중된다. 운전원의 업무 부담은 실수를 유발할 수 있으며 인적오류가 발생할 수 있다. 이에 따라 본 논문에서 운전원의 잠재적인 인적오류를 방지하고 원전의 안전성을 향상시키기 위해 운전지원 시스템의 일환으로 잔여 트립 시간 예측 시스템을 개발하였다.

잔여 트립 시간 예측 시스템은 데이터 전처리 기능, 훈련 상태 진단 기능, 비정상 시나리오 진단 기능, 잔여 트립 시간 예측 기능, 잔여 트립 시간 예측 결과 해석 기능 총 5가지 기능으로 구성된 알고리즘으로 개발되었다. 잔여 트립 시간 예측 알고리즘에는 인공지능 방법인 Auto-Encoder (AE), Light Gradient Boosting Machine (LightGBM)이 활용되었으며, 잔여 트립 시간 예측 모델의 신뢰성을 향상시키기 위해 설명 가능한 인공지능 방법인 Shapley Additive Explanation (SHAP)을 활용하였다.

알고리즘의 각 기능의 인공지능 모델을 구축하기 위해 Compact Nuclear Simulator (CNS)를 이용하여 정상 데이터 및 비정상 데이터를 취득하였다. 취득된 데이터는 데이터 전처리 기능을 통해 변수 선택 및 정규화된 데이터로 활용된다. 또한 인공지능의 훈련 및 테스트를 위해 분할되어 활용되었다. 테스트 데이터를 활용하여 각 기능의 성능을 확인하며, 높은 정확도의 진단 결과와 예측 성능을 확인하였다. 또한 가압기 PORV 고장 ‘열림’ 시나리오에 대한 잔여 트립 시간 예측 결과 해석 기능을 테스트하였다. 테스트 결과 인공지능이 예측 결과를 도출함에 있어 중요하게 보는 변수들에 대한 정보가 해당 시나리오의 증상과 일부 일치하는 것을 확인하였다. 이에 따라 인공지능이 예측 결과를 도출함에 있어서 현재 상태에 대한 주요 증상들을 주요 인자로 활용하는 것을 확인하였다.

본 연구에서는 위의 알고리즘의 기능들의 결과를 효과적으로 제공하기 위해 인터페이스를 설계하였다. 또한 인터페이스는 CNS와 연동을 통해 실시간 정보에 대한 잔여 트립 시간 예측 시스템의 결과를 확인할 수 있다. Case Study에서는 인터페이스를 통해 완화 조치와 인적오류를 가정한 잘못된 조치를 수행하며 잔여 트립 시간 예측 시스템의 결과를 확인하는 실험을 수행하였다. 실험 결과 완화 조치가 수행될 때 잔여 트

립 시간 예측 결과가 상승하는 추세를 보였으며, 반대로 인적오류가 발생될 경우에는 잔여 트립 시간 예측 결과가 감소하는 추세를 나타내는 것을 확인하였다. 또한 잔여 트립 시간 예측 결과 해석 기능을 통해 예측 결과에 대한 주요 영향 변수들의 정보를 확인할 수 있었다. 또한 주요 영향 변수의 추세를 인터페이스로 제공하며 원전의 상태에 대한 정보를 나타내었다.

최종적으로, 본 논문은 비정상 상황에서 잔여 트립 시간을 예측하여 원전의 상태에 대한 정보를 제시해 줌으로써 운전원의 작업을 지원하는 것을 목적으로 잔여 트립 시간 예측 시스템을 개발하였다. 개발된 시스템의 인공지능 모델들은 높은 성능을 나타내는 것으로 판단되며, 실험을 통해 잔여 트립 시간의 예측 정보가 원전의 상태에 대한 정보로 활용될 수 있음을 보았다. 따라서 본 연구를 통해 비정상 상황에서 운전원의 작업에 도움을 주며 인적오류를 저감할 수 있을 것으로 기대된다.

참고문헌

- [1] K. H. Yoo, J. H. Back, M. G. Na, S. Hur, H. M. Kim, Smart support system for diagnosing severe accidents in nuclear power plants, *Nuclear Engineering and Technology*, Vol. 50, pp. 562-569, 2018.
- [2] H. J. Kim, A. M. Arigi, J. H. Kim, Development of a diagnostic algorithm for abnormal situations using long short-term memory and variational autoencoder, *Annals of Nuclear Energy*, Vol. 153, 108077, 2021.
- [3] J. H. Park, H. S. Jo, S. H. Lee, S. W. Oh, M. G. Na, A reliable intelligent diagnostic assistant for nuclear power plants using explainable artificial intelligence of GRU-AE, LightGBM and SHAP, *Nuclear Engineering and Technology*, Vol. 54, pp. 1271-1287, 2022.
- [4] H. J. Kim, J. H. Kim, Long-term prediction of safety parameters with uncertainty estimation in emergency situations at nuclear power plants, *Nuclear Engineering and Technology*, Vol. 55, pp. 1630-1643, 2023.
- [5] H. S. Jo, S. H. Lee, M. G. Na, Prediction of reactor trip variables using LSTM with attention mechanism in abnormal situations, *International Symposium on Future I&C for Nuclear Power Plants*, Nov. 15-17, 2021.
- [6] M. A. Kramer, Nonlinear principal component analysis using autoassociative neural networks, *AIChE JOURNAL*, Vol. 37, pp. 233-243, 1991.
- [7] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T. Y. Liu, LightGBM: A highly efficient gradient boosting decision tree, *Advances in Neural Information Processing Systems*, pp. 3149-3157, 2017.
- [8] D. Gunning, D. W. Aha, DARPA's explainable artificial intelligence program, *AI Magazine*, Vol. 40, pp. 44-58, 2019.
- [9] V. A. Dev, M. R. Eden, Formation lithology classification using scalable gradient boosted decision trees, *Computers and Chemical Engineering*, Vol. 128, pp. 392-404, 2019.
- [10] S. M. Lundberg, S. I. Lee, A unified approach to interpreting model

- predictions, *Advances in Neural Information Processing Systems*, pp. 4765–4774, 2017.
- [11] J. C. Park, K. C. Kwon, H. H. Cha, W. M. Park, S. J. Song, K. W. Seh, Y. C. Joo, Equipment and performance upgrade of compact nuclear simulator, KAERI/RR-1967/1999 KAERI:Daejeon, 1999.
- [12] M. Kuhn, K. Johnson, *Applied predictive modeling*, New York: Springer, ISBN: 9781461468493, 2013.
- [13] S. Garcia, J. Luengo, F. Herrera, *Data preprocessing in data mining*, Springer, 2015.
- [14] 솔라리스. 텐서플로로 배우는 딥러닝. 영진닷컴, 2018.
- [15] Y. G. Kim, D. S. Park, Consideration on the use of Explainable AI in Operator Support System, *Proceedings of the Korean Nuclear Society Virtual Autumn Meeting*, Dec. 17–18, 2020.