



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](#)

2021년 2월

석사학위 논문

보일러 미세먼지 저감을 위한 운전 변수  
예측 인공지능 모델의 강화학습 응용

조선대학교 대학원

산업공학과

이 채 교

보일러 미세먼지 저감을 위한 운전 변수  
예측 인공지능 모델의 강화학습 응용

Reinforcement Learning Application of Operation Variable  
Prediction AI Model to Reduce Boiler Micro Dust

2021년 2월 25일

조선대학교 대학원

산업공학과

이 채 교

보일러 미세먼지 저감을 위한 운전 변수  
예측 인공지능 모델의 강화학습 응용

지도교수                    신    종    호

이 논문을 공학 석사학위신청 논문으로 제출함

2020년 10월

조선대학교 대학원

산업공학과

이    채    교

## 이채교의 석사학위논문을 인준함

위원장    조선대학교 교수    백 정 우 (인)

위    원    조선대학교 교수    신 종 호 (인)

위    원    조선대학교 교수    김 성 준 (인)

2020년 11월

조선대학교 대학원

## 목차

목차 .....	i
그림 목차 .....	iii
표 목차 .....	iv
ABSTRACT .....	v
제 1 장 서론 .....	1
1.1 연구 배경 .....	1
1.2 관련 연구 .....	2
1.3 연구 목표 .....	4
1.4 연구 구성 .....	5
제 2 장 배경 이론 .....	6
2.1 Automated Machine Learning .....	6
2.1.1 Random Forest .....	8
2.1.2 Extremely Randomized Tree Forest .....	9
2.1.3 GBM .....	10
2.1.4 Deep Neural Network .....	11
2.1.5 GLM .....	11
2.1.6 Stacked Ensemble .....	12
2.2 강화학습 .....	12
2.3 deep Q-learning .....	17

제 3 장 AutoML 을 이용한 SOx, NOx 예측 .....	19
3.1 화력 발전 보일러의 운전변수 예측 .....	19
3.1.1 변수 중요도 기반 독립변수 정의 .....	21
3.1.2 최적 독립변수 예측 모델 개발 .....	23
3.2 예측 독립변수를 이용한 SOx, NOx 예측 .....	31
제 4 장 deep Q-learning 을 이용한 최적 혼소조합 제어 ..	33
4.1 deep Q-learning 구성 요소 정의 .....	33
4.1.1 상태 .....	33
4.1.2 행동 .....	34
4.1.3 보상 .....	35
4.2 deep Q-learning 기반 최적 혼소조합 제어 알고리즘	36
4.2.1 수행방법 .....	36
4.2.2 수행결과 .....	38
제 5 장 결론 및 토의 .....	41
참고문헌 .....	43

## 그림 목차

그림 1. 1973년과 2015년 세계 전기 발전 중 연료별 차지 비율 비교[1].....	1
그림 2. 자동 기계학습 시스템(Auto ML: Automated Machine Learning).....	6
그림 3. 그리드 탐색과 랜덤 탐색[18].....	7
그림 4. 랜덤 포레스트[19].....	8
그림 5. 심층신경망(Deep Neural Network).....	11
그림 6. 강화학습 환경 .....	12
그림 7. Markov decision process.....	13
그림 8. Deep Q-learning .....	17
그림 9. 석탄화력발전소 계통도[28].....	19
그림 10. 데이터 전처리.....	20
그림 11. 예측 독립변수를 이용한 종속변수 예측 프로세스.....	22
그림 12. 예측 모델의 변수 중요도 그래프.....	27
그림 13. SO <sub>x</sub> , NO <sub>x</sub> 예측 모델 프로세스 .....	31
그림 14. deep Q-learning 기반 최적 혼소조합 제어 구조.....	36
그림 15. 강화학습 정책 발전 결과.....	38



## 표 목차

표 1. AI 모델 학습 실험 환경.....	23
표 2. SO <sub>x</sub> , NO <sub>x</sub> 성능 결과 값.....	25
표 3. 독립변수로 활용하기 위한 종속변수 후보(상태정보).....	28
표 4. 예측에 사용된 독립변수(연료성상 + 출력).....	28
표 5. 예측독립변수 후보 모델 성능 결과.....	29
표 6. 독립변수 그룹 별 SO <sub>x</sub> , NO <sub>x</sub> 성능 비교.....	32
표 7. 강화학습을 위한 혼소조합 제어 모델의 상태.....	34
표 8. 강화학습을 위한 혼소조합 제어 모델의 행동.....	35
표 9. 강화학습 모델 학습 실험 환경.....	37
표 10. 강화 학습 파라미터.....	37
표 11. 강화학습 기반 최적 혼소조합 제어.....	39

## ABSTRACT

### Reinforcement Learning Application of Operation Variable Prediction AI Model to Reduce Boiler Micro Dust

Chae-Kyo Lee

Advisor: Prof. Jongho Shin, Ph.D.

Department of Industrial Engineering

Graduate School of Chosun University

Recently, social issues caused by fine dust have threatened the health of people around the world. In particular, sulfur oxides (SO<sub>x</sub>) and nitrogen oxides (NO<sub>x</sub>) generated in the process of burning coal are declared as fine dust first-class carcinogens. Already, advanced countries are shutting down coal-fired power plants in stages without expanding them due to damage caused by accelerating global warming and continuous emission of air pollutants. In South Korea, the government plans to suspend new construction of coal power plants and reduce the operation of existing facilities in order to reduce fine dust under the 2020-2040 National Environmental Comprehensive Plan. Recently, major domestic power generation companies have been pushing to reduce fine dust emissions from coal-fired power plants, which are further strengthened than government measures. In addition, it is actively investing in R&D such as development of technologies for improving

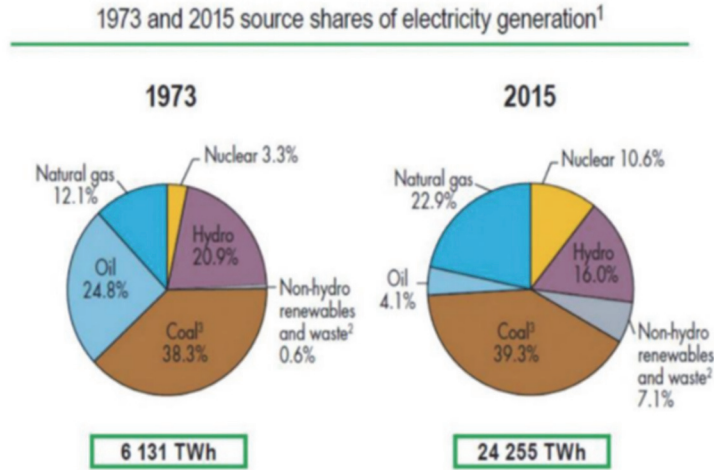
efficiency of reduction facilities to reduce sulfur oxides (SO<sub>x</sub>) and nitrogen oxides (NO<sub>x</sub>), which are major precursors of fine dust.

In this study, operating variables are predicted to reduce fine dust in thermal power generation boilers. Locate the value of the operation variable that minimizes SO<sub>x</sub> and NO<sub>x</sub> by linking the SO<sub>x</sub> and NO<sub>x</sub> prediction models to the Reinforcement learning. Create a NO<sub>x</sub> and SO<sub>x</sub> prediction model using the operation information and status information that are the operation variables of the boiler. AutoML (Automated Machine Learning) was used to create the prediction model. AutoML is used as a tool to instead perform experimental actions and trial and error that adjust various parameters that occur in the machine learning model process. In addition, in order to improve the prediction of SO<sub>x</sub> and NO<sub>x</sub>, a model that predicts other dependent variables is created and the predicted value is reused as an independent variable. The process of selecting the optimal independent variable was selected by utilizing Variable Importance. It was confirmed that the SO<sub>x</sub> and NO<sub>x</sub> model prediction performance is improved when the values of the optimal independent variable and the dependent variable are reused as independent variables. The NO<sub>x</sub> and SO<sub>x</sub> prediction models were linked with deep Q-learning, which is one of the reinforcement learning algorithms, to optimize the reduction of fine dust. Within the deep Q-learning environment, SO<sub>x</sub> and NO<sub>x</sub> prediction models were used to update status. The action controls the mixture ratio of the boiler combustion carbon model, which is the driving variable. Depending on the action, SO<sub>x</sub> and NO<sub>x</sub> values will be returned as rewards. Compensation develops policy in the direction of minimizing SO<sub>x</sub> and NO<sub>x</sub> values. For the development of the policy, walked with the ratio of

main coal type and auxiliary coal type and the ratio of high calorie bullet and low-calorie coal as constraints. As a result, the amount of main coal species used increased by an average of 3%, and the proportion of low-calorie coal used increased by an average of about 24%. As shown, the SO<sub>x</sub> value decreased by 8% on average, and the NO<sub>x</sub> value also decreased by 8.3% on average. Through the study of reinforcement learning able to consider both economic and environmental aspects in the operation of thermal power boilers at the same time.

## 제 1 장 서론

### 1.1 연구 배경



**그림 1. 1973년과 2015년 세계 전기 발전 중 연료별 차지 비율 비교[1]**

석탄화력발전은 오랜 시간 동안 다른 화석 연료 대비 저렴한 가격과 풍부한 자원을 바탕으로 에너지를 생산하기 위한 다양한 수단 중 하나로 사용되고 있으며, 전 세계적으로 생산되는 에너지의 30% 이상의 비중을 차지하고 있다.[1]

그러나 석탄화력발전은 다른 화석 연료를 이용한 에너지 생산 방식에 비해 상대적으로 오염물질 배출량이 많은 문제가 있다. 특히 석탄을 연소하는 과정에서 발생하는 황산화물(SOx), 질소산화물(NOx)은 1급 발암물질로 규정되는 미세먼지를 생성하는 원인 물질로 지목되면서 심각한 사회문제로 인식되고 있다. 따라서 선진국에서는 지구온난화 가속 및 지속적인 대기오염물질 배출에 따른 피해로 석탄화력발전소를 증설하지 않거나 단계적으로 폐쇄하고 있는 실정이다.

환경부 미세먼지 관리 종합대책(2017. 09)[2]에 따르면 우리나라의 전체 미세먼지 배출량의 72%가 간접배출로 인해 발생하고 있다. 2017년 7월 NASA와 공동연구로 진행

된 미세먼지 성분분석 결과에서도 2차 생성물질이 전체 성분의 75% 이상을 차지한 것으로 확인되었다. 특히 2차 생성 원인물질 중 NOx, Sox로 인한 생성량이 가장 많아 대기오염물질의 통합적 관리를 필요로 하고 있다.[3] 국내에서는 미세먼지의 주요 배출원으로 석탄화력발전소가 주목받고 있다.

제 5차 국가환경종합계획(2020 ~ 2040)에 따라 미세먼지 저감을 위하여 석탄발전소의 신규건설 중단 및 기존 시설 운영 감축을 추진중이다. 최종 목표인 ‘탈 석탄 사회’로의 전환을 위해 ‘탈 석탄 이행안(로드맵)’에 대한 사회적 대화를 추진하고, 석탄발전소의 감축과 함께 친환경 연료전환, 배출시설 관리 강화 등을 통해 2040년까지 초미세먼지(PM2.5) 농도를 세계보건기구(WHO)의 권고 수준인  $10\mu\text{g}/\text{m}^3$ 으로 저감한다는 계획이다.[4]

정부 정책에 많은 영향을 받는 국내 주요 발전사들은 석탄화력발전소의 미세먼지 배출 저감을 위해 정부대책보다 한층 강화된 미세먼지 감축을 추진하고 있다.[5] 이와 더불어 미세먼지의 주요 전구물질인 황산화물(SOx), 질소산화물(NOx) 저감을 위하여 저감설비 효율향상 기술개발 등 R&D(연구개발)에 적극적으로 투자하고 있다. 또한 4차 산업혁명 핵심기술(IoT, 빅데이터, AI 등)을 활용한 본격적인 스마트 발전소 구축을 통해 미세먼지 저감 문제에 적극적으로 대응하려고 하고 있다. 스마트 발전소는 수많은 센서에서 축적된 데이터와 운전정보를 분석하고 실시간으로 발전소의 운전상황을 판단하여 최적화된 의사결정을 내리는 지능화된 발전소 구현을 목표로 하고 있다.

## 1.2 관련 연구

본 논문에서는 운전변수 예측 인공지능 모델과 강화학습을 이용하여 황산화물(SOx), 질소산화물(NOx)을 저감하는 자동화 기술을 개발하고자 한다. 기존의 화력발전소에서는 초미세먼지, SOx, NOx를 저감하는 방법으로는 정부 주도로 1종 사업장(배기가스 방출량 80톤/년 이상)에 해당하는 발전소는 배출되는 유해가스 중 순차적으로

NO<sub>x</sub>, 미세먼지, SO<sub>x</sub>, 초미세먼지처리를 위해 300~400℃에서 선택적촉매환원법(암모니아(혹은 우레아)+SCR), 건식 전기집진기술, 습식 탈황공정, 습식 전기집진기술등이 적용되어 운전 중에 있다.[6] 이러한 기술은 현재 안정적으로 잘 운영되고 있으나 2~4종 미세먼지 배출 사업장과 해양선박 후처리 분야에 적용하는 것은 공간제약 및 운영비용에 관한 부담으로 인해 어렵다. 따라서 연구자들은 기존 설비를 유지한 채 인공지능을 이용한 보일러 공정을 모델링하여 에너지 효율을 높이거나 미세먼지, NO<sub>x</sub>, SO<sub>x</sub>를 저감 시키는 연구를 진행했다. 특히 이러한 연구는 해외를 중심으로 활발히 진행되고 있다.

보일러 공정 모델링에 관한 연구는 다음과 같다. Gu et al.[7]은 보일러 연소를 모델링하기 위한 적응형 LSSVM(Least Square Support Vector Machine) 알고리즘의 2단계 학습 전략을 제안하고 유효성 확인을 위해 시간 변동 비선형 함수를 사용한 사례 연구를 수행했다. Lv et al.[8]은 훈련 데이터 세트를 FCM(fuzzy c-means)과 부분 최소 제곱으로 분해 및 결합하여 NO<sub>x</sub> 배출 예측을 위한 기본 알고리즘으로 LSSVM을 이용한 앙상블 모델을 적용했다. Lv et al.[9]은 LSSVM(Least Square Support Vector Machine) 및 자연 주문을 위한 Swarm Optimization 기법을 사용하여 300MW CFB 보일러의 베드 온도를 예측하기위한 동적 모델을 개발 했다. Li et al.[10]은 300MW CFB에서 보일러 효율 및 NO<sub>x</sub> 배출량을 예측하기위한 심층 양방향 학습 머신을 소개 했다. 위 연구들은 석탄화력발전의 보일러 공정을 모델링을 통한 보일러 효율성 및 NO<sub>x</sub> 배출 예측에 관한 연구를 진행하였다. 이러한 연구를 바탕으로 다양한 알고리즘을 적용한 운영 최적화에 관한 연구가 진행되었다. Shi et al.[11]은 과거 플랜트 데이터(부하, 과잉 공기, 회분 및 수분) 및 전산 유체 역학 시뮬레이션 데이터를 사용하여 열 효율 및 NO<sub>x</sub> 배출 측면에서 연소 최적화를 위한 유전 알고리즘과 ANN의 결합 모델을 평가 하였다. Tuttle et al[12]은 ANN을 사용하여 2년 동안 실제 운영되는 490 MW 접선 석탄 화력 발전소의 NO<sub>x</sub> 감소를 위한 온라인 연소 최적화 시스템을 발표했다. Tan et al.[13]은 Combined advanced extreme 기계학습과 harmony search 알고리즘을 결합하여 700MW 석탄 화력 발전소의 운영 매개 변수를 최적화하여 NO<sub>x</sub> 배출을 제어하고자

했다. Tang and Zhang[14]은 660MW 석탄 화력 발전소의 보일러 효율 및 NOx 배출량의 최적 값을 얻기 위해 deep belief network 및 JAYA 알고리즘의 프레임 워크를 개발하였다. Cheng et al.[15]은 화력 발전소의 인증 된 DCS 데이터를 사용하여 보일러 효율을 높이는 동시에 폐가스 배출량을 줄이기 위해 LSTM-ConvNet 예측기와 DQN 강화 학습을 통해 석탄 연소 보일러 용 온라인 최적화 시스템인 Thermal Net을 개발 했다. 보일러 공정을 모델화하고 예측하기 위한 기존의 연구들은 보일러 효율, 연소상태, NOx 배출 제어에 연구 목표가 집중되어 있어 NOx와 SOx를 동시에 고려한 예측 모델 및 미세먼지 저감을 위한 보일러 운용변수 최적제어에 관한 연구가 부족하다.

### 1.3 연구 목표

본 연구에서는 보일러의 미세먼지 저감을 위해 AutoML(Automated Machine Learning)을 이용하여 NOx, SOx 예측 모델을 만든다. AutoML은 기계학습 모델링 과정에 발생하는 다양한 파라미터들을 조정하는 실험적 행위와 시행착오를 대신 수행하는 도구로 사용된다. 인공지능 전문가를 대체하는 것이 아닌 일부 모델 생성 과정을 도와주는 자동화 도구이다. AutoML을 이용한 NOx, SOx 예측 모델을 생성하기 위한 독립 변수의 선정은 전체 독립변수를 기반으로 학습한 모델에서 종속변수의 중요도를 파악하고 최소한의 독립 변수를 선택하게 하였다. 생성된 NOx, SOx 예측 모델은 미세먼지 저감 최적화를 위해 강화학습 알고리즘 중 하나인 deep Q-learning의 최적화에 이용된다. deep Q-learning은 Q-learning과 deep neural network가 결합된 방식이다. Q-learning은 현재 상태, 현재 상태에서의 행동, 행동에 대한 보상, 현재 상태에서 취한 행동을 통해 결정된 다음 상태에 대한 경험적 샘플들을 활용하여 정책을 발전시킨다. 또한, 상태와 행동에 대한 가치함수는 deep neural network로 근사(approximation)하며, deep neural network가 최적의 행동을 결정하는 정책 역할을 수행하도록 한다. 정책의 발전은 기억 메모리(replay memory)에서 무작위로 추출된 다양한 경험적 샘플들을 활용하게 된다[16]. 본 연구에서는 deep Q-learning을 이용한 강화학습 정책 발



전을 위해 NOx, SOx 예측 모델을 연동하여 상태, 행동, 보상과정 중 실시간으로 데이터 전달이 가능하도록 하였다. 본 연구에서 개발한 deep Q-learning 기반 최적 혼소조합 제어 모델은 주어진 보일러 연소 탄종의 혼소 비율을 효율적으로 제어하게 되는데, 주탄종과 보조탄종의 비율 및 고열량탄과 저열량탄의 비율 등을 경제적 측면에서 고려하면서 NOx, SOx 배출을 최소화하는 정책을 탐색하도록 구현하였다.

## 1.4 연구 구성

1장 이후 본 연구의 구성은 다음과 같다. 2장에서는 본 연구에서 사용되는 AutoML과 강화학습 deep Q-learning의 배경 이론에 대해 설명한다. 3 장에서는 deep Q-learning에 사용되는 SOx, NOx 예측 모델 개발을 위해 Variable Importance를 사용한다. Variable Importance는 종속변수 예측 값을 독립변수로 쓰는 예측 독립변수를 선정하는 데 사용되었다. 기존 독립변수와 예측 독립변수를 추가로 사용했을 때의 SOx, NOx 예측 모델의 성능 값을 비교한다. 4장에서는 본 연구의 목표인 미세먼지 저감을 위해 운전변수 중 하나인 SOx, NOx 예측 모델을 deep Q-learning의 환경과 연동시켜 상태를 업데이트하는데 사용한다. 상태안에서 혼소조합이라는 액션에 따라 SOx, NOx는 상태 값을 업데이트 시키고 보상을 반환한다. 보상은 SOx, NOx 수치를 최소화하는 방향으로 정책 학습을 한다. 그 후 특정 날짜의 혼소 비율을 최적화하여 결과를 비교해 본다. 5장에서는 본 연구의 결론 및 향후 연구 방향에 대해 기술한다.

## 제 2 장 배경 이론

본 장에서는 본 논문의 배경이 되는 AutoML(Automated Machine Learning)과 강화 학습 그리고 deep Q-learning에 대해 소개한다.

### 2.1 Automated Machine Learning

최근 인공지능을 이용한 연구는 이미지, 비디오, 자연어, 수치 기반 태스크의 추론 정확도를 높이는 것에서 더 나아가 신경망의 파라미터 최적화 및 자동 구조화 분야로 그 폭을 넓혀 가고 있다. 인공 신경망 기반 데이터 분석 태스크는 데이터 탐색 - 데이터 전처리/정제 - 특징 추출 - 모델 선택 - 모델 훈련 및 최적화의 단계로 진행된다. 특히 목표 데이터별로 모델을 선정하거나 훈련 및 최적화 단계에 많은 시간과 자원 그리고 도메인 전문지식 등이 소요된다. 따라서 다양한 목표 데이터 및 다른 태스크에서 같은 작업이 반복 수행되어야 한다는 문제점이 발생한다. 이런 문제점을 극복하고자 많은 AI 연구자들이 자동 기계학습(AutoML: Automated Machine Learning) 기술에 주목하고 있다. AutoML은 데이터 특징 추출, 모델 훈련의 하이퍼 파라미터 최적화(HPO:Hyper Parameter Optimization) 및 인공지능 모델 탐색(AI Model Search)과 같은 머신러닝 학습의 주요 프로세스를 자동화하는데 사용된다.

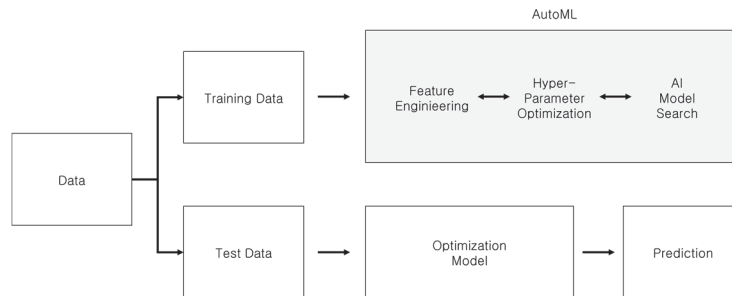
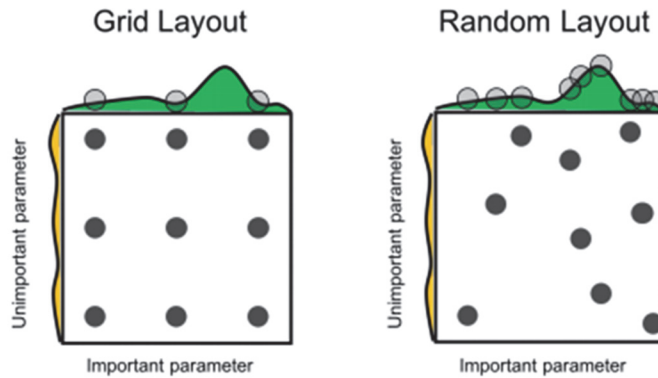


그림 2. 자동 기계학습 시스템(Auto ML: Automated Machine Learning)

본 연구에서는 AutoML을 구현하기 위해 파이썬 오픈소스 라이브러리 중 하나인 java 기반의 패키지인 H2O의 AutoML을 사용했다.[17] H2O의 AutoML은 지정된 인공지능 모델그룹(Random Forest, Extremely Randomized Tree Forest, GBM, Deep Neural Network, GLM, Stacked Ensemble)안에서 하이퍼 파라미터 최적화 과정을 거쳐 최적의 모델을 도출한다. 하이퍼 파라미터는 머신러닝 모델을 구성하는 변수로 학습률, 학습률의 스케줄링 방법, 손실함수, 훈련 횟수, 가중치 초기화 방법, 정규화 방법, 적층 계수 등과 같이 훈련성능에 직접적인 영향을 미치는 다양한 변수이다. 즉, 개별 변수를 조율하는 방식에 따라서 다양한 하이퍼 파라미터 설정이 도출되고, 이런 수많은 조합중 최적의 조합을 탐색하는 기술이 요구된다. 하이퍼 파라미터를 최적화하기 위한 탐색방식으로 H2O는 그리드 탐색(Grid Search)와 랜덤 탐색(Randomized Search)를 지원한다.



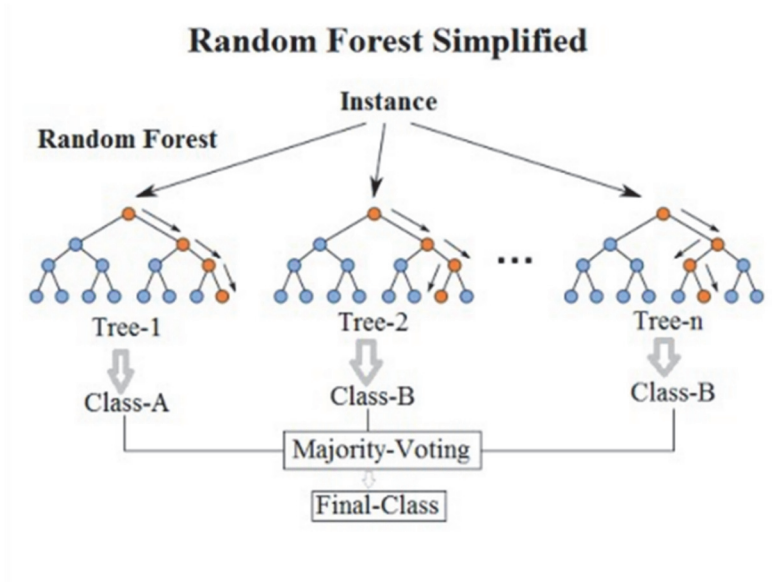
**그림 3. 그리드 탐색과 랜덤 탐색[18]**

그리드 탐색은 특정 하이퍼 파라미터 구간에서 일정 간격으로 지정된 하이퍼 파라미터 값을 선택하여 성능을 측정하고 비교하여 가장 높은 성능을 도출하는 하이퍼 파라미터 값을 최적해로 설정한다. 그리드 탐색은 구간 전역을 탐색하기 때문에 하이퍼 파라미터 변수 조합을 증가시킬수록 탐색 시간이 기하급수적으로 증가한다는 단점이 발생한다. 또한, 균일한 간격으로 탐색과정을 진행하기 때문에 그림 3의 좌측처럼

최적 하이퍼 파라미터 값을 찾지 못하는 경우가 발생할 수 있다. 이와 같은 문제를 차단하기 위해 랜덤 탐색이 제안되었다.[18] 랜덤 탐색은 그리드 탐색과 달리 **그림3**의 우측처럼 하이퍼 파라미터 구간내에서 임의의 값을 선택하기 때문에 불필요한 반복탐색을 줄이고 보다 빠르게 최적 하이퍼 파라미터를 발견할 수 있는 가능성을 높였다.

다음은 AutoML에 사용된 인공지능 모델을 간략하게 설명한다. 인공지능 모델은 지도학습의 주요 모델 인 Random Forest, Extremely Randomized Tree Forest, GBM, Deep Neural Network, GLM 로 이루어져 있다. 또한 예측 성능 향상을 위해 학습된 예측 모델을 기반으로 한 Stacked Ensemble 모델을 사용한다.

### 2.1.1 Random Forest



**그림 4. 랜덤 포레스트[19]**

랜덤포레스트(Random Forest)는 단일 머신러닝의 알고리즘 예측력을 높이기 위해 여러 개의 머신러닝 알고리즘을 융합하는 앙상블 학습 방법의 일종이다. 훈련 과정에

서 구성한 다수의 의사결정 나무로부터 출력되는 각각의 추정 결과를 이용한 분류 방법이다.[20] 랜덤포레스트는 데이터 집합의 원본으로부터 각각의 부트스트랩 샘플을 생성한다. 만들어진 의사결정 나무모델은 입력 데이터가 주어질 시에 각각의 출력결과를 종합하여 예측하게 된다. 랜덤포레스트는 데이터 집합의 원본을 카테고리별로 나눈 후 이를 무작위로 추출하는 행위인 반복 추출방식을 적용하여 원래 데이터의 수만큼 크기를 갖도록 샘플링하는 부트스트랩 방식을 채용한다. 생성된 샘플은 Majority Voting방식을 이용하여 입력된 데이터를 분류하게 된다. 이러한 방식은 각각의 예측된 결과를 종합하고 입력된 데이터에 대한 예측을 출력하게 되면 아래의 식과 같이 표현할 수 있다.

$$\text{Ensemble}(y') = \operatorname{argmax} \sum_{j=1}^n (y'_j)$$

$y'_j$ 은  $j$ 번째 의사결정나무에서 입력데이터에 대한 예측 결과이다.  $\text{Ensemble}(y')$ 에서는 각각의 의사결정나무가 예측한 결과를 종합하여 가장 많은 빈도수를 가지는 예측 결과값을 최종적인 예측 값으로 출력한다.

### 2.1.2 Extremely Randomized Tree Forest

극단적 랜덤 트리(Extremely Randomized Tree Forest)는 랜덤포레스트를 기반으로 연구된 기법이다. 랜덤포레스트보다 더 랜덤화된 요소를 가지고 있는 트리로 트리 분기시 컷 포인트(Cut-Point)를 찾을 때 이를 랜덤으로 결정하게 된다.[21] 극단적 랜덤 트리는 편향을 최소화하기 위해 랜덤 포레스트와 달리 부트 스트랩 복제본이 아닌 전체 학습 샘플을 사용하여 의사 결정 나무를 증가시킨다. 또한 분기되는 지점을 랜덤으로 선택하기 때문에 훈련속도가 랜덤 포레스트보다 약 3배정도 빠른 편이며 분류 문제는 Voting방식으로 결과를 산출하고, 회귀문제는 나온 수치 값의 평균으로 결과

치를 산출한다.

### 2.1.3 GBM

Gradient Boosting Algorithm(GBM)은 회귀분석 또는 분류분석을 수행할 수 있는 예측모형이며 예측모형의 앙상블 방법론 중 부스팅 계열에 속하는 알고리즘이다.[22] GBM은 특히 Tabular format 데이터에 대한 예측에서 높은 성능을 보여주고, 머신러닝 알고리즘 중에서도 가장 예측 성능이 높다고 알려져 있다. 특히 Kaggle에서 많은 연구자들이 사용하는 XGBoost도 GBM을 기반으로 개발된 알고리즘이다. GBM은 계산량이 상당히 많이 필요한 알고리즘이기 때문에, 이를 하드웨어 효율적으로 구현하는 것이 필요한데, LightGBM [23], CatBoost[24], XGBoost[25] 같은 알고리즘이 GBM을 기반으로 한 향상된 알고리즘 기법이다. GBM은 경사 하강법(Gradient Descent)을 이용해 가중치 업데이트를 수행하는 기법이다. 분류의 실제 결과 값을  $y$  설명변수를  $x_1, x_2, \dots, x_n$ 이라 정의하고, 이 설명변수에 기반한 예측 함수를  $F(x)$ 라고 하면 아래와 같은 식이 된다.

$$h(x) = y - F(x)$$

위의 실제 결과 값  $y$ 에서 예측 함수  $F(x)$ 를 이용한 예측 값을 뺀 값을 잔차(Residual) 값이라고 한다. 위 잔차식  $h(x)$ 를 최소화하는 방향성을 가지고 반복적으로 가중치 값을 업데이트 하는 것이 경사 하강법이다. GBM은 A라는 모델을 생성하여  $y$ 를 예측하고 남은 잔차를 다시 B라는 모델을 통해 예측하고 A+B 모델을 통해  $y$ 를 예측하고 이러한 모델을 여러 개 만들어 잔차를 계속 줄여 나간다. 하지만 이러한 방식은 bias는 줄일 수 있어도, 과적합이 일어날 수 있다는 단점이 있다. 따라서 GBM을 사용할 때 Sampling, Penalizing등의 Regularization 기법이 사용된다.

## 2.1.4 Deep Neural Network

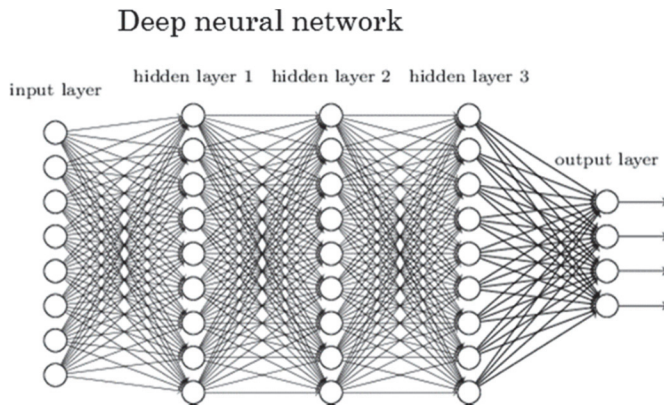


그림 5. 심층신경망(Deep Neural Network)

심층 신경망(DNN: Deep Neural Network)은 인간 두뇌를 모방한 계층적 네트워크에서 수 많은 레이어에 입력 데이터를 훈련 함으로써 분류가 수행되는 인공지능 기술이다. 심층 신경망은 입력 레이어(input layer), 출력 레이어(output layer), 은닉 레이어(hidden layer)로 구성되며 이를 통하여 특징의 추출을 자동으로 수행한다. 심층 신경망은 입력된 데이터간의 의존관계 확인 없이 복잡한 비선형함수의 처리가 가능하다. 일반적으로 학습 데이터의 양이 많을수록 성능이 향상된다.

## 2.1.5 GLM

일반화선형모형(GLM: Generalized Linear Models)은 일반선형모형(General Linear Model)의 가정들이 적용될 수 없는 경우를 위해 일반선형모형을 보다 확장시킨 모형이다. 다중 회귀분석에 쓰이는 일반선형모형은 독립변수와 종속변수 사이의 선형성, 오차항의 정규성, 독립성과 등분산성의 4가지를 기본 가정으로 한다. 하지만 이러한 기본 가정들이 적용되지 않는 비연속형 자료, 범주형 자료에 적용하기 위해 일반화선형모형에서는 종속변수를 적절한 함수로 변화시킨  $f(x)$ 와 독립변수를 선형 결합으로 모형화 한다. 일반화선형모형은 선형회귀분석의 많은 장점들을 그대로 수용한다. 여

러 독립변수들을 하나의 회귀식으로 설명함으로써 독립변수들 간의 영향을 모두 보정한 다변수 분석이 가능하며, 독립변수가 종속변수에 미치는 영향의 정도를 회귀계수를 이용하여 설명할 수 있다.

### 2.1.6 Stacked Ensemble

Stacked Ensemble 기법은 David H. Wolpert의 Stacked generalization[26]에서 출발한 개념으로 여러 예측 모델의 정보를 조합하여 새로운 예측모델을 생성하는 모델 앙상블 기술이다. 1계층 모델의 결과물을 고루 섞어 특정 모델의 강점을 부각하고, 특정 모델의 약점을 보완하는 Stacking 모델(2계층 모델)은 대부분의 경우 단일 1계층 모델보다 좋은 성능을 보인다. Stacked Ensemble 모델은 1계층 모델들이 서로 매우 다를 때 가장 큰 효과를 얻는다. H2O에서 제공하는 Stacked Ensemble 모델은 ALL MODEL과 Best of Family 두가지 방식의 Stacked Ensemble 모델을 제공한다. ALL MODEL 방식은 H2O AutoML에서 생성된 모든 모델을 이용하여 Stacked Ensemble 모델을 생성한다. Best of Family는 AutoML 학습시 지정된 인공지능 모델그룹(Random Forest, Extremely Randomized Tree Forest, GBM, Deep Neural Network)중 각 그룹에서 가장 좋은 성능을 나타내는 모델들을 결합하여 Stacked Ensemble 모델을 생성한다.

## 2.2 강화학습

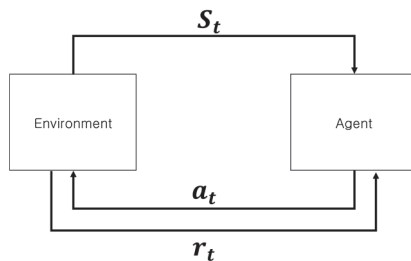


그림 6. 강화학습 환경



강화학습(Reinforcement Learning)은 환경(Environmnet)과 에이전트(Agent)의 상호작용을 통해 학습된다. 그림 6에서 에이전트는 환경으로부터 획득한 상태( $s_t$ )를 기반으로 동작 가능한 행동( $a_t$ ) 중 하나를 선택한다. 환경은 이에 대한 피드백으로 보상( $r_t$ )을 다시 에이전트에게 전달한다. 이러한 과정을 반복하면서 에이전트는 보상을 통해 보상을 최대화하도록 학습이 이루어지는 구조이다. 대부분의 강화학습 문제는 MDP(Markov Decision Process)로 표현된다.

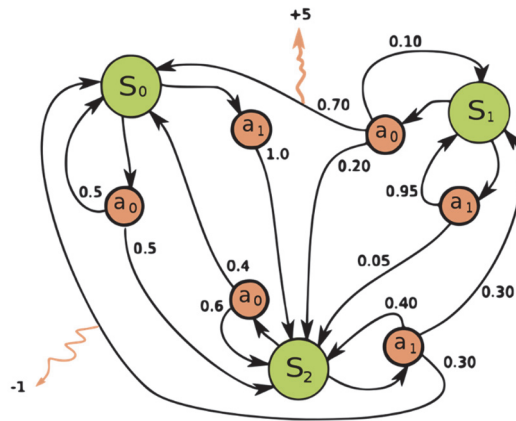


그림 7. Markov decision process

MDP는 의사 결정 과정을 수학적 모델로 나타낸 방법으로 모든 환경이 Markov Property를 가진다고 가정하였다. Markov Property는 미래 상태에 대한 조건부 확률이 과거 상태에 독립적이고 현재 상태에는 의존적인 속성을 의미한다. 이를 식으로 표현하게 되면 다음과 같다.

$$\mathbb{P}\{S_{t+1}|S_t\} = \mathbb{P}\{S_{t+1}|S_1, \dots, S_t\}$$

Markov 상태인  $s$ 와 다음 상태  $s'$ 의 상태 전이 확률(State Transition Probability)과 모든 상태에 대한 상태 전이 확률은 상태 전이 행렬(State Transition Matrix)  $P$ 로 나타낼 수 있으며 이를 식으로 표현하면 다음과 같다.

$$P_{ss'} = \mathbb{P}\{S_{t+1} = s' | S_t = s\}$$

$$P = \begin{matrix} & \text{to} \\ \text{from} & \begin{bmatrix} P_{11} & \cdots & P_{1n} \\ \vdots & & \vdots \\ P_{n1} & \cdots & P_{nn} \end{bmatrix} \end{matrix}$$

MDP는 Markov Property를 가지는 상태의 전이에 대해서 상태공간  $S$ , 행동 공간  $A$ , 보상 함수  $R$ , 할인인자  $\gamma$ 의 나열로 나타내며 다음과 같이 표현할 수 있다.

$$\langle S, A, P, R, \gamma \rangle$$

$$P = P_s^a = \mathbb{P}\{S_{t+1} = s' | S_t = s, A_t = a\}$$

$$R = R_s^a = \mathbb{E}\{R_{t+1} | S_t = s, A_t = a\}$$

$$\gamma \in [0, 1]$$

강화학습의 목표는 미래에 받게 될 보상을 최대화하는 것이다. 이를 위해서 미래에 받을 보상의 합  $G_t$ 는 아래의 식과 같이 나타낼 수 있다.

$$G_t = R_{t+1} + \gamma R_{t+2} + \cdots = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1}$$

할인 인자인  $\gamma$ 이 0에 가까우면 에이전트는 근시안적인 행동을 하게 된다. 반대로 1에 가까우면 에이전트는 원시안 적인 행동을 하게 된다. 에이전트의 행동을 결정하는 정책  $\pi$ 는 임의의 상태인  $s$ 에 대하여 어떤 행동  $a$ 를 취할 확률을 나타내며 다음과

같이 표현할 수 있다.

$$\pi(a|s) = \mathbb{P}\{A_t = a | S_t = s\}$$

정책  $\pi$ 는 보상의 합  $G_t$ 를 이용하여 평가한다. 이 값을 가치함수라고 부른다. 가치함수는 상태 가치함수  $V_\pi(s)$ 와 행동 가치함수  $Q_\pi(s, a)$ 가 있으며 다음과 같이 표현할 수 있다.

$$V_\pi(s) = \mathbb{E}\{G_t | S_t = s\}$$

$$Q_\pi(s, a) = \mathbb{E}\{G_t | S_t = s, A_t = a\}$$

강화학습의 목표를 위해 미래의 받을 보상을 최대화하는 최적의 정책  $\pi^*$ 를 찾는다. 최적의 정책  $\pi^*$ 는 모든 정책  $\pi$  중 모든 상태와 행동에 대해 최댓값을 갖는 가치함수를 의미하며 아래와 같이 표현된다.

$$V_*(s) = \max_\pi V_\pi(s)$$

$$Q_*(s, a) = \max_\pi Q_\pi(s, a)$$

최적의 가치함수를 구하기 위해 Bellman 방정식을 이용한다. Bellman 방정식은 현재 상태  $s$ 의 가치 값과 다음 상태  $s'$ 에서의 가치 값 사이의 관계를 나타내는 식이다. Bellman 방정식을 이용한 최적 가치함수를 구하는 식은 아래와 같다.

$$V_*(s) = \max_\pi R_s^a + \gamma \sum_{s' \in S} P_{ss'}^a V_*(s')$$

$$Q_*(s, a) = R_s^a + \gamma \sum_{s' \in S} P_{ss'}^a \max_{a'} Q_*(s', a')$$

이를 이용하여 최적 정책인  $\pi^*$ 을 구할 수 있다.

강화학습 문제해결은 주로 동적 프로그래밍(Dynamic Programming), 몬테 카를로 기법(Monte-Carlo Methods), 시간차 기법(Temporal Difference Methods), Q-learning등을 사용해 왔다.

동적 프로그래밍은 복잡한 문제를 해결하기 위해 여러 개의 하위 문제로 나누어 해결하는 기법이다. 동적 프로그래밍을 강화학습에 적용하기 위해 MDP의 모든 상황을 알고 있어야 하며 Bellman 최적식을 반복적으로 계산하여 최적의 정책을 도출한다.

몬테 카를로 기법은 동적 프로그래밍과 달리 MDP에 대한 사전 정보가 없어도 학습이 가능하다. 현재의 정책을 따라서 환경과 상호작용을 하게 되고 한 에피소드가 끝난 후에 얻을 보상 값을 통해 다음 학습을 수행한다. 학습이 끝나면 보상 값을 얻기 위해 도달한 모든 상태  $s$ 에 대해 가치함수를 갱신하게 된다. 이 과정을 반복적으로 수행하여 특정 정책  $\pi$ 에 대한 가치함수가 계산된다.

시간차 기법은 동적 프로그래밍과 몬테카를로 기법을 조합한 방법이다. 환경에 대한 사전 정보 없이 학습이 가능하며 학습 도중에 값을 갱신한다.

Q-learning은 행동 가치함수  $Q_\pi$ 을 이용하여 최적 정책  $\pi^*$ 을 찾는다. 행동 가치함수를 업데이트 하는 식은 아래와 같다.

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_{t+1} + \gamma \max_{a_{max}} Q(s_{t+1}, a_{max}) - Q(s_t, a_t)]$$

Q-learning이 최적 정책을 수렴하기 위해 상태 공간과 행동 공간의 모든 상태와 행동이 짝을 이루어야 하기 때문에 이를 위한 탐색 과정을 수행한다. 가장 대표적인 탐색 과정은  $\epsilon - greedy$  방법을 사용한다.  $\epsilon - greedy$  방법은 학습의 초기 단계에서는 탐색의 빈도를 높이고 학습이 진행될수록 이용의 빈도를 높여 최적 정책에 수렴하도록 유도하는 방법이다.

## 2.3 deep Q-learning

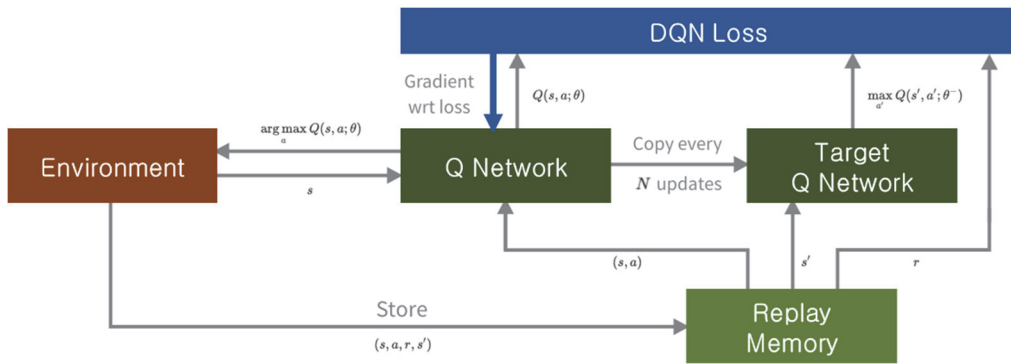


그림 8. Deep Q-learning

deep Q-learning[27] 은 Q-learning과 deep neural network가 결합된 방식이다. deep Q-learning은 행동 가치함수를 deep neural network로 근사하여 상태와 행동의 조합이 형성하는 차원의 저주(Curse of dimensionality) 문제를 해결한다. 또한 기억 메모리(replay memory)를 활용하여 불완전한 정책을 따라 수집되는 시계열 샘플들의 상관관계를 없앤다. 기억 메모리는 제한적인 경험의 수를 극복하기 위해 샘플을 랜덤으로 기억 메모리에서 추출하여 새롭고 다양한 경험을 생성하며(experience replay) 이를 통해 정책을 효과적으로 탐색 및 발전시킨다.

deep Q-learning은 deep neural network 모델 2개를 사용하여 하나는 행동 가치함수를 최대화하는 행동 결정 역할의 Q-network로 활용하고, 다른 하나는 앞의 Q-network를 업데이트하기 위해 타겟 값을 생성하는 역할을 수행하는 target Q-network이다. Target Q-network는 일정 주기로 Q-network의 가중치가 복제된다. 일정 주기로 복제하여 사용하는 이유는 충분히 학습되지 않은 Q-network를 매 업데이트 시점마다 복제한다면 Q-network는 불안정한 출력이 나타날 수 있으며, 이는 Q-network가 제대

로 업데이트 되지 않는 원인이 될 수 있다. 이러한 deep-Q learning의 강화학습 방법은 고전 게임과 바둑 등에 적용되어, 정책의 자가발전(self-learning) 가능성을 보여주었고, 수많은 경우의 수가 존재하는 다양한 상황에 대해서도 사람 수준 이상의 의사결정이 가능함이 입증되었다.

### 제 3 장 AutoML을 이용한 SOx, NOx 예측

본 장에서는 화력 발전 보일러에서 발생하는 SOx, NOx를 예측하기 위해 보일러의 운전변수를 예측하고, 변수 중요도(Variable Importance)를 기반으로 한 최적의 독립 변수를 선정방법을 평가한다.

#### 3.1 화력 발전 보일러의 운전변수 예측

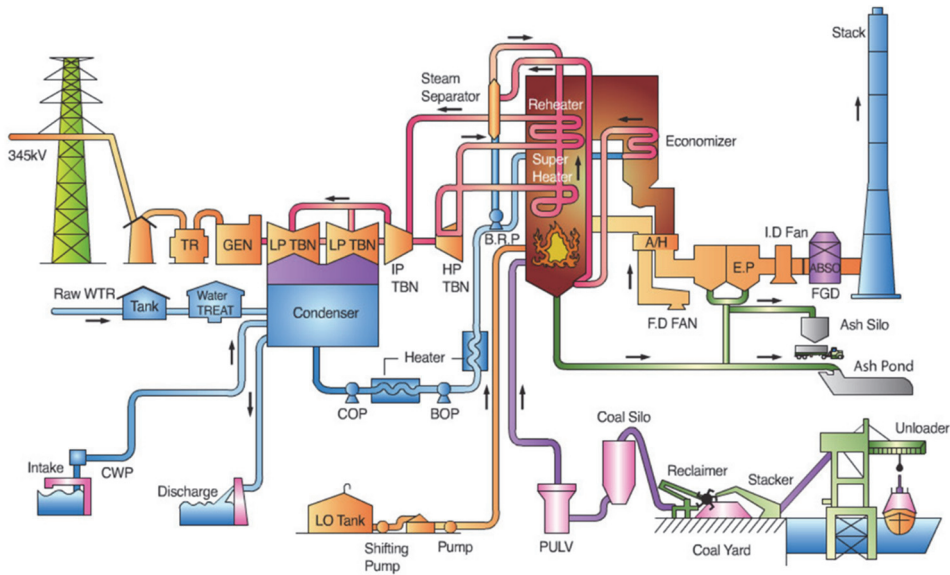


그림 9. 석탄화력발전소 계통도[28]

본 연구의 대상은 석탄화력발전소(Coal-fired Power Plant)의 보일러(Boiler)이다. 발전소의 주기기인 보일러, 터빈, 발전기 중 보일러는 연료(coal), 공기(Air), 연소가스(combustion/flue gas), 그리고 급수(feedwater)/증기(steam)로 연결되는 에너지 변환을 통해 터빈을 회전하는 고온/고압의 증기를 발생하는 장치다.

본 연구에 사용된 원천 데이터는 국내 ○○발전 ○○화력본부의 초초임계압(USC 급) ○호기에서 약 9개월간 수집된 보일러의 연료정보, 상태정보로 보일러의 운전과정에 발생하는 다양한 변수 값 들이다. 이 두 정보 값을 보일러의 운전변수로 통합해서 정의한다. 연료정보는 연소탄종 285종에 대한 상탄시간, 탄종(A,B,C,D), 성상(발열량, 원소분석, 공업분석) 값이며 상태정보는 센서데이터를 이용, 10분 간격으로 수집된 195개의 tag 값이다. 주요 변수로는 출력(MW), 미분기 차압 및 급탄량(Ton/hr), 1,2차 공기량(Ton/hr), 버너각도, 급수/증기계통 부위별 튜브 온도, 탈질설비 온도, inlet/outlet 증기온도, 과열저감수 유량 및 밸브 포지션, 탈황설비 입구 SOx 값, 탈질설비 입구 NOx 등이 있다.

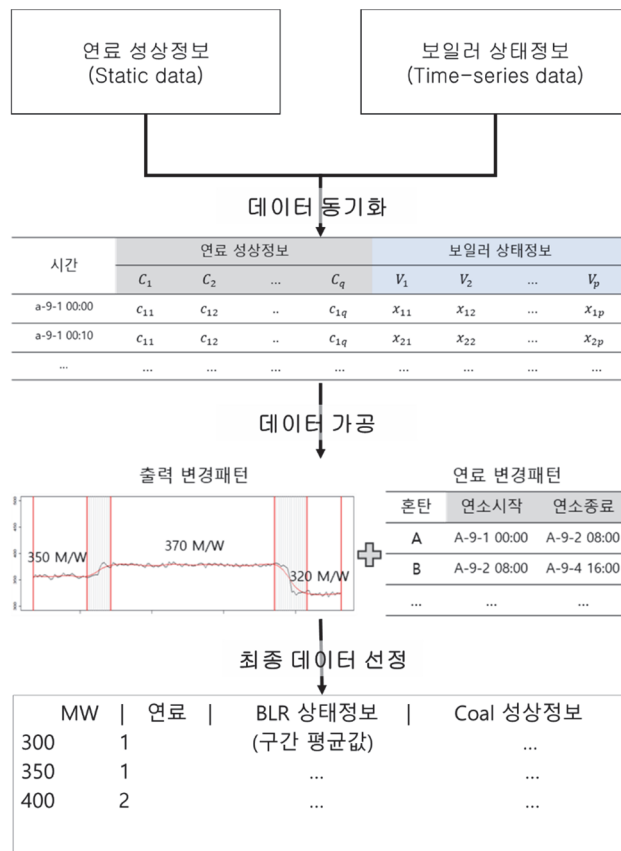


그림 10. 데이터 전처리



그림 10에서 보일러의 연료정보(정적 데이터)와 상태정보(시계열 데이터)는 두 가지 다른 유형의 데이터로 분리되어 있어 동기화 과정을 통한 하나의 학습 데이터로 통합이 필요하다. 또한 상태정보에서 수집되는 센서 데이터는 10분간격으로 수집되어 출력 변화구간(Load transient-phase)이나 연료변경 시점에서 일시적으로 보일러 상태의 변동성이 발생한다. 이를 해결하기 위해 출력 변화구간과 유지구간의 시점 정보를 파악하고 연료 변경시점 정보를 반영하여 시계열 데이터를 총 2,396 구간으로 나눈다. 나눈 구간의 상태정보 값은 구간 내 평균 값으로 적용되며, 1구간이 1개의 평균 데이터로 바뀌게 된다. 이를 통해 보일러 상태의 변동성에 따른 영향을 줄일 수 있으며 결과적으로 상태정보 데이터는 10분 단위의 25,994개에서 데이터 가공 후 2,396개의 데이터로 축소되었다.

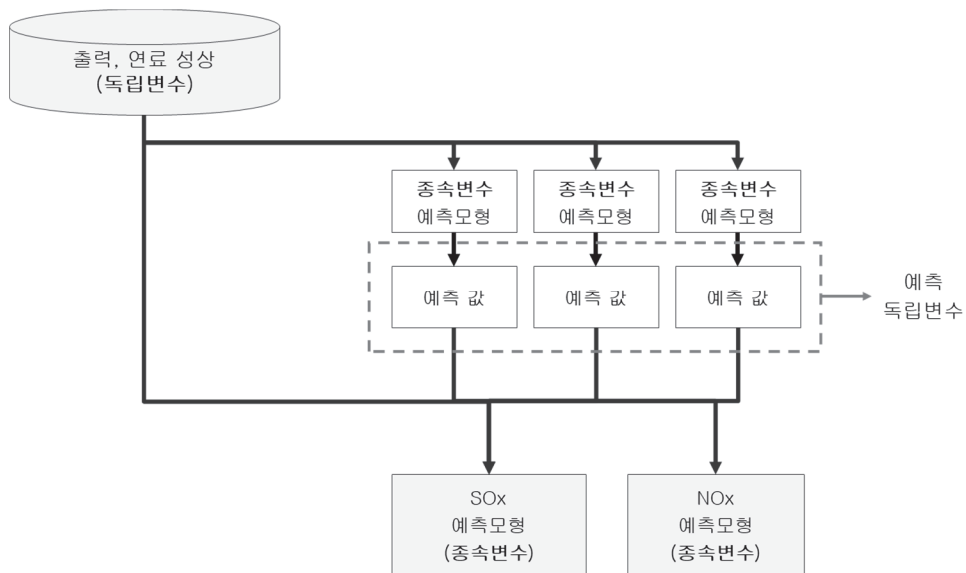
### 3.1.1 변수 중요도 기반 독립변수 정의

보일러의 운전변수인 연료정보(정적 데이터)와 상태정보(시계열 데이터)가 결합된 최종 데이터는 196개의 독립 변수로 이루어져 있다. 인공지능 모델의 학습에서 주어진 독립변수가 많을수록 예측하고자 하는 종속변수를 설명하는데 유리하다. 하지만 많은 독립 변수는 깊은 신경망 구조를 요구하여 인공지능 모델을 무겁게 만들게 된다. 또한 데이터 수집단계에서도 수집해야 하는 독립변수가 적을수록 관리와 운영에 효율적이다.

따라서 우리는 학습된 인공지능 모델에서 변수 중요도(Variable Importance)를 계산하여 최적의 독립변수를 선별한다. 여기서 최적의 독립변수의 종속 변수 대상이 되는 값은 본 연구의 주요 목표인 미세먼지 저감 대상인 SO<sub>x</sub>, NO<sub>x</sub> 수치다. SO<sub>x</sub>, NO<sub>x</sub> 수치는 탈황 설비전에 측정된 SO<sub>x</sub> 수치 값과 탈질 설비전에 측정된 NO<sub>x</sub> 수치 값으로 정의된다. 이 두 수치값은 보일러 상태정보에서 수집된 것으로 상태 정보의 데이터들은 종속 변수로 활용된다. 그 이유는 연료정보가 목표 출력을 발전하기 위해 탄종 혼소

비율, 성상정보 등의 값이 미리 정해져 있기 때문이다. 값이 미리 정해지는 정적인 연료 정보에 비해 상태 정보는 연료정보와 출력에 의해 발생하는 여러 반응들을 센서로 측정된 동적인 수치 값이다. 따라서 상태 정보는 연료정보와 출력이라는 독립변수들을 이용해 예측해야 하는 종속 변수로 정의된다.

본 연구에서는 독립변수인 연료 정보와 출력만을 이용해 종속 변수인 SOx, NOx 값을 예측해야 한다. 하지만 종속변수로 활용 되는 상태정보 중에서는 실질적인 종속변수인 SOx, NOx 와 관련성이 높은 변수들이 존재 할 것이다. 이런 상태정보 값을 이용한 최종 생성 모델의 예측력을 높이기 위한 방법을 제안한다. 이 방법의 프로세스는 **그림 11**에서 확인할 수 있다.



**그림 11. 예측 독립변수를 이용한 종속변수 예측 프로세스**

최종 목표 예측인 종속변수 성능 향상을 위하여 **그림 11**과 같은 프로세스를 따른다. 기존에 제공되는 독립변수인 출력, 연료 성상 정보 만으로 종속변수인 상태정보의 SOx, NOx 값을 예측한다. 더 나아가, 종속변수인 상태정보 중 SOx, NOx와 관련성이 높은 종속변수를 찾아 예측모형을 만들고 모델이 예측한 값을 독립변수로 재활용한다.

본 연구에서는 이 값을 예측 독립변수라고 칭한다. 독립변수인 출력, 연료 성상과 함께 예측 독립변수는 다시 독립변수로 활용하여 SOx, NOx 값을 예측해 본다. 그 후 기존 독립변수를 사용한 종속변수 예측 결과 값과 기존 독립변수와 추가 예측독립변수를 활용한 종속변수 예측 결과 값을 비교해 본다.

### 3.1.2 최적 독립변수 예측 모델 개발

본 연구에서 사용된 데이터 2,396개는 70% : 15% : 15%(1685 : 353 : 358)의 비율로 학습 : 검증 : 테스트에 사용되었다. 모델의 성능 일반화를 위해 테스트 데이터는 2396개의 데이터 중 마지막 15%의 데이터로 358개이며 이를 데이터 수집 기간인 9개월로 환산하였을 때 약 1개월 반 동안 수집된 데이터로 8월 중순부터 9월 말에 해당되는 데이터이다. 이렇게 테스트데이터를 따로 분리하여 예측한 이유는 테스트데이터가 학습에 사용될 경우 학습데이터에 오버피팅(overfitting)될 수 있기 때문이다. 나머지 학습데이터 1685개와 검증데이터 353개는 나머지 기간인 1월 ~ 8월 중순까지의 데이터를 무작위로 샘플링 하여 분할하였다.

본 연구에서 AutoML 학습을 위한 머신 러닝 환경은 아래 표와 같다.

**표 1. AI 모델 학습 실험 환경**

실험 환경	
CPU	Intel Xeon E5-2640 V4 2.4Ghz
RAM	128Gb
GPU	NVIDIA TITAN RTX(24G Memory)
OS	Windows
Programming Language	Python 2.8
Machine Learning Library	H2O 3.3

모델을 평가하는 지표는 회귀 모델의 평가에 주로 쓰이는 MSE, RMSE, RMSLE, MAE,  $R^2$  등을 사용한다. MSE는 회귀 모델의 주요 손실함수로 예측 값과 실제 값의 차이인 오차들의 제곱 평균으로 정의된다. 제곱을 하기 때문에 특이치에 민감하다. MAE는 실제 값과 예측 값의 차이인 오차들의 절댓값 평균으로 MSE보다는 특이치에 덜 민감하다. RMSE는 MSE에 루트를 씌운 값으로 오류 지표를 실제 값과 유사한 단위로 다시 변환하기 때문에 해석이 다소 용이 해진다. RMSLE는 RMSE와 다르게 실제 값과 예측 값에 + 1을 더한 후 ( $\log 0$ 를 방지) 로그 값을 취한다. 이로 인해 특이치가 있더라도 값의 변동폭이 크지 않으며  $\log$ 를 취하기 때문에 Scale을 고려한 상대적 차이를 확인할 수 있다. MSE, MAE, RMSE, RMSLE은 손실함수로 0에 가까울수록 모델 성능이 좋으며 식은 다음과 같이 나타낼 수 있다.

$$\begin{aligned}
 MSE &= \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \\
 MAE &= \frac{1}{n} \sum_{i=1}^n (|y_i - \hat{y}_i|) \\
 RMSE &= \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \\
 RMSLE &= \sqrt{\frac{1}{n} \sum_{i=1}^n (\log \hat{y}_i + 1) - \log(y_i + 1))^2}
 \end{aligned}$$

$R^2$ (R-Squared) 결정계수는 분산기반 예측 성능 평가로 모델이 얼마나 데이터를 잘 설명할 수 있는가에 대한 점수이다. 0~1 사이의 값을 가지고 1에 가까울수록 모델에 대한 설명력이 높음을 의미한다. 결정계수는 아래의 식과 같이 표현된다.

$$R^2 = 1 - \frac{RSS}{TSS}$$

여기서 총제곱합(TSS: total sum of squares)은 종속변수  $y$ 의 총분산을 측정하며 회귀

가 수행되기 전에 반응변수에 내재하는 변동량이다. 잔차제곱합(RSS:residual sum of squares)은 회귀가 수행된 후에 설명되지 않고 남아있는 변동량을 측정한다. TSS와 RSS는 아래와 같은 식으로 표현된다.

$$RSS = \sum_{l=1}^n (y_l - \hat{y}_l)^2$$

$$TSS = \sum_{i=1}^n (y_i - \bar{y})^2$$

MAE, MSE, RMSE와 같은 지표들은 데이터의 Scale에 영향을 받기 때문에 본 연구에서는 RMSE, RMSLE, R<sup>2</sup> 평가 지표를 확인한다. SOx 와 NOx 모델 예측에 관련한 변수중요도를 파악하기 위해 연료정보와 상태정보 모두를 활용하여 SOx와 NOx 예측 모델을 만들었다. AutoML의 모델 생성 개수는 최대 20개로 지정하였고 AutoML에 의해 생성된 SOx와 NOx 예측 모델 중 가장 높은 성능을 보인 모델의 예측 결과 값은 표 2와 같다.

**표 2. SOx, NOx 성능 결과 값**

종속변수	MODEL	MSE	RMSE	RMSLE	MAE	R <sup>2</sup>
SOx	Stacked Ensemble (ALLMODEL)	124.37	11.15	0.02	6.31	0.98
NOx	Stacked Ensemble (ALLMODEL)	17.74	4.21	0.04	2.88	0.93

각각의 SOx, NOx 예측 모델은 AutoML이 만든 20개의 모델을 이용한 Stacked Ensemble 모델이 최고 예측 성능을 보여주었다. SOx와 NOx의 MSE, RMSE, MAE는 각각 다른 Scale의 수치를 보여주는데 앞에서 기술한 대로 데이터의 Scale의 영향을 받기 때문이다. SOx의 모델은 0.02의 RMSLE 성능을 가지고 0.98의 R<sup>2</sup>을 가지고 있다. NOx 모델은 0.04의 RMSLE와 0.93의 R<sup>2</sup>를 나타내고 있다. NOx 모델의 성능이 좀더 떨어지지만 모든

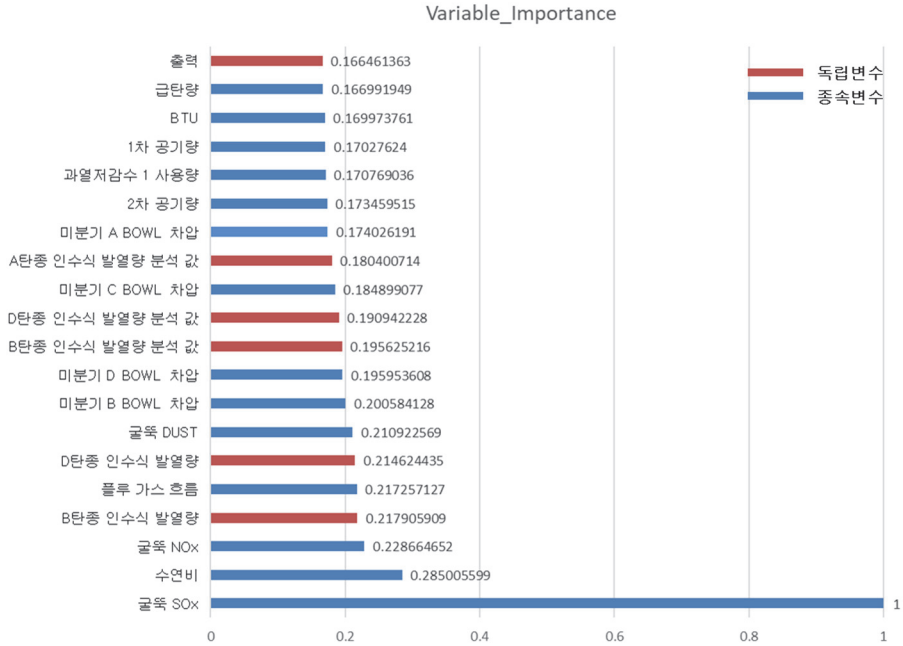
변수를 사용 하였을때 두 모델 모두 높은 성능으로 SOx, NOx를 예측하고 있음을 확인했다.

SOx, NOx의 변수 중요도를 추출하기 위해 우리는 SOx, NOx의 AutoML 모델 중 DNN 모델이나 의사 결정 나무 계열 모델을 이용하였다. 변수 중요도는 **그림12**와 같다.

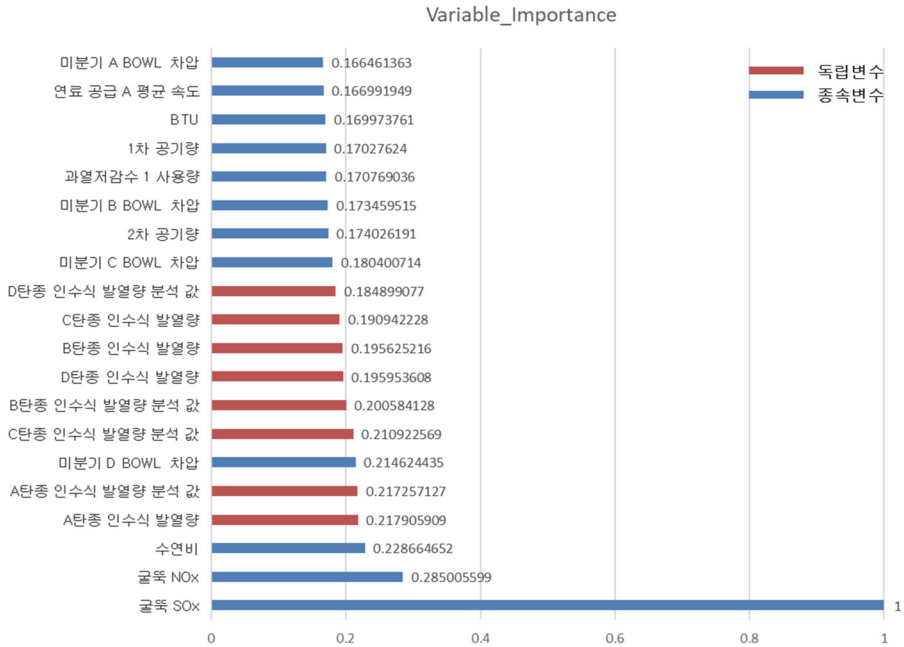
**그림 12**는 SOx와 NOx 모델을 기반으로 한 상위 20개 변수의 변수 중요도를 그래프로 나타낸 것이다. 변수 중요도에서 Y축은 각 변수들의 이름이며 X축은 변수의 중요도를 0 ~ 1로 표현하고 있으며 1에 가까울수록 영향도가 높다는 것이다. 또한 그래프의 색상은 독립변수로 사용된 변수들의 원래 변수속성을 나타낸다. 파란색 그래프는 종속변수(상태정보)들로 운전정보(독립변수)를 이용해 예측해야 하는 변수들이다. 빨간색 그래프는 보일러 운전시 주어지는 운전정보(독립변수)이다.

본 연구에서는 SOx, NOx모델 예측에 상태정보가 얼마나 영향을 주는지 파악하기 위해 상위 20개의 변수 중에서 독립변수와 종속변수가 얼마나 차지하고 있는지 확인했다. **그림 12(a)**의 SOx 예측 변수 중요도에서 SOx를 예측 하는데 영향을 주는 상위 20개의 변수 중 14개의 종속변수(상태정보)와 6개의 독립변수(운전정보)로 이루어져 있다. **그림 12(b)**의 NOx 예측에 영향을 주는 상위 20개의 변수 중 11개의 종속변수(상태정보)와 9개의 독립변수(운정정보)로 이루어져 있다. 모델 예측에 영향을 주는 변수가 두 모델 다 예측을 해야 하는 상태정보가 더 많이 분포하고 있다. 따라서 SOx와 NOx 예측 모델의 성능을 높이기 위해서는 본 연구에서 제안한 예측 독립변수를 활용한 종속변수 예측 방법(**그림 11**)을 이용해야 한다.

따라서 예측 독립변수 후보를 정하기 위해 변수 중요도를 기반으로 한 상위 20개의 변수 중에서 종속변수로서 독립변수로 활용 될 변수 후보는 14개로 **표3**에서 확인할 수 있다.



(a) SOx 모델 Variable Importance



(b) NOx 모델 Variable Importance

그림 12. 예측 모델의 변수 중요도 그래프

**표 3. 독립변수로 활용하기 위한 종속변수 후보(상태정보)**

운전변수	변수 이름	운전변수	변수 이름
상태정보	굴뚝 SOx 수치	상태정보	수연비
상태정보	굴뚝 NOx 수치	상태정보	BTU
상태정보	굴뚝 DUST 수치	상태정보	과열저감수 1 사용량
상태정보	미분기 A BOWL 차압	상태정보	1차 공기량
상태정보	미분기 B BOWL 차압	상태정보	2차 공기량
상태정보	미분기 C BOWL 차압	상태정보	연료 공급 A 평균 속도
상태정보	미분기 D BOWL 차압	상태정보	급탄량

예측 독립변수를 종속변수로 예측하기 위해 사용되는 독립변수는 보일러의 운전정보이다. 운전정보의 변수들 또한 변수 중요도 과정에서 영향력이 낮은 변수들을 제거 하였다. 제거 후 독립변수로 사용된 65개의 변수는 표 4를 통해 확인할 수 있다.

**표 4. 예측에 사용된 독립변수(연료성상 + 출력)**

운전변수	변수 이름	변수 수
상태정보	출력	1
운전정보	A,B,C,D 탄종 혼탄비율	4
운전정보	A,B,C,D 탄종 인수식 발열량 분석값	4
운전정보	A,B,C,D 탄종 총수분	4



운전정보	A,B,C,D 탄종 유황분	4
운전정보	A,B,C,D 탄종 분쇄도	4
운전정보	A,B,C,D 탄종 휘발분	4
운전정보	A,B,C,D 탄종 회분	4
운전정보	A,B,C,D 탄종 고정 탄소	4
운전정보	A,B,C,D 탄종 원소 탄소	4
운전정보	A,B,C,D 탄종 원소 질소	4
운전정보	A,B,C,D 탄종 회성분 SiO <sub>2</sub>	4
운전정보	A,B,C,D 탄종 회성분 CaO	4
운전정보	A,B,C,D 탄종 착화온도	4
운전정보	A,B,C,D 탄종 회융점 IDT	4
운전정보	A,B,C,D 탄종 입도	4

표 4의 독립변수를 이용하여 표 3의 종속변수 후보들을 예측 하였다. 종속변수 후보의 결과 값은 표5와 같다.

**표 5. 예측독립변수 후보 모델 성능 결과**

종속변수	MODEL	MSE	RMSE	RMSLE	MAE	R <sup>2</sup>
1차 공기량	Stacked Ensemble (BestOfFamily)	39.67	6.29	0.01	3.37	0.98
급탄량	Stacked Ensemble (AllModels)	11.8	3.43	0.02	2.53	0.97

2차 공기량	Stacked Ensemble (AllModels)	523.76	22.88	0.01	16.98	0.97
미분기 A BOWL 차압	GBM	730.79	26.99	0.03	16.74	0.84
미분기 C BOWL 차압	Stacked Ensemble (BestOfFamily)	523.51	22.88	0.07	16.06	0.83
BTU	Stacked Ensemble (BestOfFamily)	7.71	2.77	0.03	2.12	0.77
수연비	Stacked Ensemble (AllModels)	7.71	2.77	0.03	2.12	0.77
미분기 B BOWL 차압	Stacked Ensemble (AllModels)	294.22	17.15	0.07	12.67	0.7
미분기 D BOWL 차압	Stacked Ensemble (BestOfFamily)	341.39	18.47	0.07	13.74	0.67
굴뚝 DUST 수치	Stacked Ensemble (BestOfFamily)	0.07	0.27	0.05	0.14	0.54
과열저감수 1 사용량	DeepLearning_grid	2.15	1.44	0.002	0.62	0.53
굴뚝 SOx 수치	Stacked Ensemble (BestOfFamily)	0	3.12	0.1	2.12	0.52
연료 공급 a 평균속도	DeepLearning_grid	0.8	0.35	0.06	1.12	0.34
굴뚝 NOx 수치	Stacked Ensemble (AllModels)	27.92	5.28	0.15	3.84	0.1

표 5에 따르면 예측독립변수 후보의 결정계수( $R^2$ )의 성능에 따라 내림차순으로 모델을 정리 하였다. 모델 성능은 최소 0.1부터 0.98까지의 크게 차이를 보였다. 특히, 그림 12에서 변수 중요도가 높게 나온 굴뚝 NOx 수치, 굴뚝 SOx 수치는 각각  $R^2$  가 0.1, 0.52로 모델의 예측 성능이 매우 떨어지는 것을 확인할 수 있었다. 본 연구에서는 예측독립변수로 사용되는 종속변수의 성능이 매우 우수해야 예측의 오류를 줄일 수 있고, 최소화된 오류는 예측독립변수로 예측되는 종속변수의 결과 값에 영향을 미치기 때문에  $R^2$ 의 성능이 0.9 이상인 1차공기량, 2차공기량, 급탄량 모델을 최종 예측독립변수로 사용했다.

### 3.2 예측 독립변수를 이용한 SOx, NOx 예측

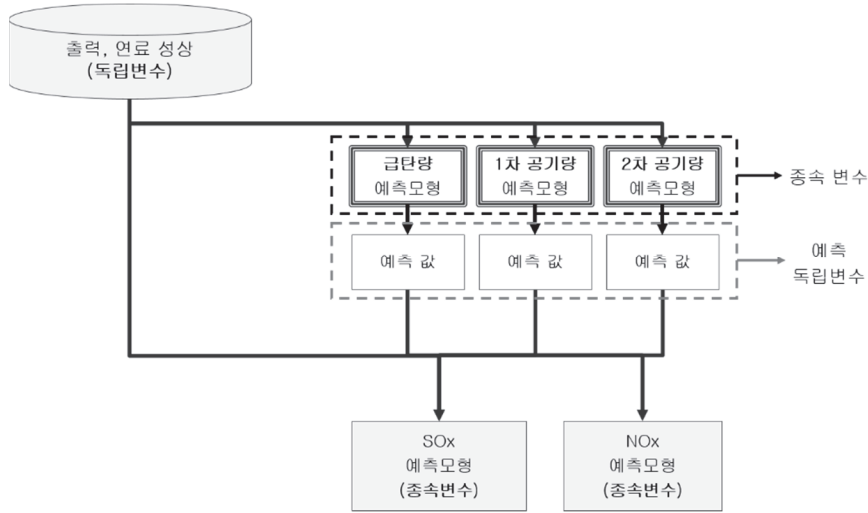


그림 13. SOx, NOx 예측 모델 프로세스

본 연구의 주요 목표 중 하나인 미세먼지 운전변수인 SOx, NOx를 예측하기 위해 앞 절에서 1차공기량, 2차공기량, 급탄량 예측 독립변수 모델을 생성했다. 예측 독립변수가 정해져 그림11은 그림13와 같이 업데이트 된다. 본 절에서는 예측 독립변수를 이용한 성능 개선을 확인하기 위해 표 4에서 정의한 연료성상+출력 변수를 이용한 SOx, NOx 예측 성능과 연료성상+출력 변수 와 예측독립변수(상태정보의 1차공기량,2차공기량, 급탄량)를 이용한 SOx, NOx 예측 성능을 비교한다. AutoML을 이용한 SOx, NOx 예측 성능은 표 6과 같다.

**표 6. 독립변수 그룹 별 SOx, NOx 성능 비교**

	연료성상+출력	연료성상+출력 +예측독립변수	연료성상+출력	연료성상+출력 +예측독립변수
종속변수	SOx	SOx	NOx	NOx
MODEL	Stacked Ensemble (BestOfFamily)	Stacked Ensemble (BestOfFamily)	Stacked Ensemble (BestOfFamily)	Stacked Ensemble (BestOfFamily)
MSE	1512.47	<b>1296.87</b>	39.25	<b>29.33</b>
RMSE	38.89	<b>36.01</b>	6.26	<b>5.41</b>
RMSLE	0.09	<b>0.08</b>	0.06	<b>0.05</b>
MAE	26.36	<b>24.35</b>	4.49	<b>3.90</b>
R <sup>2</sup>	0.78	<b>0.83</b>	0.85	<b>0.89</b>

표 6에 따르면 모든 모델이 Stacked Ensemble(BestOfFamily)모델에서 최고 성능을 나타냈다. 또한, 예측 독립변수를 활용할 경우 SOx, NOx 예측 모델 모두 모델 성능이 향상됨을 확인했다. 결정 계수 기준으로 SOx의 경우 0.78에서 0.83으로 성능이 향상되었고 NOx의 경우 0.85에서 0.89로 향상되었다. 표 2에서 SOx, NOx는 모든 변수를 활용했을 경우에 0.98과 0.93의 높은 성능을 보여주었지만 이는 모든 운전정보와 상태정보(195개의 변수)를 사용해야하는 제약이 있다. 하지만 본 연구의 예측독립변수를 활용한다면 195개의 변수에서 68개의 변수만을 이용하여 예측할 수 있다. 이는 약 65%의 독립변수 감소 효과를 가져오며, 그림 12에서 확인할 수 있듯이 상태정보에서 굴뚝 SOx 수치와 굴뚝 NOx 변수에 의존적이기 때문에 두 변수로 인한 모델 성능의 과대 평가를 고려한다면 예측 독립변수를 이용한 SOx, NOx 예측 모델은 최적 독립변수를 이용하여 모델이 학습되었음을 확인할 수 있다.

## 제 4 장 deep Q-learning을 이용한 최적 혼소조합 제어

본 장에서는 미세먼지 저감을 위해 보일러의 운전변수 중 하나인 석탄의 혼소비율 조합을 제어하기 위해 deep Q-learning을 이용한다. 본 연구는 deep Q-learning의 Environment에 SOx, NOx 예측 모델을 결합 한다. SOx, NOx 예측 모델은 상태 업데이트에 쓰이며 미세먼지 지수인 SOx, NOx 값을 보상으로 활용한다. 보상 획득 조건에서 주탄종과 보조 탄종 비율 또는 고열량탄과 저열량탄의 비율 등을 고려한다. 검증을 위해 특정 날짜의 혼소 조합에서 최적의 조합을 제시하고 기존의 혼소 조합 비율과 비교해본다.

### 4.1 deep Q-learning 구성 요소 정의

#### 4.1.1 상태

상태 공간은 현재 탄종의 혼소비율과 혼소비율에 따른 SOx, NOx 예상 수치이다. 앞 장에서 개발한 SOx, NOx 예측 모델은 연료성상 정보와 출력만 주어지게 된다면 SOx와 NOx 수치를 예측할 수 있다. 탄종에 따라서 연료성상 정보는 동일하기 때문에 학습을 위해서 A,B,C,D 탄종과 출력을 고정하여 학습한다. 최적 혼소조합 제어 모델의 상태 정보는 표7과 같다.

표 7. 강화학습을 위한 혼소조합 제어 모델의 상태

구분	상태
$s_1(t)$	고열량 탄 A 비율
$s_2(t)$	저열량 탄 B 비율
$s_3(t)$	고열량 탄 C 비율
$s_4(t)$	저열량 탄 D 비율
$s_5(t)$	SOx
$s_6(t)$	NOx

#### 4.1.2 행동

본 연구에서 미세먼지 저감을 위해 최적 혼소 조합을 찾아야 한다. 탄종 A,B,C,D의 혼소 비율에 따라 조합이 바뀌기 때문에 본 연구에서의 행동은 각 탄종의 비율에 대한 증가와 감소 행위라고 정의 하였다. 각 탄종에 대해서 5%로 증가하고 감소하는데 증감 크기가 클 경우 혼소 비율의 조합 범위가 작아져 빠르게 학습되지만 혼소조합이 단순해져 미세먼지 저감의 최적 비율을 찾지 못할 수 있다. 반대로 증감의 크기가 작을 경우 혼소조합이 복잡해져 미세먼지 저감의 최적비율을 찾을 수 있지만 많은 학습시간이 소요되거나 모델의 학습이 이루어지지 않을 수 있다. 행동은 표8과 같이 정의된다.

표 8. 강화학습을 위한 혼소조합 제어 모델의 행동

구분	행동	단위	값
$a_1(t)$	고열량 탄 A 비율 증가	%	5
$a_2(t)$	고열량 탄 A 비율 감소	%	5
$a_3(t)$	저열량 탄 B 비율 증가	%	5
$a_4(t)$	저열량 탄 B 비율 감소	%	5
$a_5(t)$	고열량 탄 C 비율 증가	%	5
$a_6(t)$	고열량 탄 C 비율 감소	%	5
$a_7(t)$	저열량 탄 D 비율 증가	%	5
$a_8(t)$	저열량 탄 D 비율 감소	%	5

### 4.1.3 보상

보상은 SOx와 NOx 수치 값의 합으로 정의하였다. SOx, NOx 수치 값의 합이 최소가 되도록 학습을 진행하였으며 이를 위해 SOx와 NOx의 수치 값을 음수로 변환하였다. 주력 혼탄종인 탄종 A와 B의 비율이 50% 이상이 되면 보상을 주었다. 또한, 저열량탄의 비율을 늘리기 위해 탄종 C와 D의 비율이 최소 30% 부터 시작하여 70%에 가까워질 수록 보상을 증가시켰다. 보상은 다음과 같이 표현된다.

$$Reward = \begin{cases} [s_5(t) + s_6(t)] * 0.5, & \text{if } s_1(t) + s_2(t) \geq 50 \\ [s_5(t) + s_6(t)] * 0.9, & \text{if } 40 \leq s_2(t) + s_4(t) \leq 50 \\ [s_5(t) + s_6(t)] * 0.7, & \text{if } 50 \leq s_2(t) + s_4(t) \leq 60 \\ [s_5(t) + s_6(t)] * 0.5, & \text{if } 60 \leq s_2(t) + s_4(t) \leq 70 \end{cases}$$

## 4.2 deep Q-learning 기반 최적 혼소조합 제어 알고리즘

본 논문에서 제안한 deep Q-learning 기반 최적 혼소조합 제어 알고리즘은 그림 14와 같이 구성하였다.

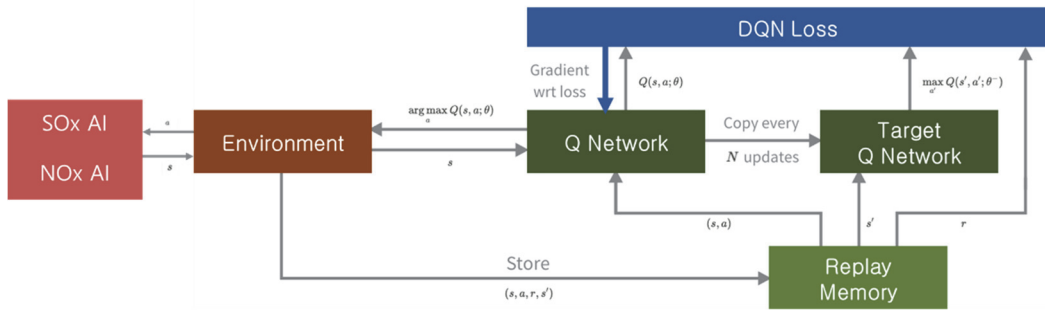


그림 14. deep Q-learning 기반 최적 혼소조합 제어 구조

그림 14에서 Q Network 및 Target Q Network 외에도 본 연구에서는 환경에서 State를 업데이트하기 위해 인공지능 모델을 이용한다. 다만 이 인공지능 모델은 3장에서 만든 SOx, NOx 예측 모델로 미리 학습된 인공지능 모델이다. 이 모델은 단순히 환경 내에서 정해진 행동에 따른 상태 값을 업데이트 하는데 사용된다.

### 4.2.1 수행방법

최적 혼소조합의 학습을 위해 우리는 4.1.1 절에서 탄종을 고정하여야 한다고 서술했다. 따라서 SOx, NOx 예측 학습에 사용되었던 테스트 데이터 중 가장 많이 사용된 탄종 혼소 조합을 기반으로 학습을 진행하였다. 강화 학습에 사용된 머신 러닝 환경은 표 9와 같다.



**표 9. 강화학습 모델 학습 실험 환경**

실험 환경	
CPU	Intel Xeon E5-2640 V4 2.4Ghz
RAM	128Gb
GPU	NVIDIA TITAN RTX(24G Memory)
OS	Windows
Programming Language	Python 2.8
Machine Learning Library	H2O 3.3 KERAS 2.4.3 Tensorflow - GPU 2.3.1

또한 학습을 위해 설정한 강화학습 Parameter는 표 10과 같다.

**표 10. 강화 학습 파라미터**

Parameter	Value
Learning rate	0.0001
Discount factor	0.99
Replay memory	5000
Mini batch	32
Episode	2000
Target Q Network Update Cycle	50

학습은 2000 에피소드(Episode)에서 수행되며, Target Q Network는 50에피소드 마다 가중치를 업데이트 하게 하였다. 리플레이 메모리는 총 5000 개의 값을 저장할 수 있으며 5000개 이상부터 새로운 값이 들어오면 가장 먼저 들어온 값을 삭제하게 된다. 학습에는 32개의 임의의 샘플(Mini batch)를 뽑아 강화학습 모델을 학습시킨다.

### 4.2.2 수행결과

그림 15는 주어진 학습환경에서 강화학습이 정책 발전을 통해 보상을 최대화한 값이다.

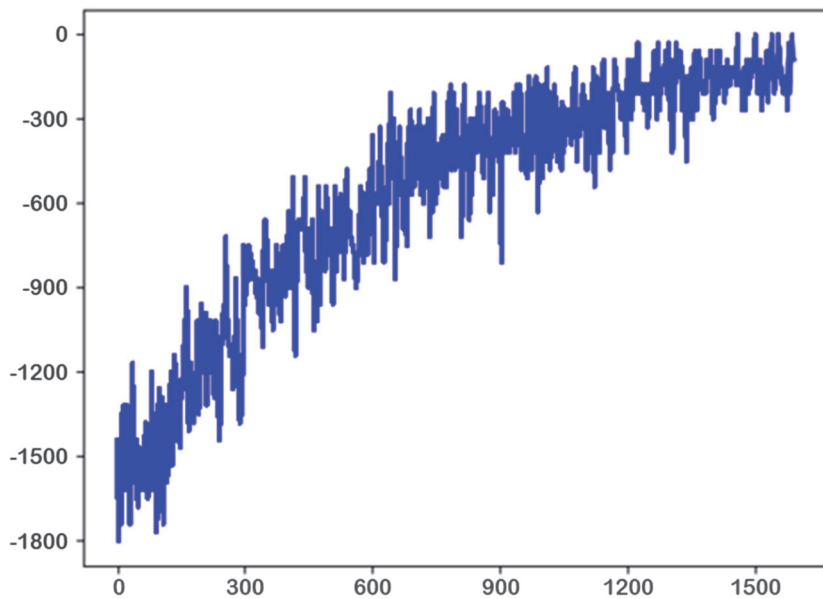


그림 15. 강화학습 정책 발전 결과

그림 15에서 X축은 진행된 에피소드를 말하며 Y축은 보상 값을 의미한다. 4.1.3절

에서  $S0x$ ,  $N0x$  수치를 최소화 하기위해 보상을 마이너스로 변환하였기 때문에 음수의 보상 값에서 시작한 강화학습은 보상은 정책이 발전할수록 0에 가까이 수렴하는 모습을 볼 수 있다. 학습은 100번째 에피소드에서 시작되었으며,  $\epsilon - greedy$  기법을 이용하여 학습 초기에는 임의의 행동을 취하도록 하여 상태와 행동에 대해 탐사 (Exploration) 과정을 거치도록 하였다. 학습이 진행될수록  $\epsilon$  값이 점점 줄어들어 각각의 상태에 따른 최적의 행동을 취하게 되어 보상 값이 점점 증가하는 것을 확인 할 수 있었다.

강화학습 모델을 평가하기 위해서 강화학습 모델에 사용된 탄종 혼소 조합 중 각각 다른 날짜 3일에 대해 강화학습 기반의 혼소 최적화를 진행하였다. 혼소 최적화 결과는 표 11에서 확인 할 수 있다.

**표 11. 강화학습 기반 최적 혼소조합 제어**

	Day1	Day1 (강화학습 최적화)	Day2	Day2 (강화학습 최적화)	Day3	Day3 (강화학습 최적화)
고열량탄 A 비율	36	20	24	20	28	20
저열량탄 B 비율	24	40	16	30	12	30
고열량탄 C 비율	26	20	39	20	39	20
저열량탄 D 비율	14	20	21	30	21	30
$S0x$	469	<b>423</b>	460	<b>412</b>	430	<b>412</b>
$N0x$	87	<b>81</b>	85	<b>80</b>	92	<b>80</b>

표 11은 각기 다른 날짜에 사용된 동일 탄종의 혼소조합을 강화학습 모델을 통해 혼소 최적화를 진행한 결과이다. Day1의 경우 강화학습을 적용할 경우 AB의 혼소비율은 동일 하지만 BD의 혼소비율이 22% 증가하였고  $S0x$  수치감소 또한 약 10%가 감소,  $N0x$  수치는 약 7% 감소하였다. Day2의 경우 AB의 혼소비율은 10% 증가하였고 BD의 혼

소비율이 23% 증가하였다. SOx 수치감소는 약 10%가 감소되었고 NOx 수치는 약 5% 감소하였다. Day 3의 경우 AB의 혼소비율은 10% 증가하였고 BD의 혼소비율이 27% 증가하였다. SOx 수치감소는 약 4%가 감소되었고 NOx 수치는 약 13% 감소하였다. 3일의 평균적인 AB의 혼소 비율 증가는 약 3%가 증가하였고 BD의 혼소 비율은 약 24%가 증가하였다. SOx 수치는 평균적으로 8%가 감소되었으며 NOx 수치 또한 평균적으로 8.3%가 감소하였다.

표 11의 결과 값에 따르면 AB, CD 혼소 비율이 보상에서 설정한 제약조건을 만족시키고 있으며 SOx와 NOx 수치를 감소시키는 걸로 보아 강화 학습의 목표인 SOx, NOx 최소화 라는 정책을 학습 모델이 정책 수렴에 잘 반영한 것으로 확인 된다. 결과적으로 본 연구를 통해 개발된 최적 혼소조합제어 모델이 화력 발전소에서 미세먼지 저감을 위해 충분히 적용될 수 있음을 확인하였다.

## 제 5 장 결론 및 토의

본 논문은 화력발전소에서 수집된 운전변수 정보인 연료정보와 상태정보를 기반으로 SOx와 NOx 예측 모델을 만들고 deep Q-learning 과 연동하여 미세먼지 저감을 위한 최적 운전변수를 예측하는 방법론을 제안하였다. 특히 본 논문에서는 SOx, NOx 예측 모델을 기반으로 한 강화학습 모델을 통해 종속변수가 어떤 독립변수 값에 의해 최적화가 될 수 있는지 확인할 수 있었다. 이는 최근 산업현장에서 자동제어 최적화 알고리즘에 유용하게 쓰일 수 있다는 점에서 의미가 크다.

SOx, NOx 예측 모델을 개발하기 위해 변수 중요도를 기반으로 최적 독립변수를 선정하는 과정을 거쳤다. 예측에만 쓰이는 종속변수들을 다시 독립변수로 활용하기 위해 종속변수 중 SOx, NOx 예측에 높은 중요도 값을 가지는 값들을 분류하고 독립변수들을 이용하여 종속변수를 개별 예측하였다. 그 중 높은 성능 값을 가지는 일부 종속변수를 다시 예측 값으로 활용하는 예측 독립변수를 만들고 기존의 독립변수와 예측 독립변수를 이용하여 SOx, NOx 예측 시 성능이 증가하는 것을 확인하였다. 이러한 방법을 통해 누락된 변수나 부족한 변수가 있는 산업 현장에서 모델의 성능을 끌어 올리기 위해 예측 값이 다시 입력 값으로 활용 될 경우 모델의 성능을 보완할 수 있을 것으로 기대된다.

또한 SOx, NOx 예측 모델을 deep Q-learning 과 연동시켜 강화학습 내의 환경 안에서 상태를 업데이트 하는 용도로 사용하였다. 상태 안에서 혼소조합이라는 액션에 따라 SOx, NOx는 예측 값으로 상태를 업데이트 하고 이 두 상태를 기반으로 보상을 반환하였다. 보상 정책은 SOx, NOx를 최소화하는 행동을 탐색하였고 학습이 완료된 강화학습을 통해 탄중 혼소조합을 최적화 해본 결과 주탄 사용량이 평균적으로 3% 증가하였고 저열량탄의 사용 비율을 평균적으로 약 24% 증가시켰다. 반대로 SOx 수치는 평균적으로 8%가 감소되었으며 NOx 수치 또한 평균적으로 8.3%가 감소되어 화력발전 보일러 운영에서 경제적인 측면과 환경적인 측면을 동시에 고려할 수 있었다.

하지만 개발된 최적 혼소조합 제어 모델은 단일 탄종 혼소조합에 최적화되어 있어 화력 발전 보일러를 실시간으로 최적 운영 하기 위해서는 다양한 탄종의 혼소조합에 대해 최적화를 할 수 있어야 한다. 이러한 문제를 해결하기 위해 향후 연구로는 deep-Q learning의 상태와 행동 구조를 개선하여 다양한 탄종의 혼소조합을 최적화할 수 있는 모델을 설계하고자 한다.

## 참고문헌

1. P. Madejski, "Thermal Power Plants : New Trends and Recent Developments", IntechOpen, Poland, 2018, pp. 1-11.
2. 환경부(2017). 미세먼지 관리 종합대책
3. 박진수, 최진수, 김현재, 전하은, 성민영, 이용환, 안준영, 이동원, 서영교, 서석준, 박승명, 오준, 강경식, 박윤서, 강소영, 이상보, 김정수 (2107). 미세먼지(PM2.5) 2차 생성량 추정에 관한 연구( I ). 국립환경과학원.
4. 관계부처합동(2019). 제5차 국가환경종합계획(2020~2040)
5. 백강수(2017). 발전부문 미세먼지 저감을 위한 노력과 대책. 한국동서발전(주). POLICY & ISSUES. Conducted at [http://www.epa.or.kr/kor/upload/magazine/201703/02161\\_01.pdf](http://www.epa.or.kr/kor/upload/magazine/201703/02161_01.pdf)
6. [ 최종원. (2018). 세 가지 미세먼지(PM/SOx/NOx)를 동시 저감하기 위한 All-in-one 시스템 개발. 설비저널, 47(10), 106-111.
7. Gu, Y., Zhao, W., Wu, Z., 2011. Online adaptive least squares support vector machine and its application in utility boiler combustion optimization systems. J. Process Contr. 21 <https://doi.org/10.1016/j.jprocont.2011.06.001>.
8. Lv, Y., Liu, J., Yang, T., Zeng, D., 2013. A novel least squares support vector

machine ensemble model for NO<sub>x</sub> emission prediction of a coal-fired boiler. Energy. <https://doi.org/10.1016/j.energy.2013.02.062>.

9. Lv, Y., Hong, F., Yang, T., Fang, F., Liu, J., 2017. A dynamic model for the bed temperature prediction of circulating fluidized bed boilers based on least squares support vector machine with real operational data. Energy 124. <https://doi.org/10.1016/j.energy.2017.02.031>.

10. Li, G.Q., Qi, X. Bin, Chan, K.C.C., Chen, B., 2017. Deep bidirectional learning machine for predicting NO<sub>x</sub> emissions and boiler efficiency from a coal-fired boiler. Energy Fuels 31. <https://doi.org/10.1021/acs.energyfuels.7b01415>.

11. Shi, Y., Zhong, W., Chen, X., Yu, A.B., Li, J., 2019. Combustion optimization of ultra supercritical boiler based on artificial intelligence. Energy 170. <https://doi.org/10.1016/j.energy.2018.12.172>.

12. Tuttle, J.F., Vesel, R., Alagarsamy, S., Blackburn, L.D., Powell, K., 2019. Sustainable NO<sub>x</sub> emission reduction at a coal-fired power station through the use of online neural network modeling and particle swarm optimization. Contr. Eng. Pract. 93 <https://doi.org/10.1016/j.conengprac.2019.104167>.

13. Tan, P., Xia, J., Zhang, C., Fang, Q., Chen, G., 2014. Modeling and optimization of NO<sub>x</sub> emission in a coal-fired power plant using advanced machine learning methods. Energy Proc. 61 <https://doi.org/10.1016/j.egypro.2014.11.1129>.

14. Tang, Z., Zhang, Z., 2019. The multi-objective optimization of combustion system operations based on deep data-driven models. Energy 182. <https://doi.org/10.1016/j.energy.2019.06.051>.



15. Cheng, Y., Huang, Y., Pang, B., Zhang, W., 2018. ThermalNet: a deep reinforcement learning-based combustion optimization system for coal-fired boiler. Eng. Appl. Artif. Intell. 74 <https://doi.org/10.1016/j.engappai.2018.07.003>.
16. V. Mnih, K. Kavukcuoglu, D. Silver, A.A. Rusu, J. Veness, M.G. Bellemare, A. Graves, M. Riedmiller, A.K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, D. Hassabis, Human-level control through deep reinforcement learning, Nature, 518 (7540) (2015) 529-533.
17. LeDell, Erin, and S. Poirier. "H2o automl: Scalable automatic machine learning." Proceedings of the AutoML Workshop at ICML. Vol. 2020. 2020.
18. J. Bergstra, Y. Bengio, "Random search for hyper-parameter optimization," J. Mach. Learning Research, vol. 13, Feb. 2012, pp. 281-305.
19. [https://en.wikipedia.org/wiki/Random\\_forest](https://en.wikipedia.org/wiki/Random_forest)
20. L. Breiman, "Random Forests," Machine Learning, Vol.45, pp.5-32, 2001. DOI: 10.1023/A:1010933404324
21. Geurts, Pierre, Damien Ernst, and Louis Wehenkel. "Extremely randomized trees." Machine learning 63.1 (2006): 3-42.
22. Friedman, Jerome H. "Greedy function approximation: a gradient boosting

machine." *Annals of statistics* (2001): 1189–1232.

23. Ke, Guolin, et al. "Lightgbm: A highly efficient gradient boosting decision tree." *Advances in neural information processing systems*. 2017.

24. Prokhorenkova, Liudmila, et al. "CatBoost: unbiased boosting with categorical features." *Advances in neural information processing systems*. 2018.

25. Chen, Tianqi, and Carlos Guestrin. "Xgboost: A scalable tree boosting system." *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. 2016.

26. Wolpert, David H. "Stacked generalization." *Neural networks* 5.2 (1992): 241–259.

27. V. Mnih, K. Kavukcuoglu, D. Silver, A.A. Rusu, J. Veness, M.G. Bellemare, A. Graves, M. Riedmiller, A.K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, D. Hassabis, Human-level control through deep reinforcement learning, *Nature*, 518 (7540) (2015) 529–533.

28. <https://www.kepco-enc.com/portal/contents.do?key=1442>, 한국전력기술