

센서 측정값과 AI를 활용한 BOG 처리 관리 방안

KNP팀

2022.12.12

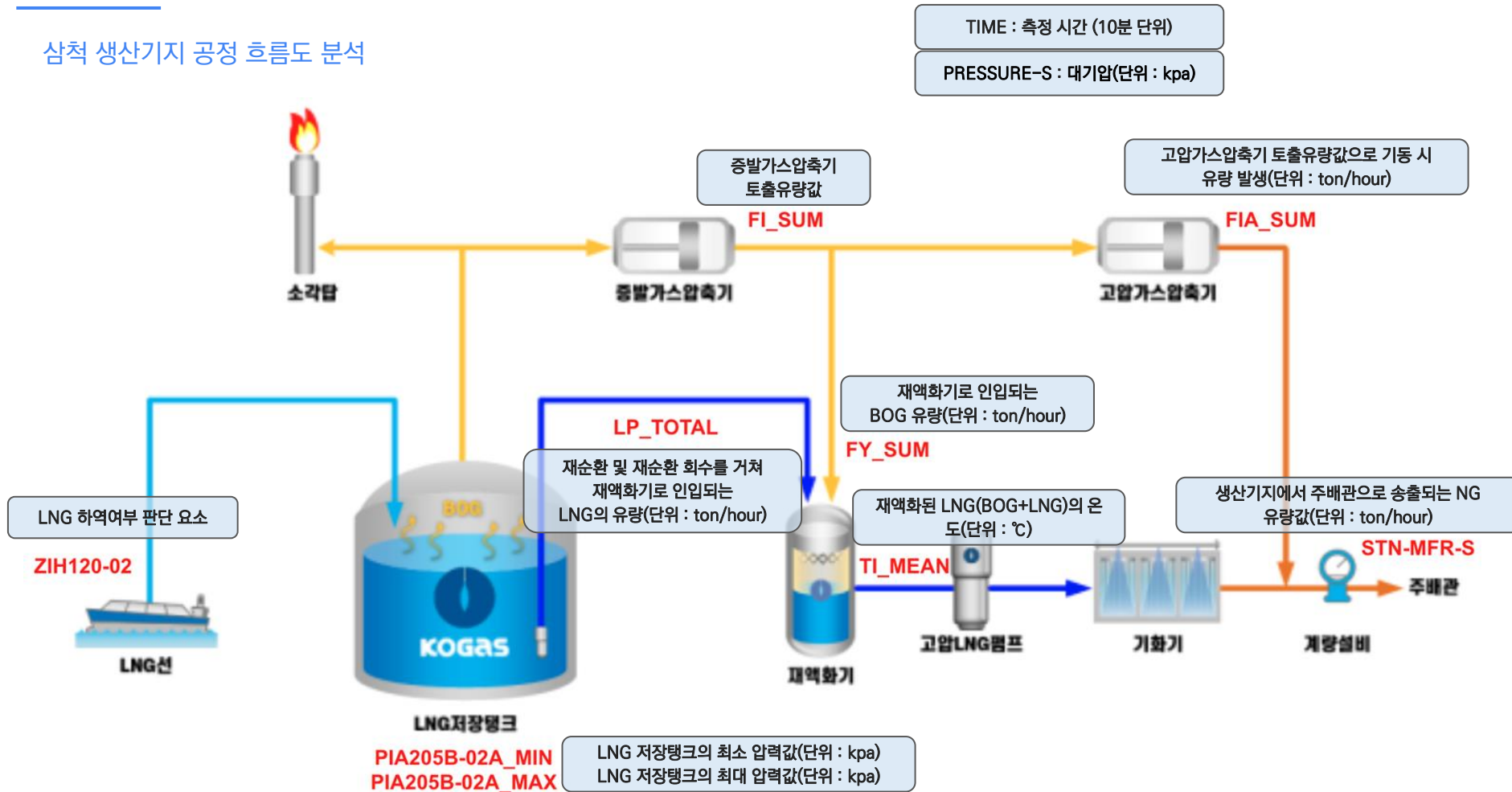
가스 압력 상승의 위험성

하역과정에서 처리량 이상으로 LNG를 주입 시 가스 압력이 상승
갑자기 저장탱크의 전체 압력이 상승하면 그만큼 가스폭발의 위험도 높아짐
이에따라 위험 관리를 위해 BOG 가스 압력의 예측이 필요

이한원 과장 | 한국가스공사 설비운영팀
LNG선 하역할 때는 하역 초기에
탱크 압력이 많이 상승을 하거든요
그래서 이상 사태가 발생할 수도 있는데...

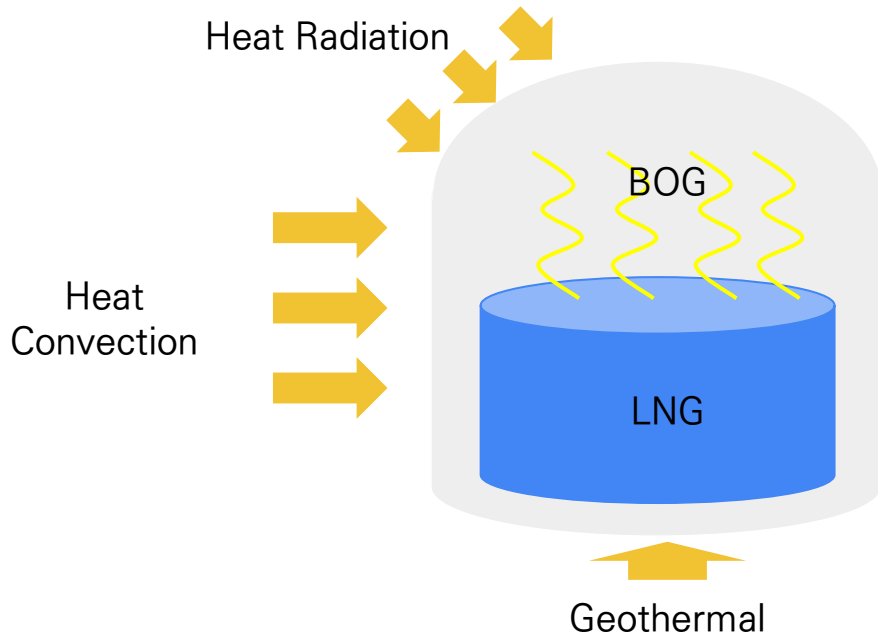
장지호 | 한국가스공사 설비운영팀
배에서 초기 하역 준비를 할 때
펌프를 일정량 돌려야 하는데
한 번에 너무 많이 돌려서
처리할 수 있는 양을 초과해서 들어오는 바람에
탱크 안에서 압력이 상승했습니다

삼척 생산기지 공정 흐름도 분석



BOG 발생량 연관 인자

하역 이외에도 BOG 발생량은 열의 전도, 복사, 대류와 상관관계가 있음
공급 기지가 있는 지역의 평균 온도, 풍량, 일사량 등은 열해석에 영향을 미침



김현승, (2018). 계절 변화에 따른 LNG 저장 Tank 내부 LNG상태변화 및 BOG 발생량 예측 (석사학위, 전북대학교 기계시스템공학과). dCollection.

기상 데이터

삼척 생산 기지와 최근접 종관기상관측(ASOS) 지점인 울진

기상청 OpenAPI를 통해 종관기상관측 데이터 활용 가능

종관기상관측(ASOS) - 자료

▪ 자료설명

자료설명

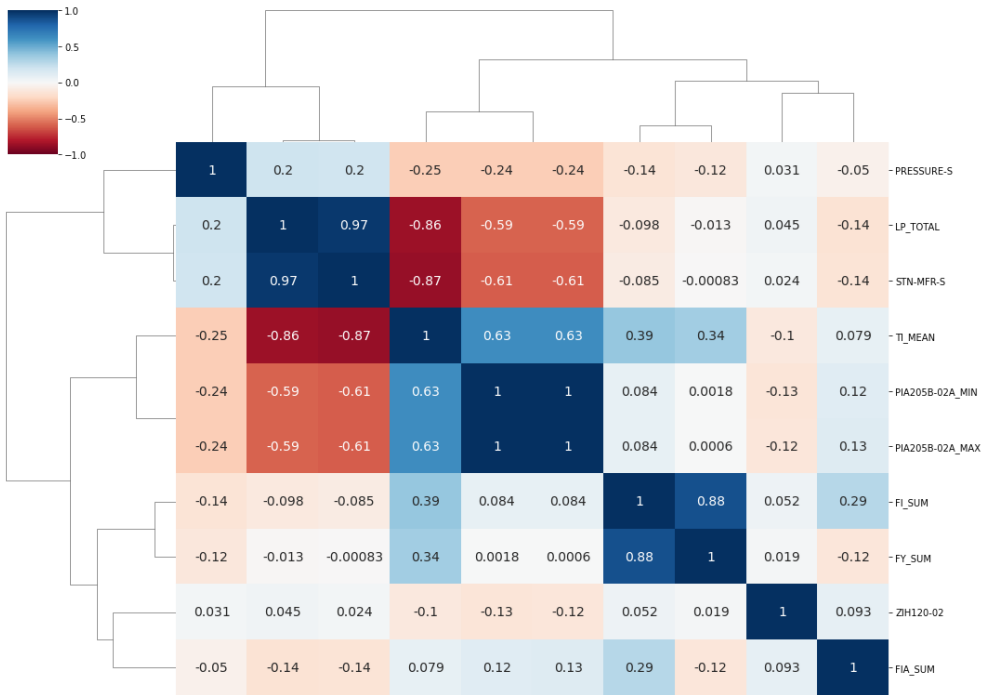
종관기상관측이란 종관규모의 날씨를 파악하기 위하여 정해진 시각에 모든 관측소에서 같은 시각에 실시하는 지상관측을 말합니다.
종관규모는 일기도에 표현되어 있는 보통의 고기압이나 저기압의 공간적 크기 및 수명을 말하며, 주로 매일의 날씨 현상을 뜻합니다.

자료형태	분, 시간(매정시), 일, 월, 연	제공기간	1904년~(지점별, 요소별 다름)
제공지점	103개 * 원하는 지점이 없는 경우, 방재기상관측(AWS) 메뉴 이용	제공요소	기온, 강수, 바람, 기압, 습도, 일사, 일조, 눈, 구름, 시정, 지면상태, 지면·초상온도, 일기현상, 증발량, 현상번호
유의사항	- 1회 조회 가능 최대 기간: 분 1일, 시간 1년, 일 10년, 월 연 제한 없음(장기간 자료는 '파일셋 조회' 메뉴 이용) - 시간/분 자료에 대해 관측값의 정상 여부를 판단하는 품질검사 플래그(QC FLAG) 정보 제공 * 제공 요소: 기온, 습도, 기압, 지면온도, 풍향, 풍속, 일조 / 플래그 종류(의미): 0(정상), 1(오류), 9(결측) - 전일 자료는 당일 10시 이후 확인 가능		
비고	- 10분 또는 1시간 최다강수시각은 최다강수가 나타난 시작 시간으로, (-) 표기가 있는 경우 전날을 뜻함 - 강수량은 겨울철(11월~익년 3월) 3시간 간격으로 제공		
지침	요소별 관측방법이나 자료 산출방식에 대한 상세 설명은 ☞ [지상기상관측지침] 참조		

1시간 간격의 기상데이터 사용
기온(°C), 지면온도(°C), 현지기압(hPa),
해면기압(hPa), 일조(hr), 풍속(m/s)

측정 데이터셋 분석

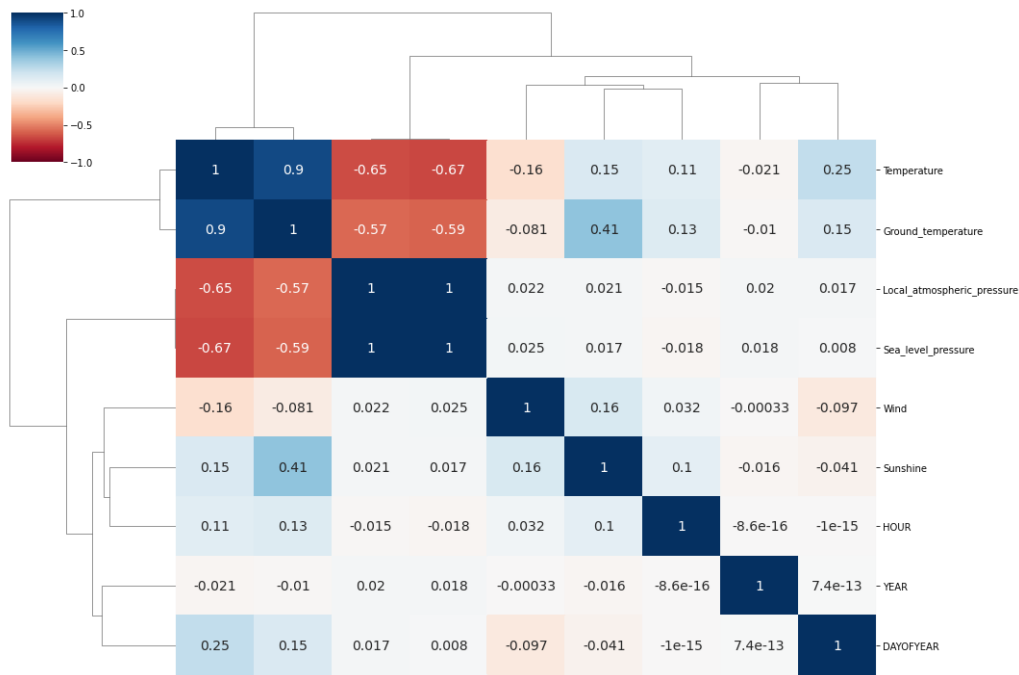
측정 데이터를 속성에 따라 분류



설명	열 이름	단위
측정 시간값	TIME	무차원 변수
0/1 binary 값	ZIH120-02	
연속값	FI_SUM	
압력	PIA205B-02A_MIN	kpa
	PIA205B-02A_MAX	
	PRESSURE-S	
온도	TI_MEAN	℃
유량	LP_TOTAL	ton/hour
	FIA_SUM	
	STN-MFR-S	
	FY_SUM	

기상 데이터셋 분석

기상 데이터를 속성에 따라 분류



설명	열 이름	단위
일조량	Sunshine	Hour
압력	Local_atmosphere_pressure	hPa
	Sea_level_pressure	
온도	Temperature	℃
	Ground_temperature	
풍속	Wind	m/s

파생변수 생성

측정 데이터와 기상 데이터를 바탕으로 파생 변수 생성

측정 데이터셋 기반 파생변수

PIA205B-02A_DIFF	탱크 내부 압력 최솟값 최대값 차이
PRESSURE_MAX_DIFF	대기압과 탱크 압력 최대값 차이
BOG	BOG가스 유량
TI_SUM	재액화기로 인입되는 유량
OUTLET_SUM	계량설비로 인입되는 유량
TI_ACC	주배관 송출 유량과 계량설비로 인입되는 유량의 차이
TI_P_MAX	재액화된 LNG 절대 온도를 탱크 내부 최대 압력으로 나눈 값 ($PV = nRT \Rightarrow T/P = V/nR$)
TI_VOL_MAX	재액화된 LNG의 부피 상관값 ($T/P \times n = V/nR \times n = V/R$)
...	

기상 데이터셋 추가 파생변수

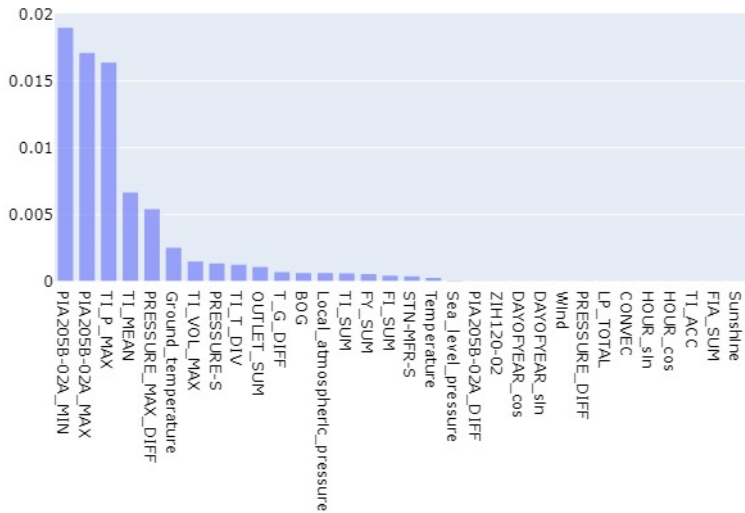
PRESSURE_DIFF	공정 데이터의 대기압과 기상관측데이터의 대기압이 차이
TI_T_DIV	탱크 내부 온도와 기온의 차이
T_G_DIFF	지표 온도와 기온의 차이
CONVEC	탱크 내외부 기온차 x 풍속 (대류 상관값)
DAYOFYEAR_sin	날짜를 1월 15일 기준으로 삼각함수 변환(연중 최저/최고 기온 날짜)
DAYOFYEAR_cos	
HOURL_sin	시간을 12시 24분을 기준으로 삼각함수 변환(속초의 태양 남중시간)
HOURL_cos	
...	

Feature 중요도 탐색

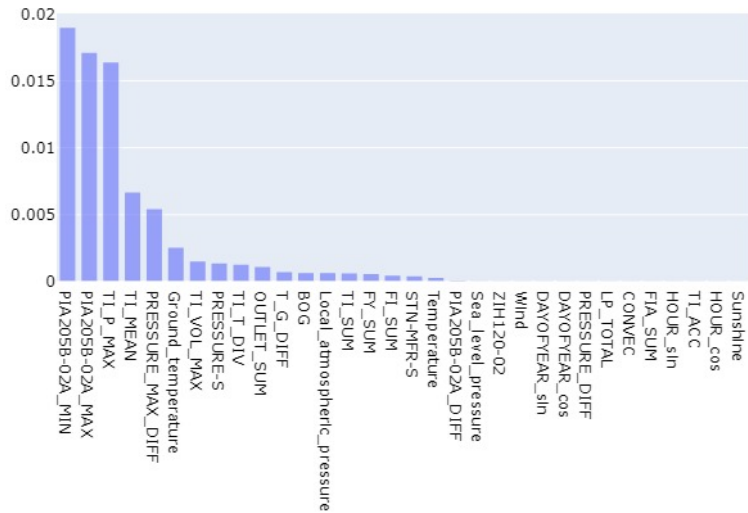
XAI(설명 가능한 AI)인 SHAP을 통해

변수 중요도가 높은 파생변수 탐색

저장 탱크의 최대 압력값 SHAP



저장 탱크의 최소 압력값 SHAP



Feature 중요도 탐색

최종적으로 32개 변수를 통해 예측 모델 개발

측정 데이터 (10개)

ZIH120-02
FI_SUM
PIA205B-02A_MIN
PIA205B-02A_MAX
PRESSURE-S
TI_MEAN
LP_TOTAL
FIA_SUM
STN-MFR-S
FY_SUM

기상 데이터 (6개)

Sunshine
Local_atmosphere_pressure
Sea_level_pressure
Temperature
Ground_temperature
Wind

측정 데이터셋
기반 파생변수 (8개)

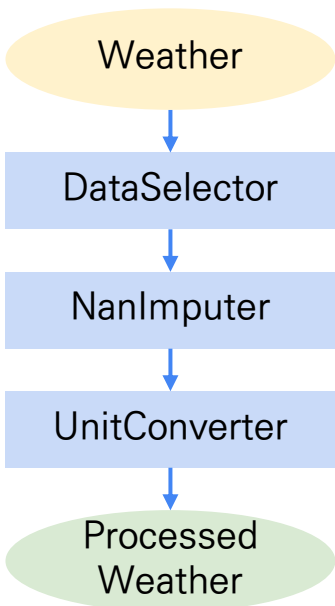
PIA205B-02A_DIFF
PRESSURE_MAX_DIFF
BOG
TI_SUM
OUTLET_SUM
TI_ACC
TI_P_MAX
TI_VOL_MAX

기상 데이터셋
추가 파생변수 (8개)

PRESSURE_DIFF
TI_T_DIV
T_G_DIFF
CONVEC
DAYOFYEAR_sin
DAYOFYEAR_cos
HOURL_sin
HOURL_cos

데이터 파이프라인 구축 - 기상데이터

울진의 2019년부터 2021년까지의
기상 Raw 데이터를 파이프라인을 통해 가공



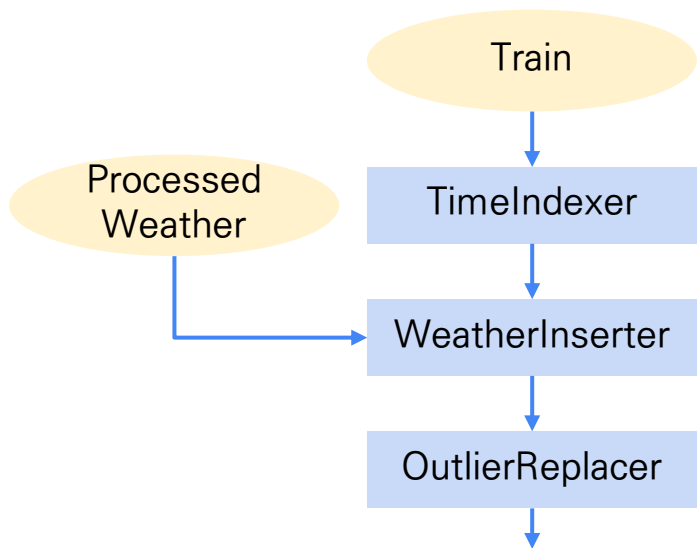
DataSelector : 컬럼 선택 및 시간 컬럼 생성

NanImputer : 결측치를 이전 데이터로 대체

UnitConverter : 압력/온도 단위 변환

데이터 파이프라인 구축 - 측정 데이터(1/2)

측정 Raw 데이터를 파이프라인을 통해 가공



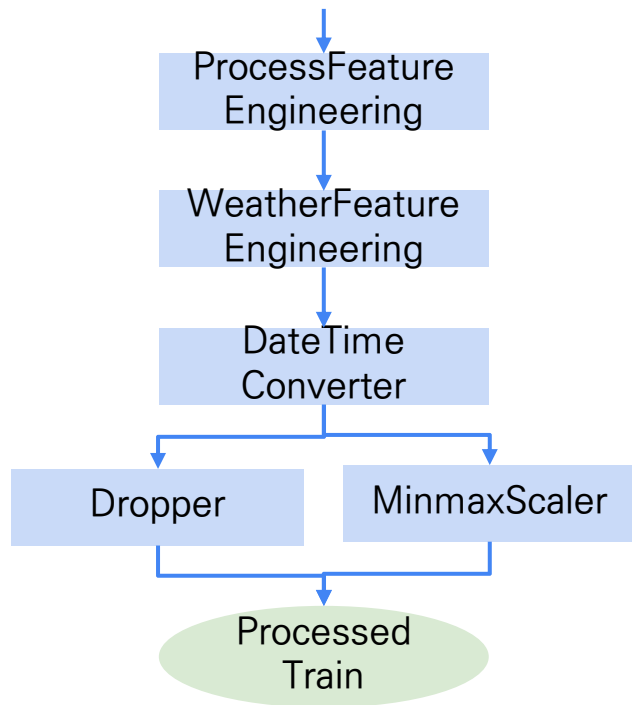
TimeIndexer : 시간 컬럼 생성

WeatherInserter : 공정 데이터와 기상관측데이터 병합
병합 시 정각 시각의 기상데이터 사용

OutlierReplacer : 공정 데이터의 대기압과 기상관측데이터의
대기압이 차이가 2.5kPa 이상일 시
대기압을 기상관측데이터로 대체

데이터 파이프라인 구축 - 측정 데이터(2/2)

측정 Raw 데이터를 파이프라인을 통해 가공



ProcessFeatureEngineering : 측정 데이터 기반 파생변수 생성

WeatherFeatureEngineering : 기상 데이터셋 추가 파생변수

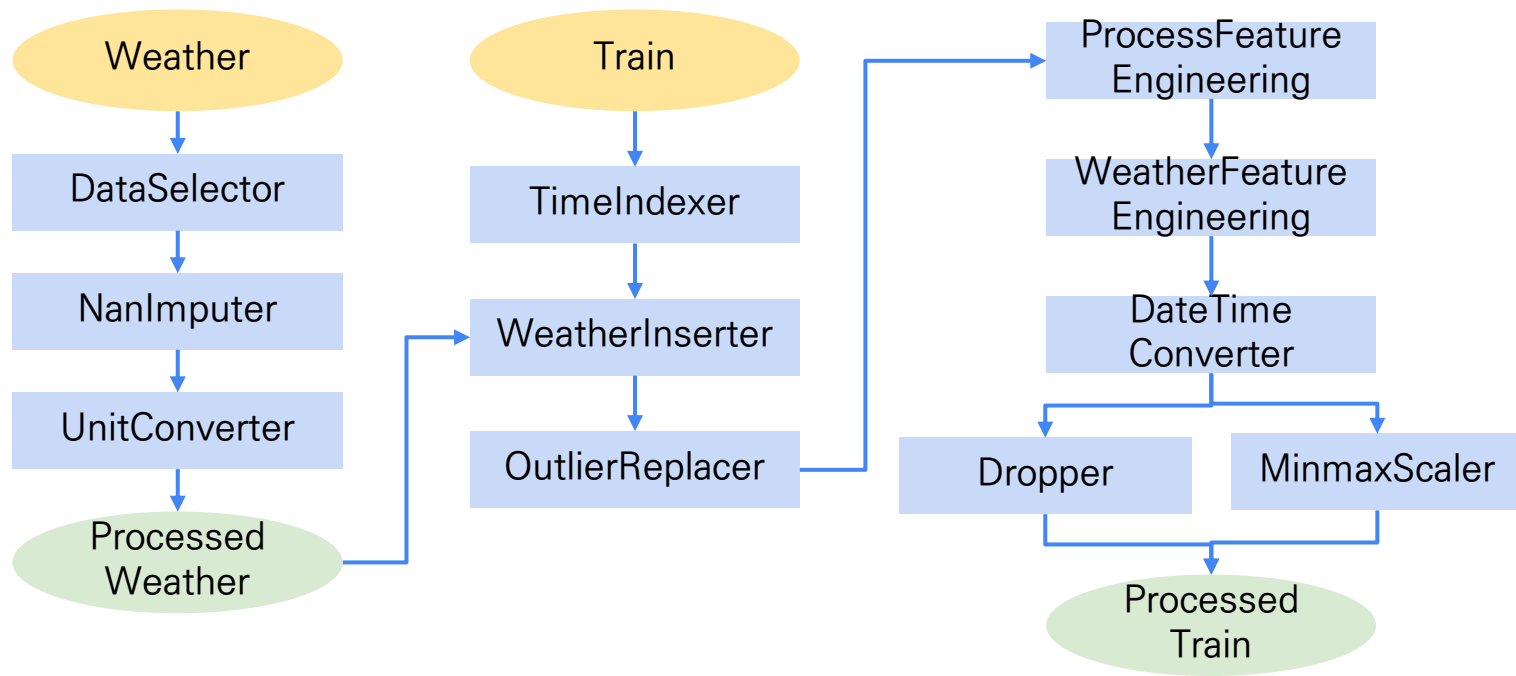
DateTimeConverter : 날짜/시간 파생변수 생성

Dropper : 필요 없는 컬럼 제거

MinmaxScaler : 데이터 0~1 범위로 스케일링

데이터 파이프라인 구축

파이프라인을 통해 Raw 데이터를 Processed 데이터로 가공



파이프라인 검증

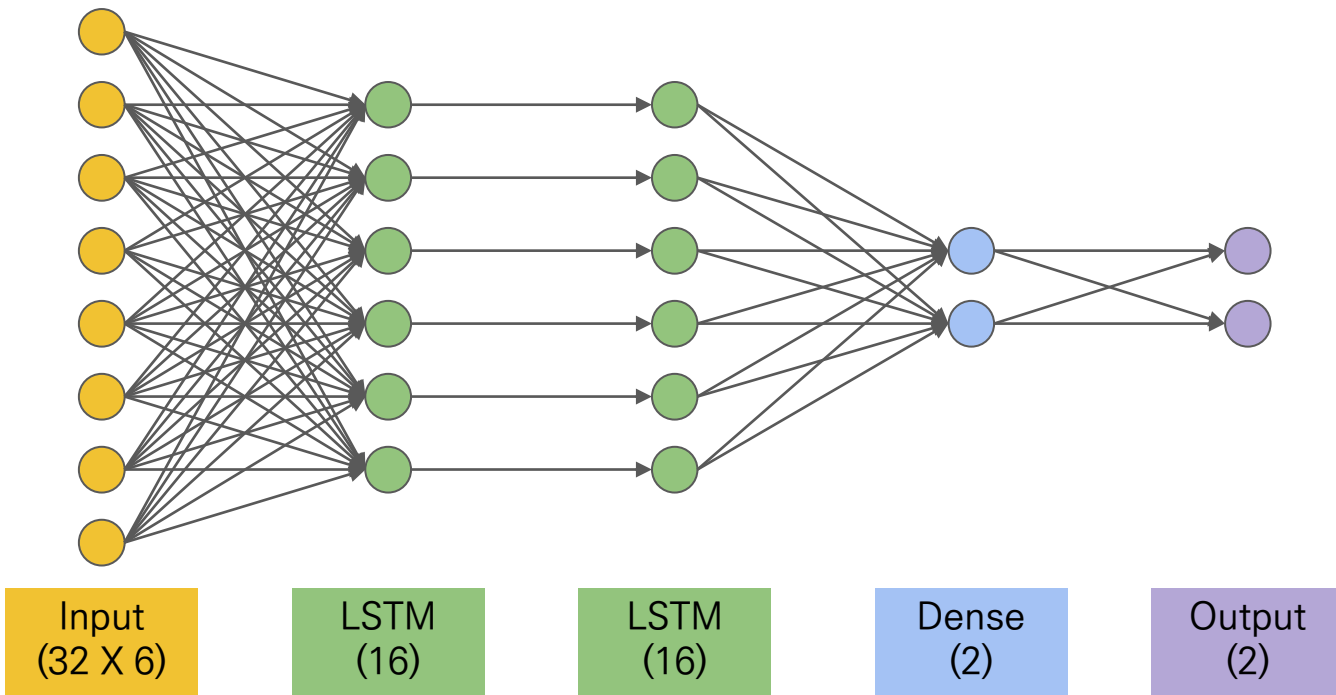
파이프라인 ON/OFF를 통해 성능 향상 검증

최초 스코어 0.026053에서 0.021848로 향상

<div><div><div></div><div></div><div></div><div></div></div></div>		X	X	X	X	O	O	O	O	ProcessFeature Engineering
		X	X	O	O	X	X	O	O	WeatherFeature Engineering
		X	O	X	O	X	O	X	O	DateTime Converter
X	X	0.026053	0.025904	-	-	0.039295	0.027714	-	-	
X	O	-	-	-	-	-	-	-	-	
O	X	0.026731	0.0263	0.026337	0.025777	0.022846	0.023866	0.034606	0.024016	
O	O	0.025716	0.027028	0.026258	0.025783	0.026218	-	0.025995	0.021848	
Weather Inserter	Outlier Replacer									

딥러닝 모델 구성

Pytorch를 통해 딥러닝 모델 구성



Custom Loss 설정

예선의 평가지표는 MAE였지만, 본선의 평가지표는 WMAE로 변경
학습을 위한 WMAE Custom Loss 설정

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - x|$$

평가지표로는 탱크 압력의 최소, 최대의 평균
MAE(Mean Absolute Error)가 사용

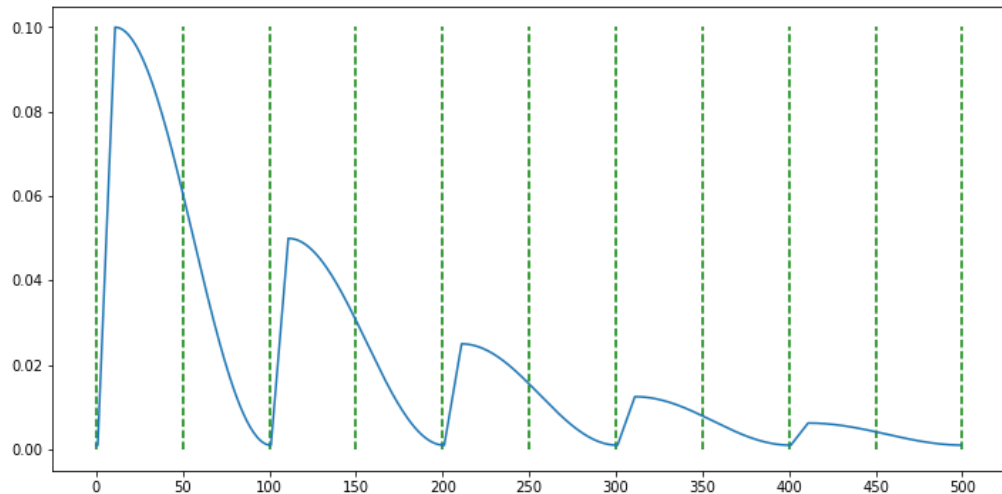
$$WMAE = \frac{1}{\sum w_i} \sum_{i=1}^n w_i |x_i - x|$$

본선 대회 MAE 점수는 10분 전 시점 대비 변화량이
큰 시점의 탱크압력 예측 정확도에 영향을 받도록 계산

Optimizer 설정

간단하고 대부분의 상황에서 효율적인 최적화 알고리즘 Adam을 사용

Overfitting을 회피하고 Generalize 하기 위해 SGDR를 사용



Stochastic Gradient Descent with Warm Restarts

Learning Rate가 주기적으로 반복되며 Learning Rate의 max값이 감소함

bag of tricks for image classification에서 사용한 방법으로 좋은 성능을 보임

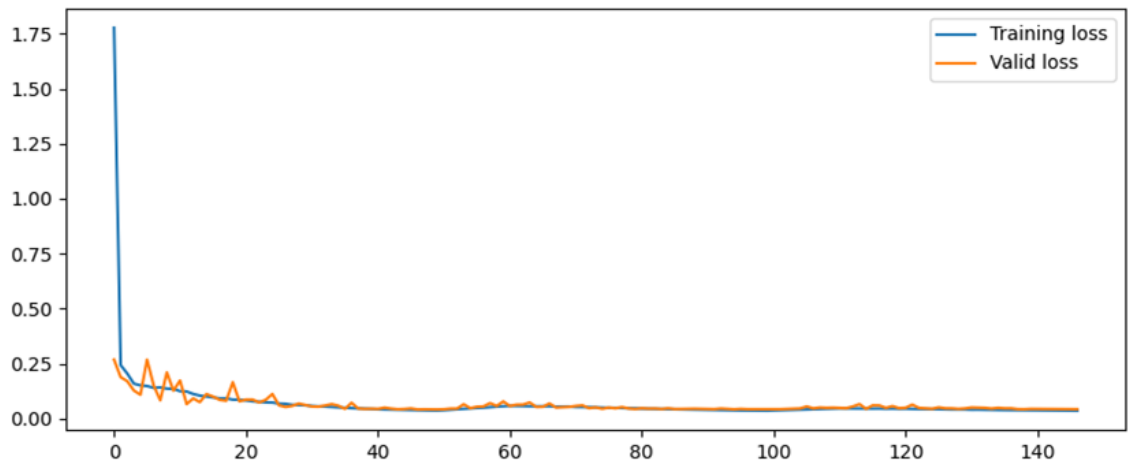
<https://arxiv.org/abs/1608.03983>

<https://arxiv.org/abs/1812.01187>

모델 학습

데이터 Train 80%, Test 20%로 분할하여 학습

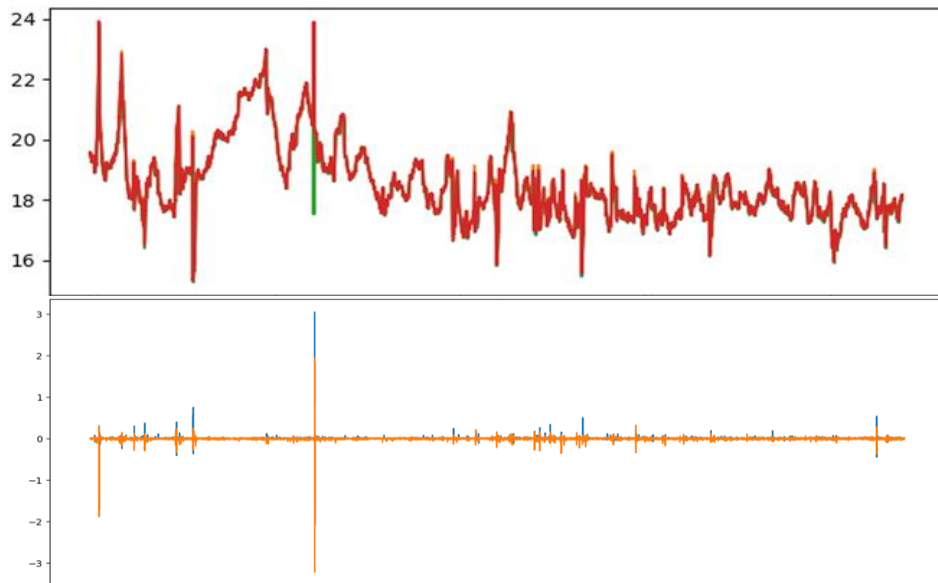
Early Stopping을 통해 Epoch 95 전후로 학습 모델 저장



딥러닝 모델의 노드 수,
은닉층 수, 활성화 함수,
데이터 분할 비율, Loss 종류,
Optimizer, Batch Size,
Learning Rate 등을
Grid Search를 활용하여 최적화

결과 분석

일반 상황에서 예측 정확도가 높으나
급작스러운 탱크 압력 변화 시 정확도 낮아짐



로컬 점수 - 0.021375

공개 점수 - 0.02465

최종 점수 - 0.02839

동작속도 및 모델 자원

88,000 길이의 데이터 학습에 걸리는 시간은 16분 내외
5,500개 예측에 걸리는 시간 37초

시스템 환경	사양
CPU	Intel(R) Xeon(R) CPU E5-2690 v4 @ 2.60GHz
GPU	Tesla V100
RAM	112GB
언어	Python 3.8
빌드 환경	MS VSCODE

총 학습 변수의 수	
변수 분류	개수
측정 데이터	10
기상 데이터	6
측정 파생변수	8
기상 파생변수	8

동작 속도	
분류	소요 시간
학습	15m 46s
예측 (5,508개)	37s
예측 (1개)	6.7ms

MLops 프로세스 설계

주어진 학습데이터로 만든 모델을 MLflow를 통해 ver1으로 배포
LIVE 데이터와 TFDV 라이브러리를 통해 Data Drift와 Skew 검증

→ LIVE Data : 실제 공정에서 나오는 가장 최신의 데이터

발생 가능한 문제

1. Data drift
2. Feature skew
3. Distribution skew



TFDV를 통해 문제 발생 시 발견하여
LIVE 데이터와 학습데이터를 합치거나
혹은 LIVE 데이터로만 모델 재학습

MLflow로 ver2로 다시 배포,
지속적으로 TFDV로 검증

MLops 프로세스 설계

TFDV를 통해 문제 발견 및 모델 재학습

MLflow로 ver2로 재배포

