Anomaly detection in industrial environment

Goal

높은 검출 성능, 빠른 학습 및 판단속도, 가벼운 모델크기

데이터셋

- 사출성형기 AI 데이터셋
- 용해탱크 AI 데이터셋

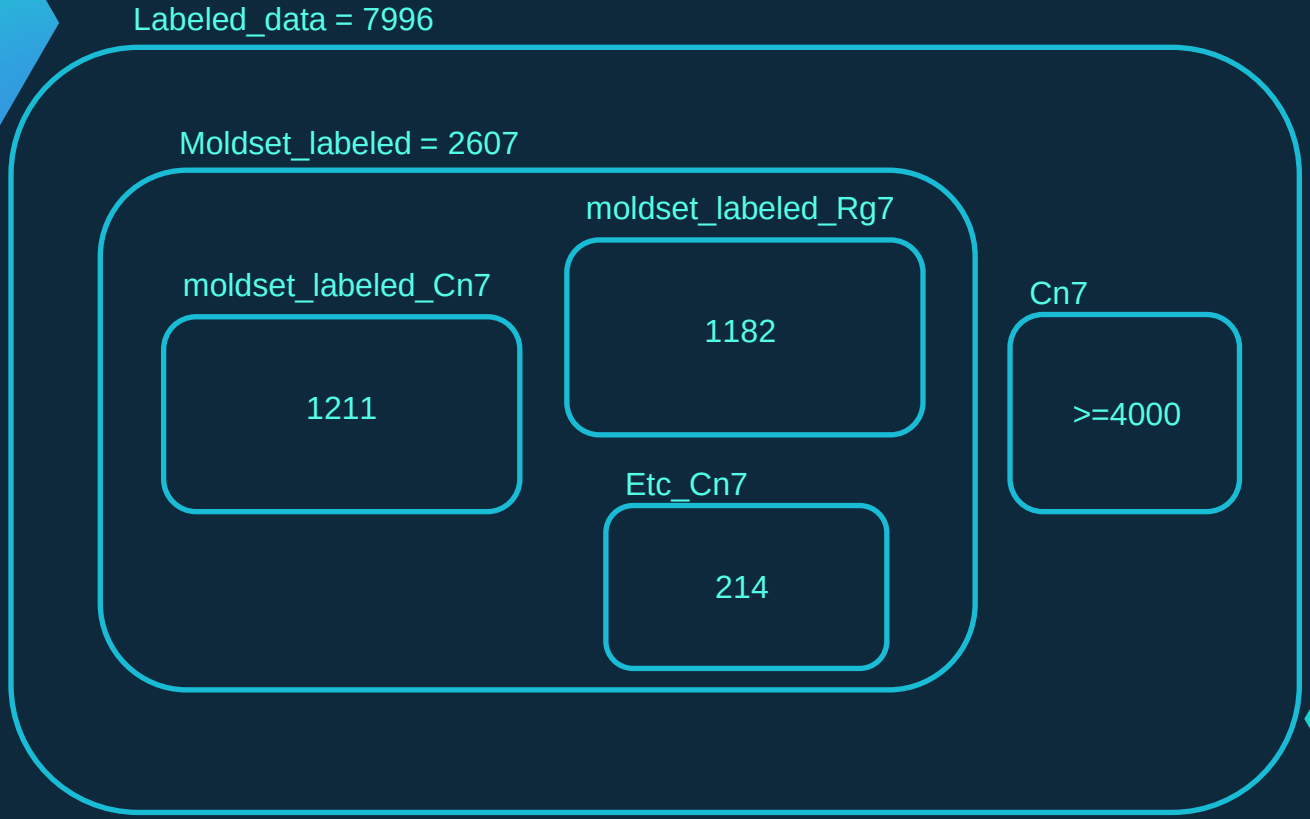
Our process is easy

데이터
구조 파악

데이터
시각화

모델
수립 및
평가

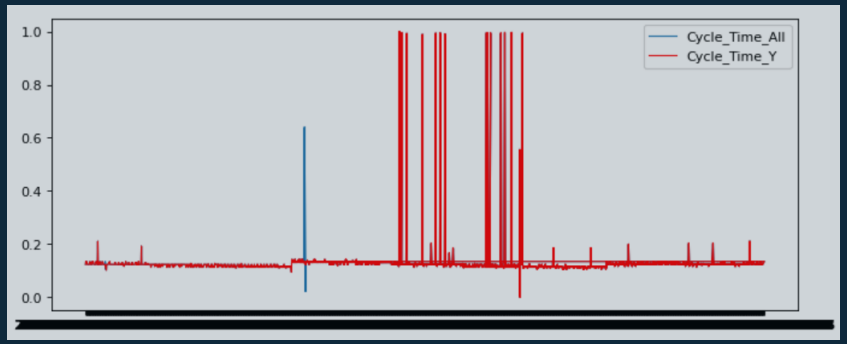
1 데이터 구조 파악


Labeled_data = 7996
Moldset_labeled = 2607
moldset_labeled_Rg7
1182
moldset_labeled_Cn7
1211
Etc_Cn7
214
Cn7
>=4000

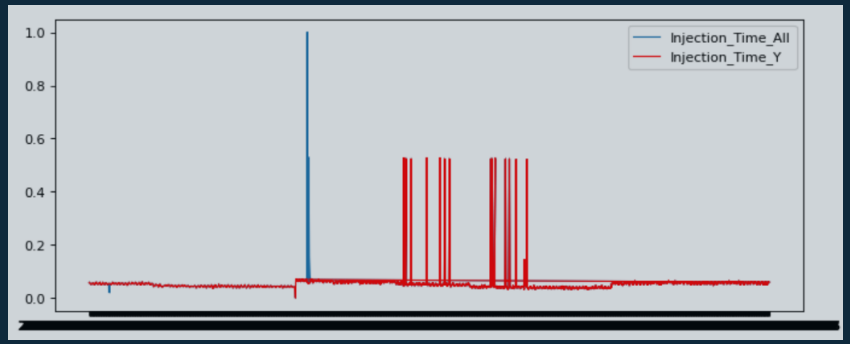
- 650톤 우진 2호기가 가장 많음
- 분석대상 : CN7, RG3

2 데이터 시각화

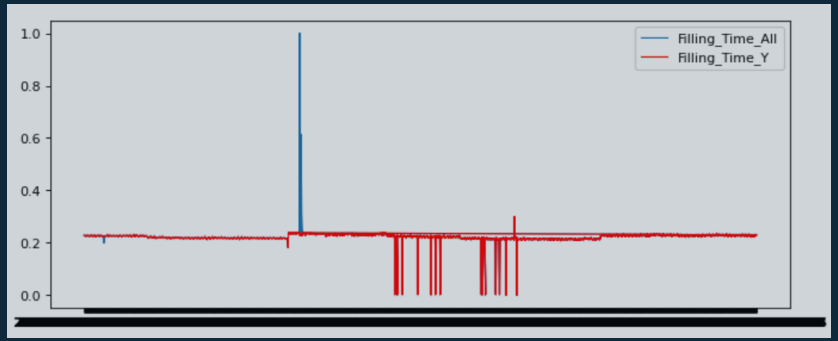
Cycle_Time



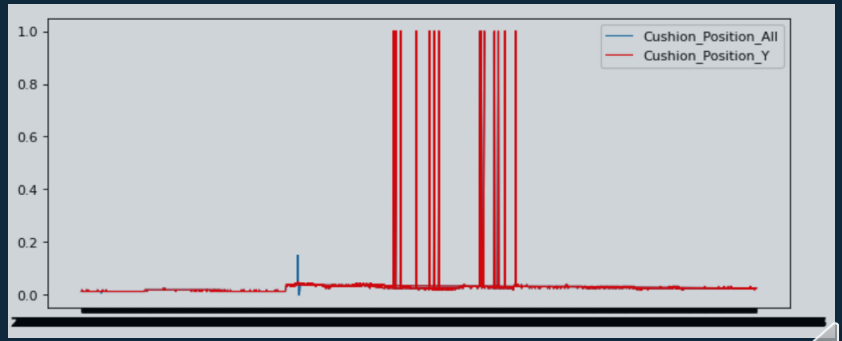
Injection_Time



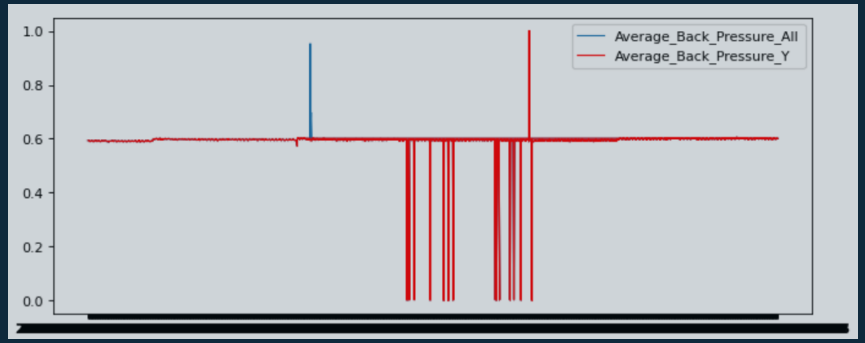
Filling_Time



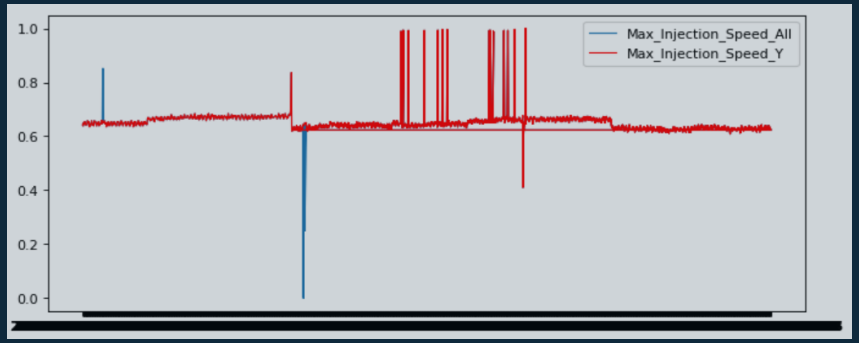
Cushion_Position



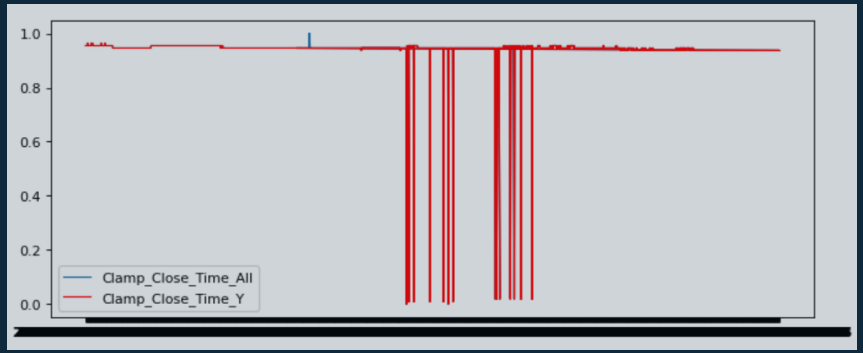
Average_Back_Pressure



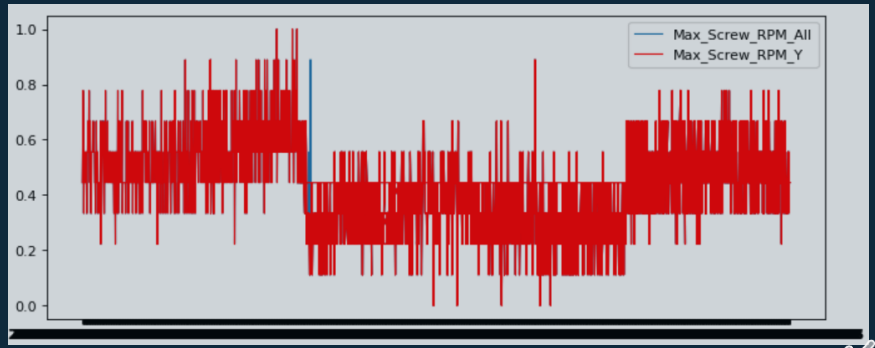
Max_Injection_Speed



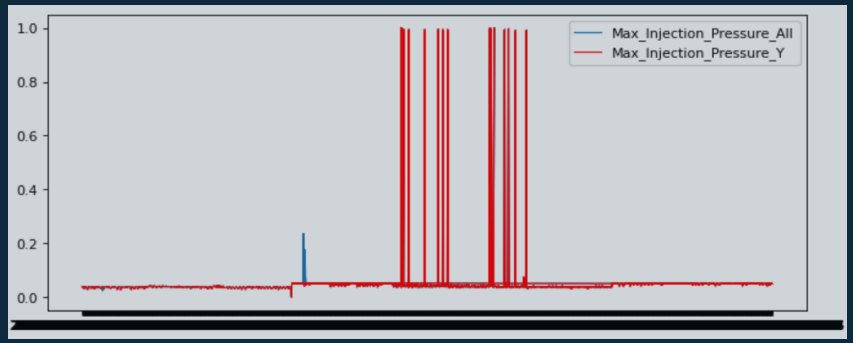
Clamp_Close_Time



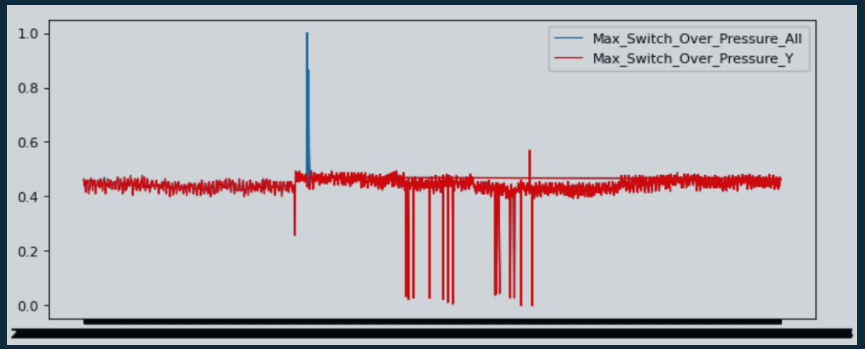
Max_Screw_RPM



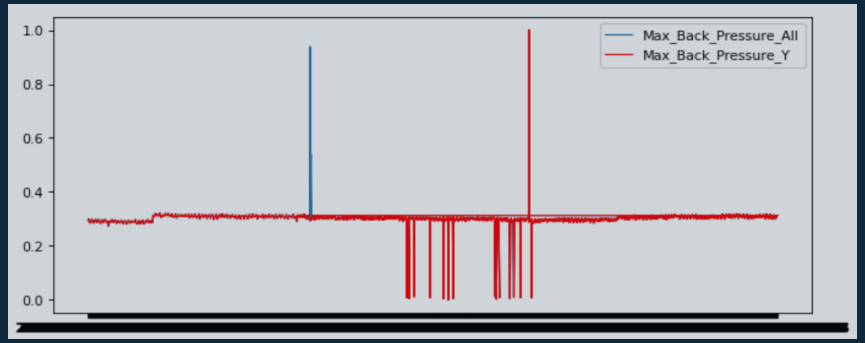
Max_Injection_Pressure



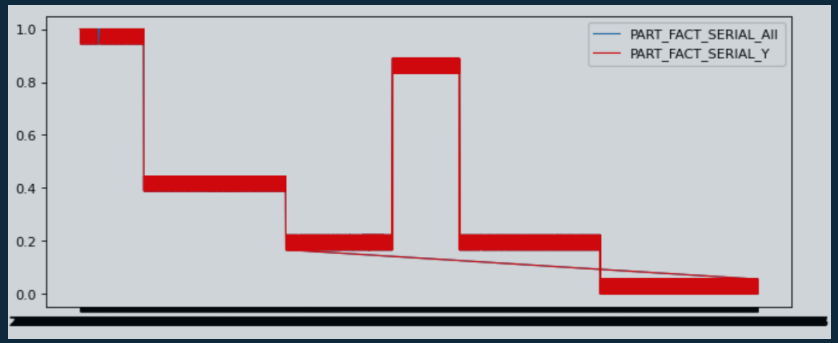
Max_Switch_Over_Pressure



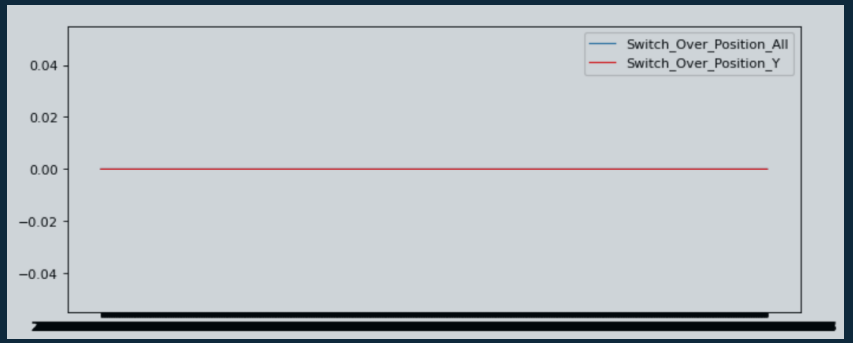
Max_Back_Pressure



PART_FACT_SERIAL

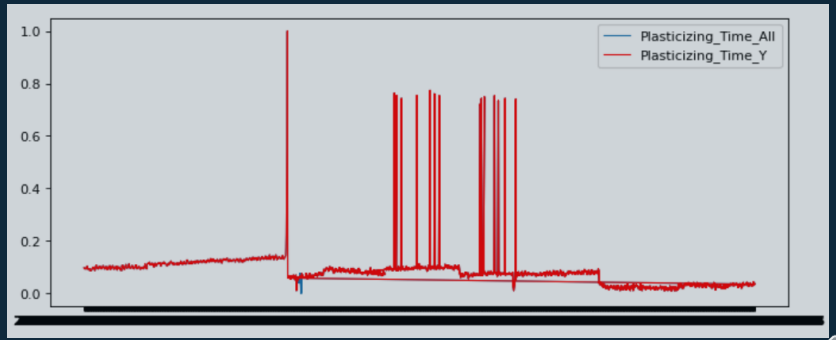


Switch_Over_Position

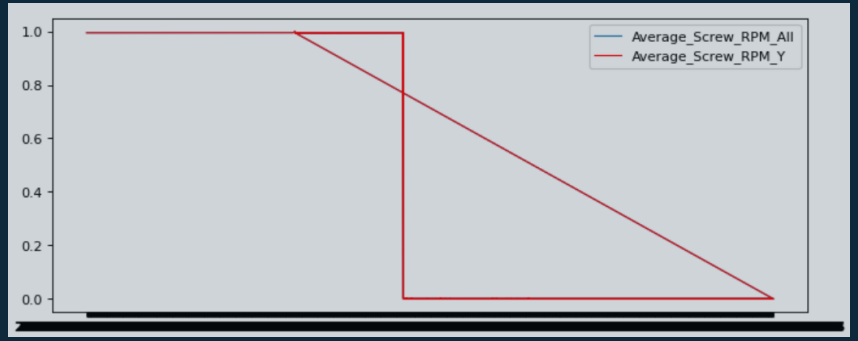


Plasticizing_Position

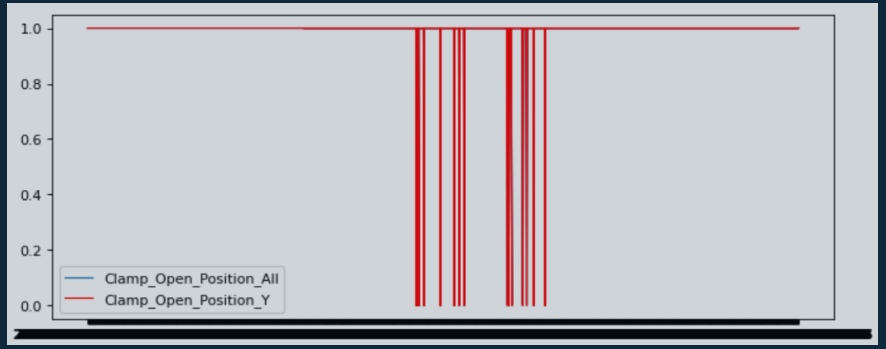
Plasticizing_Time



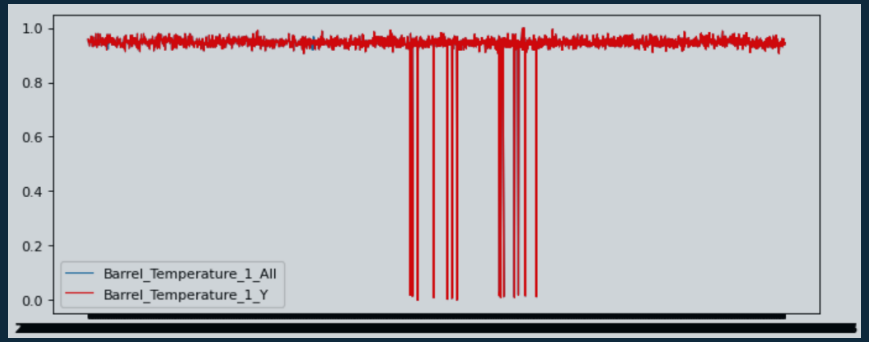
Average_Screw_RPM



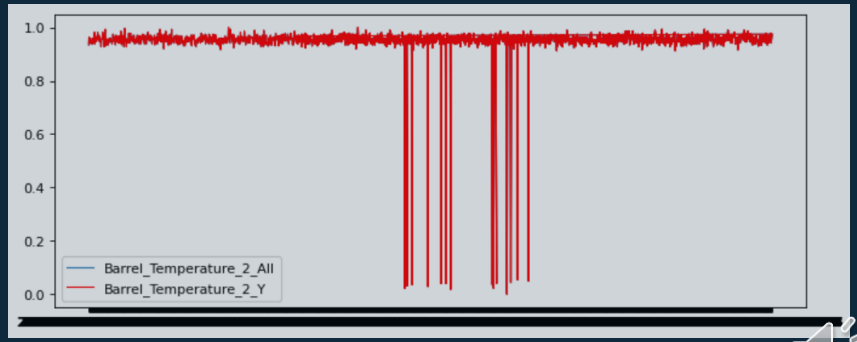
Clamp_Open_Position



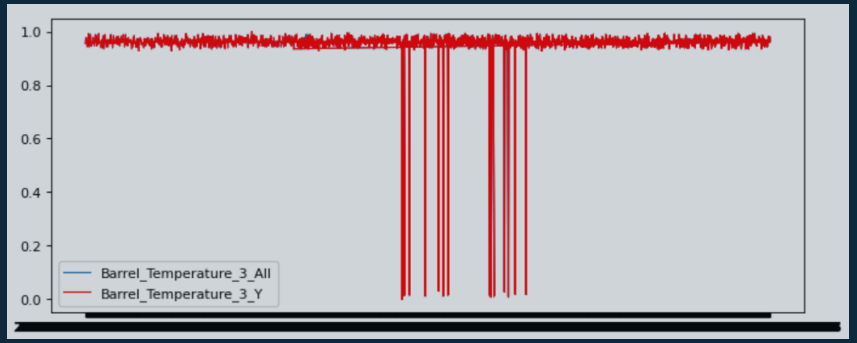
Barrel_Temperature_1



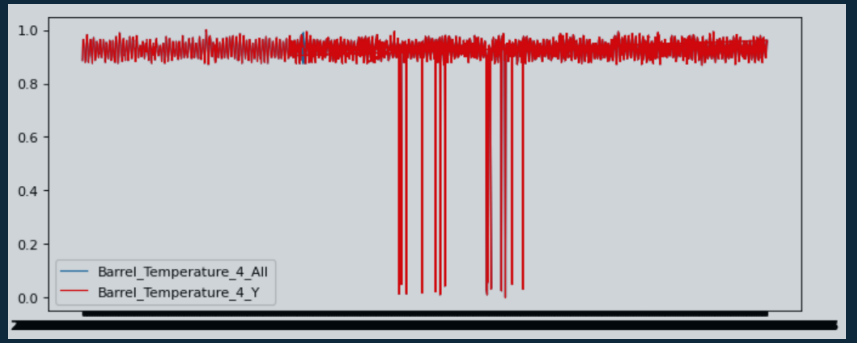
Barrel_Temperature_2



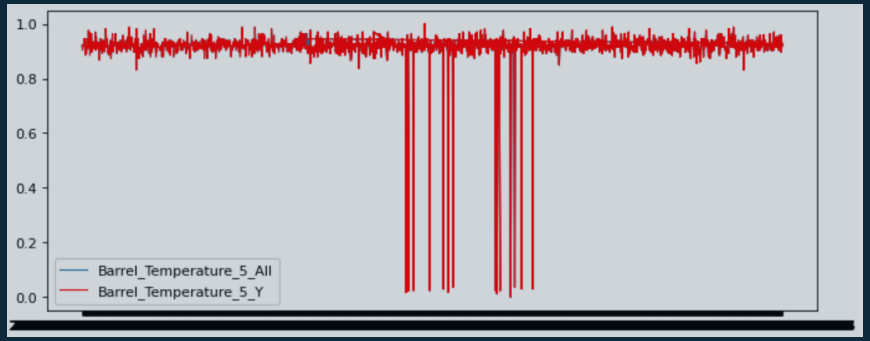
Barrel_Temperature_3



Barrel_Temperature_4

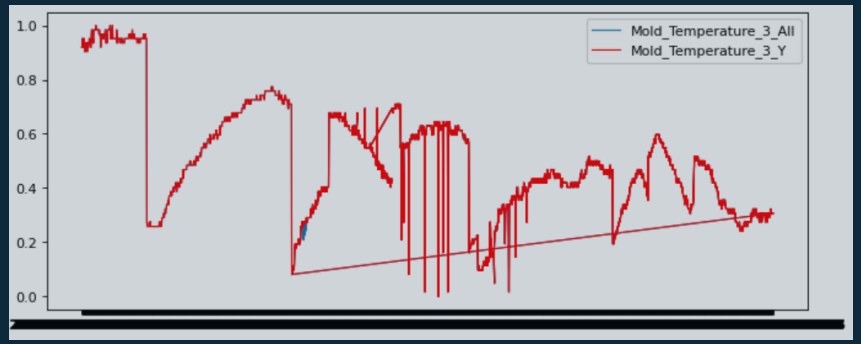


Barrel_Temperature_5

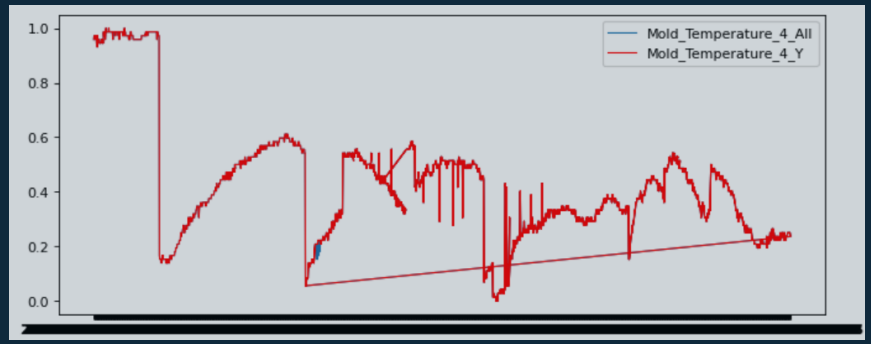


Barrel_Temperature_6

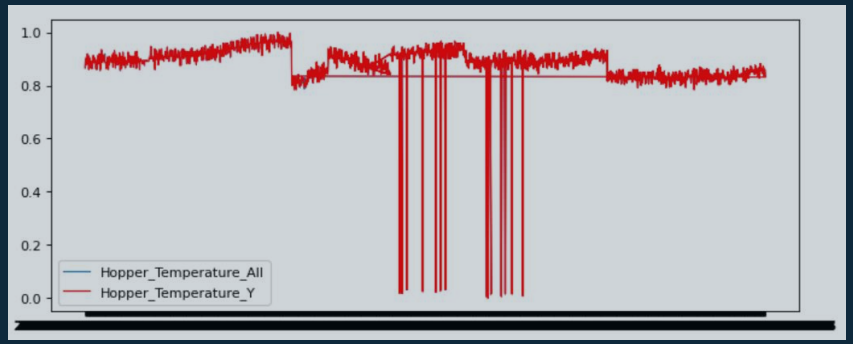
Mold_Temperature_3



Mold_Temperature_4



Hopper_Temperature



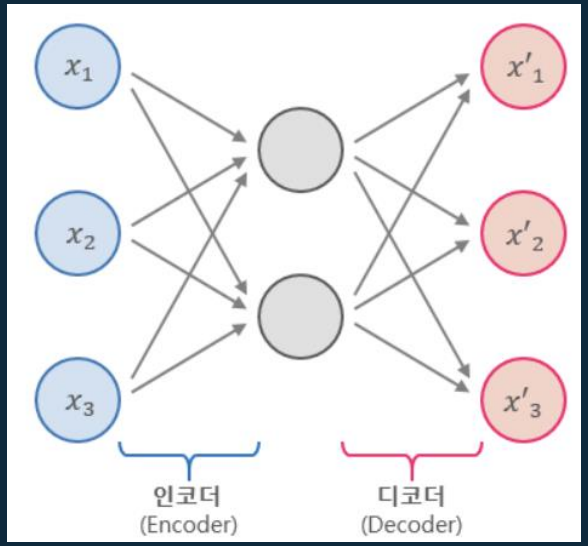
3

모델 수립 및 평가

1. AutoEncoder

■ 입력과 출력을 같도록 하는 구조를 의미하는 것으로 노이즈 제거에 탁월한 비지도 머신러닝 방법론

■ input에 최대한 똑같게 output이 나오도록 하는 것으로 노이즈가 없던 부분을 복원하도록 하는 문제를 설계 가능한 장점

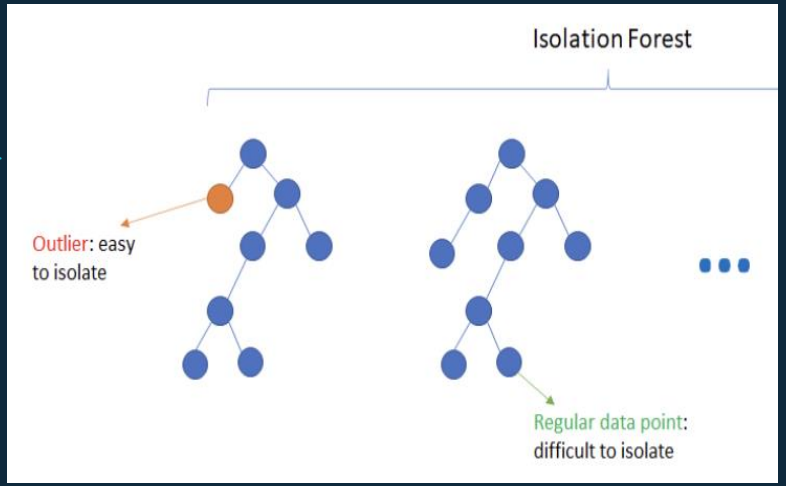



```
def autoencoder_model(X):
    inputs = Input(shape=(X.shape[1], X.shape[2]))
    L1 = LSTM(16, activation='relu', return_sequences=True,
              kernel_regularizer=regularizers.l2(0.00))(inputs)
    L2 = LSTM(4, activation='relu', return_sequences=False)(L1)
    L3 = RepeatVector(X.shape[1])(L2)
    L4 = LSTM(4, activation='relu', return_sequences=True)(L3)
    L5 = LSTM(16, activation='relu', return_sequences=True)(L4)
    output = TimeDistributed(Dense(X.shape[2]))(L5)
    model = Model(inputs=inputs, outputs=output)
    return model
```

2. Isolation Forest

■ 데이터 셋을 의사결정나무 형태로 표현해 정상값을 분리하기 위해서는 깊이가 깊고 이상치 값은 의사결정나무 상단부에서 분리된다는 머신러닝 방법론

■ 군집기반 이상탐지 알고리즘에 비해 계산량이 매우 적으며 강건한 모델 생성 가능

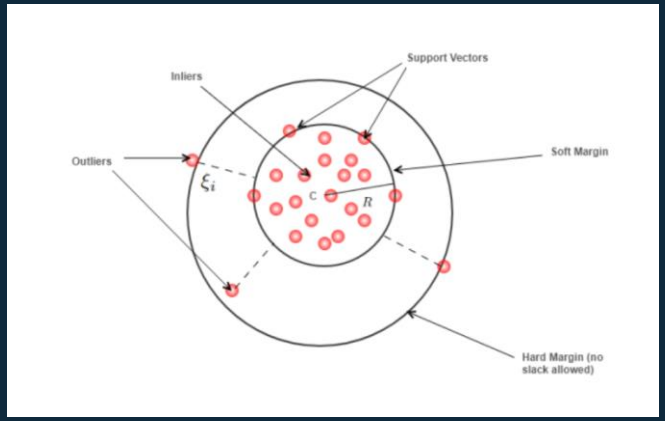


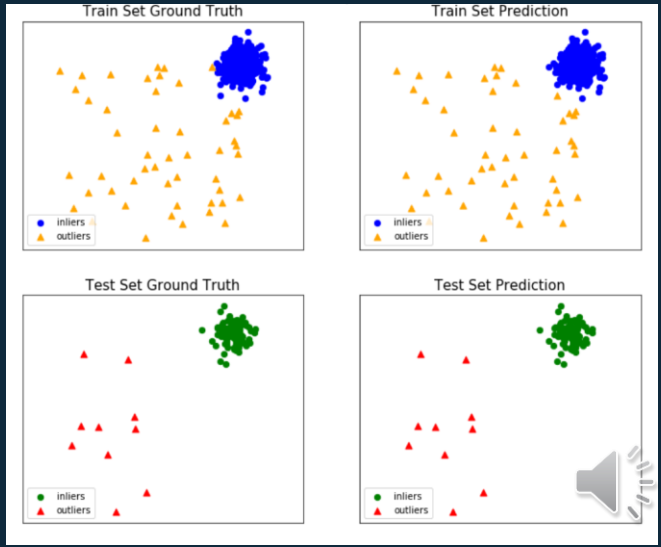

```
if_clf = IsolationForest(contamination=0.3, max_features=1.0,
                         max_samples=0.7, n_estimators=80, random_state=4)
if_clf.fit(X_train)
```

3. One-class SVM

■ One-class SVM은 novelty detection에 서포트 벡터를 사용하는 방법

■ 기존의 서포트 벡터 방식처럼 초평면을 사용하는 것이 아닌 구를 사용하며, 정상 데이터 안에서 데이터 포인트의 밀집도를 찾아 긍정 혹은 부정으로 인식하기 때문에 확률분포가 필요하지 않음





Confusion-Matrix(사출공정)

	Predicted: Y	Predicted: N	
Actual: Y	100	2	Sensitivity(0.972)
Actual: N	0	69	Specificity(1)
	Precision(1)	NPV(0.972)	Accuracy(0.988)

※ F1-Score = 0.986

※ Train Time = 0.45 sec

Confusion-Matrix(용해탱크)

	Predicted: Y	Predicted: N	
Actual: Y	195,462	14,538	Sensitivity(0.686)
Actual: N	145,289	31,778	Specificity(0.931)
	Precision(0.690)	NPV(0.686)	Accuracy(0.590)

※ F1-Score = 0.710

※ Train Time = 30 min

감사합니다!