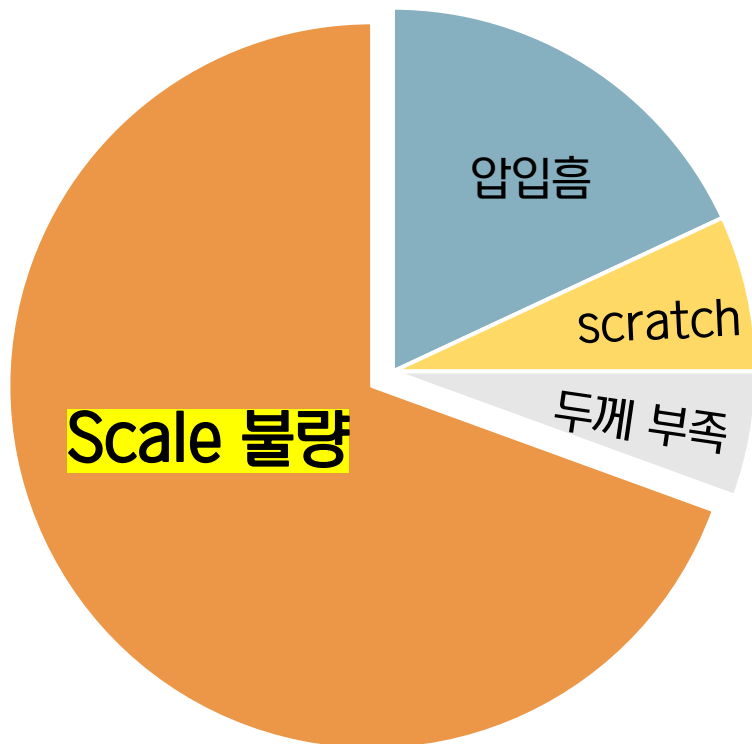


후판 압연공정 Scale불량 발생 원인 분석 및 개선안 수립

C4 고다영

A공장의 고객사에서 최근 들어 '**Scale 불량 발생 증가**' 라는 이슈가 발생했다. 그 원인을 분석해 본 결과 **압연공정**에서 Scale 불량이 급증한 것을 확인할 수 있었다. 이에 따라 압연 공정 관련 데이터를 수집하여 Scale 불량 발생의 근본 원인을 찾고 분석하여 개선 기회를 도출하고자 한다.

[A공장 불량 발생 현황]



발생원인	발생률(%)
압입흠	1.3
Scratch	0.5
두께 부족	0.4
Scale	5.0
계	7.2%

- Plate 두께가 얇을수록 Scale 불량 발생률은 증가할 것이다.
- 압연 공정 중 온도가 높을 수록 Scale 불량 발생률은 증가할 것이다.
- HSB를 적용하냐 안 하냐에 따라 Scale 불량 발생 여부에 큰 차이를 보일 것이다.
- 공정 중 descaling을 많이 할수록 Scale 불량 발생률이 감소할 것이다.

Scale 발생	가열로 가열대 온도	가열로 균열대 온도	가열로 추출 온도	HSB	압연 온도	Descaling 횟수	판 두께
없음 ↕ 발생	저 ↕ 고	저 ↕ 고	저 ↕ 고	적용 ↕ 미적용	저 ↕ 고	증가 ↕ 감소	후 ↕ 박

[분석 목표] Scale 불량 발생 원인 분석 및 개선안 수립

데이터 수집

- 후판 관련 데이터 (Plate 두께, 폭, 길이, 중량, 강종, 제품규격)
- 압연 공정 관련 데이터 (가열로 온도, 시간, 압연 온도, 작업조 등)

데이터 정제

- 데이터 품질 확인 – 결측치, 이상치 처리
- EDA – 그래프를 통한 데이터 분포 확인 및 파생변수 생성

데이터 분석

- 회귀분석 및 기계학습 기법을 통한 데이터 별 중요도 파악
- Scale 불량 분류 모델 생성 및 모델 평가

결론 도출

- Scale 불량에 대한 근본적인 원인 파악 및 해석
- 원인을 분석하여 개선방안 수립

데이터 탐색

[데이터 수집 결과]

PLATE_NO	Plate No	범주형	설명변수
ROLLING_DATE	작업시각	연속형	설명변수
SCALE	Scale불량	범주형	목표변수
SPEC	제품 규격	범주형	설명변수
STEEL_KIND	강종	범주형	설명변수
PT_THICK	Plate 두께	연속형	설명변수
PT_WIDTH	Plate 폭	연속형	설명변수
PT_LENGTH	Plate 길이	연속형	설명변수
PT_WEIGHT	Plate 중량	연속형	설명변수
FUR_NO	가열로 호기(1, 2, 3호기)	범주형	설명변수
FUR_NO_ROW	가열로 장입열(1열, 2열)	범주형	설명변수
FUR_HZ_TEMP	가열로 가열대 온도	연속형	설명변수
FUR_HZ_TIME	가열로 가열대 시간	연속형	설명변수
FUR_SZ_TEMP	가열로 균열대 온도	연속형	설명변수
FUR_SZ_TIME	가열로 균열대 시간	연속형	설명변수
FUR_TIME	가열로 재로시간	연속형	설명변수
FUR_EXTTEMP	가열로 추출온도	연속형	설명변수
ROLLING_TEMP_T5	압연온도	연속형	설명변수
HSB	HSB적용(1-적용,0-미적용)	범주형	설명변수
ROLLING_DESCALING	압연 중 Descaling 횟수	연속형	설명변수
WORK_GR	작업조	범주형	설명변수

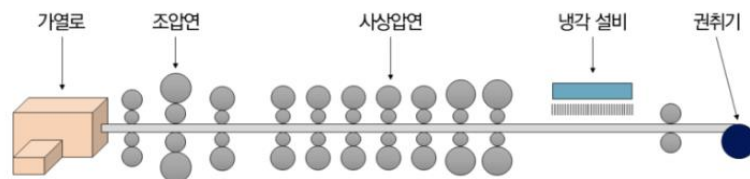
Scale 불량 원인 분석에 필요하지 않으므로 제거 결정.

전체 720개 데이터 중
양품 489개(67.9%), 불량 231개(32%)

[후판 관련 데이터]

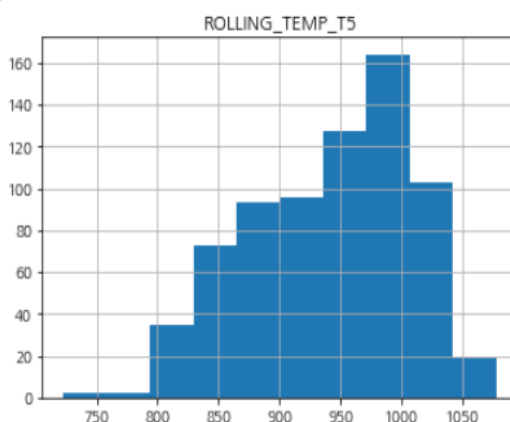
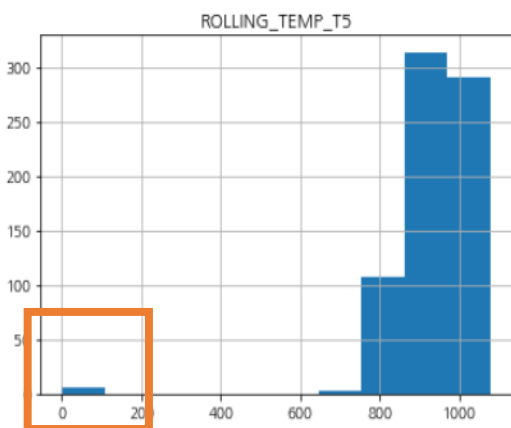
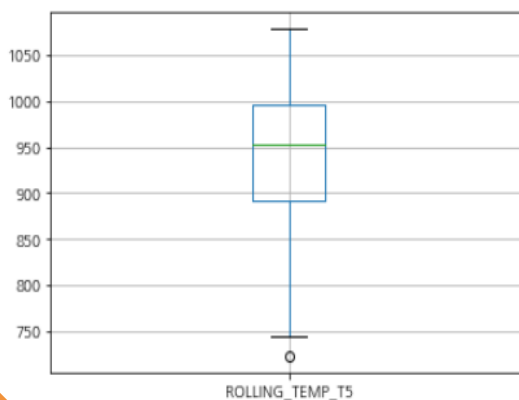
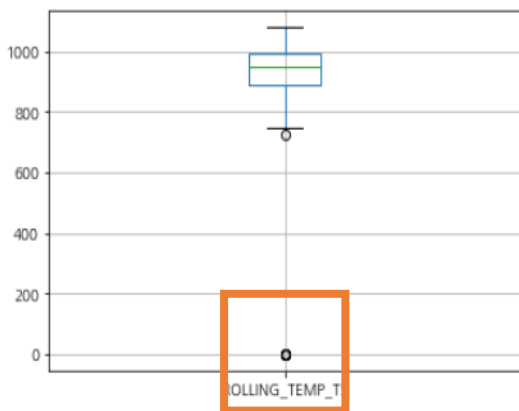
- Plate 두께, 폭, 길이, 중량은 고객사에서 주문한 plate 정보
- 강종은 C로 시작하는 경우 탄소강, T로 시작하는 경우 티타늄강을 의미.

[압연 공정 관련 데이터]



- 각 공정 과정에서 수집된 데이터
- 재로시간 > 가열대 시간 + 균열대 시간

[압연온도 Boxplot, Histogram]



① 결측치, 이상치 확인

- 결측치는 존재하지 않는다.
- 압연온도를 boxplot, histogram을 통해 분포를 확인한 결과, 압연온도는 0도가 될 수 없다.

→ 이상치 발견 및 제거

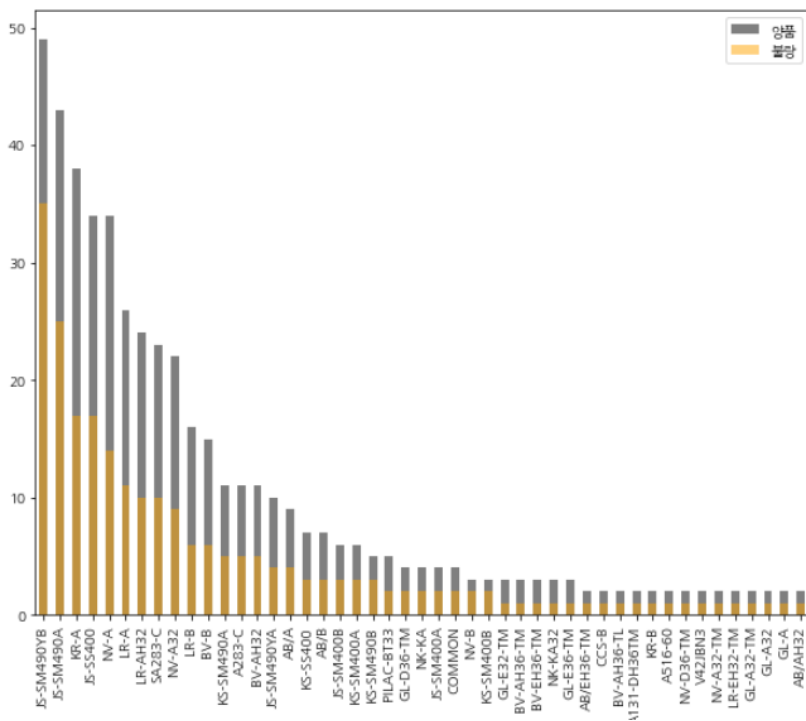
- 이 외 변수들 역시 boxplot, histogram을 통해 확인한 결과 공정 과정 중 발생할 수 있는 데이터로 이상치 존재하지 않는다.

데이터 탐색

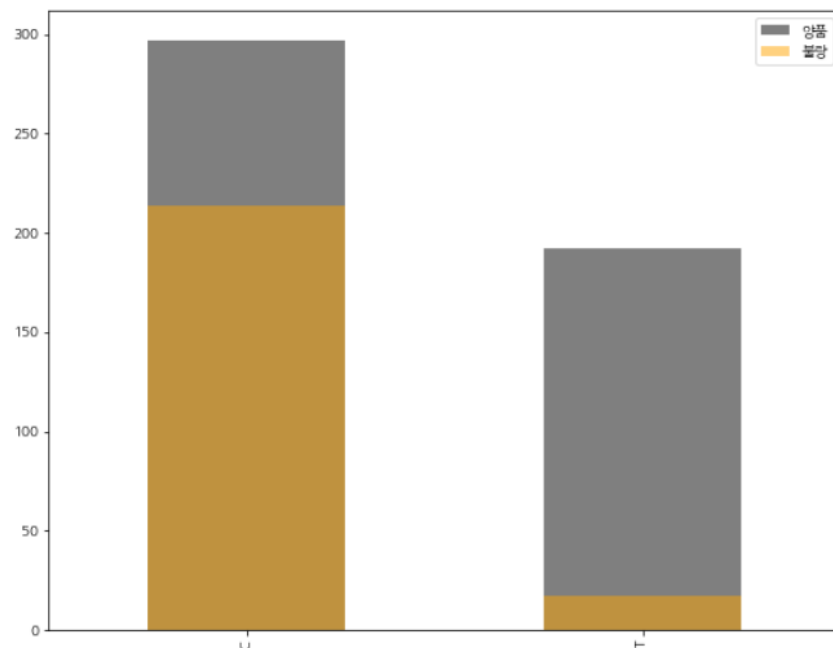
② 그래프 탐색

- SPEC(제품 규격)별 불량, 양품 추이가 목표변수의 추이와 거의 비슷하게 나타남.
→ SPEC은 유의미한 설명변수가 되지 못한다고 판단해 제거
- 강종별 불량, 양품 개수 확인 결과 탄소강에서 더 많은 불량률 확인
→ 강종은 Scale 불량에 유의미한 설명변수이다.

[제품 규격 별 양품, 불량 개수]



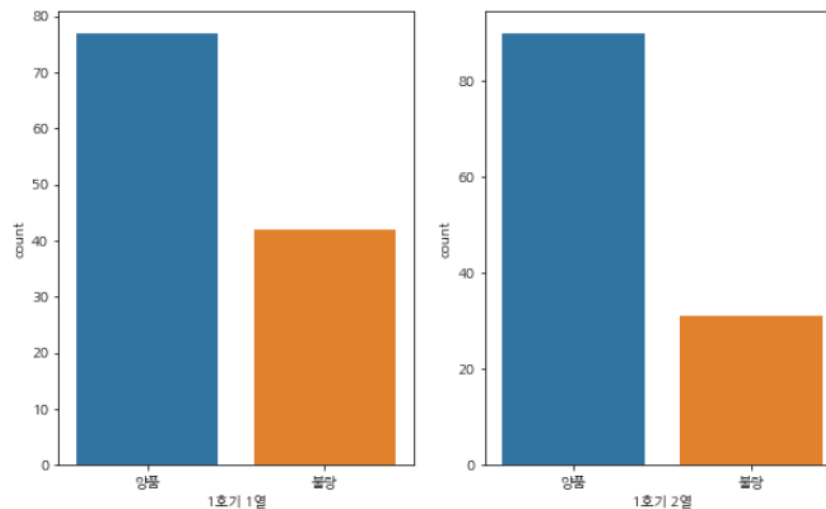
[강종별(탄소강, 티타늄강) 양품, 불량 개수]



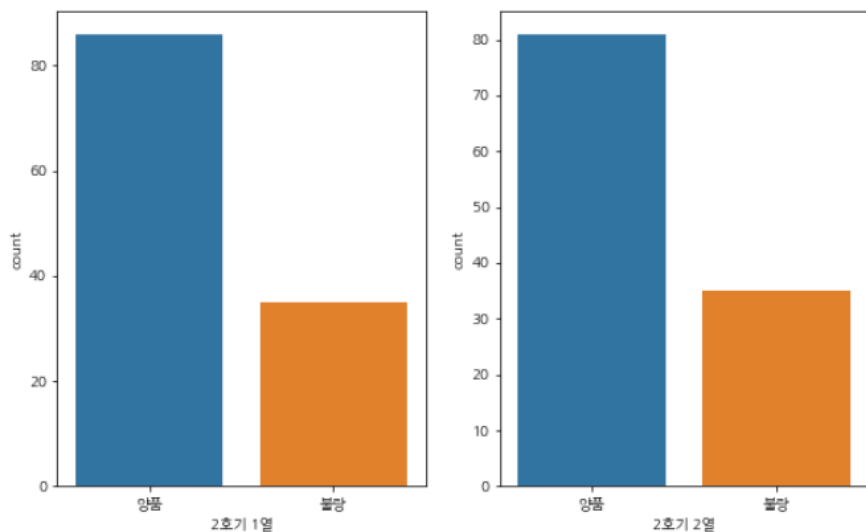
② 그래프 탐색

- 가열로 호기와 열 별 불량/양품 개수를 비교해 본 결과,
 - 1호기 1열의 불량률 \approx 1호기 2열의 불량률
 - 2호기/3호기 1열 불량률 = 2호기/3호기 2열 불량률
- 범주형 변수인 '가열로 호기'와 '가열로 열' Scale 불량에 유의미한 설명변수가 되지 못하므로 제거.

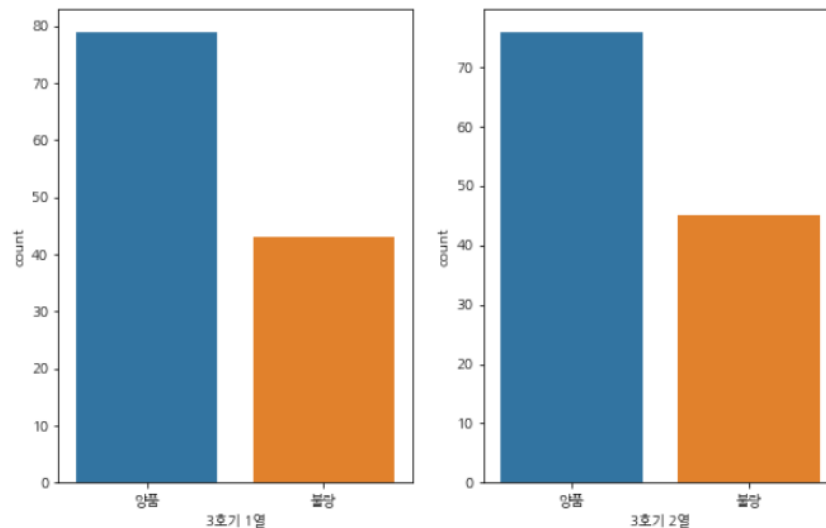
[1호기 1열과 1호기 2열 비교]



[2호기 1열과 2호기 2열 비교]

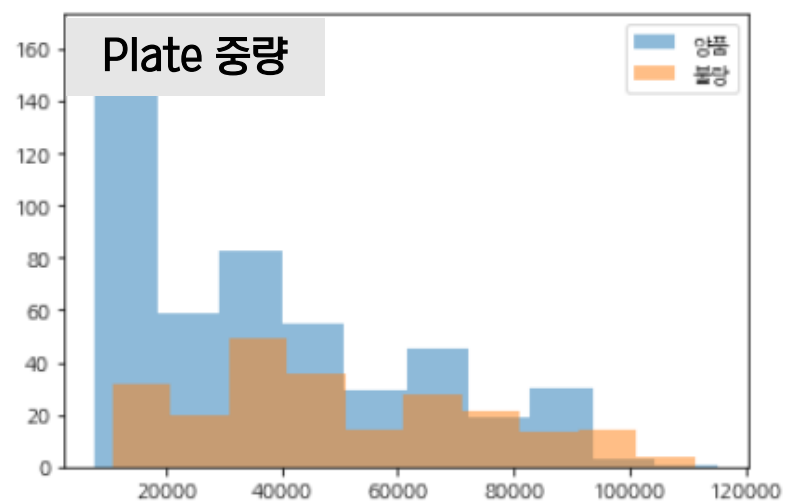
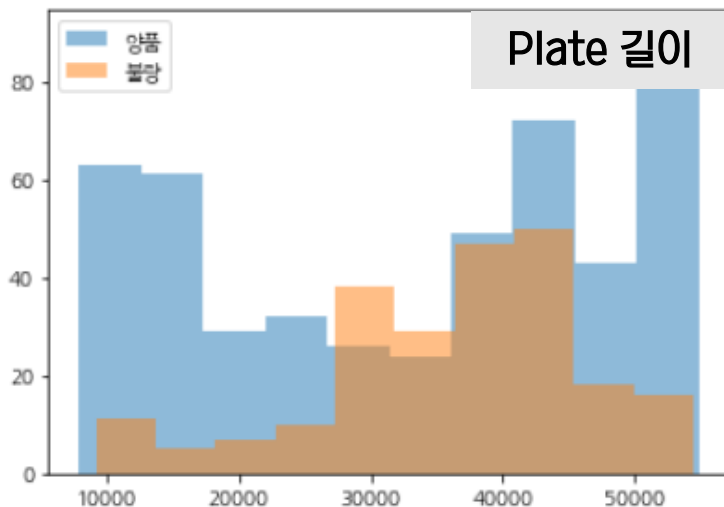
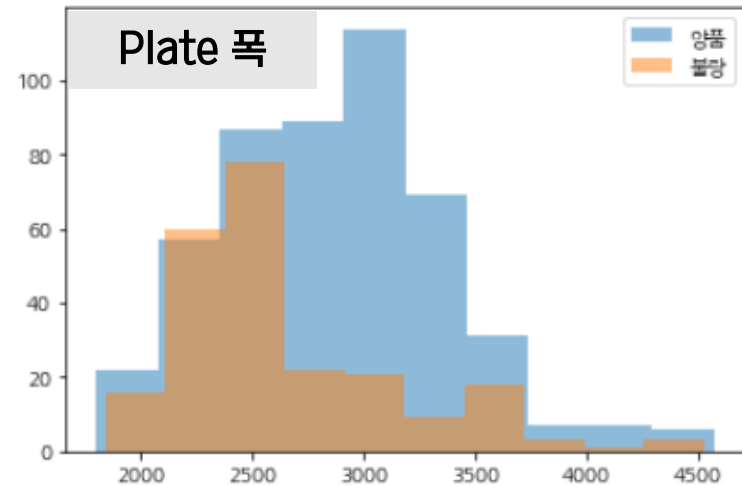
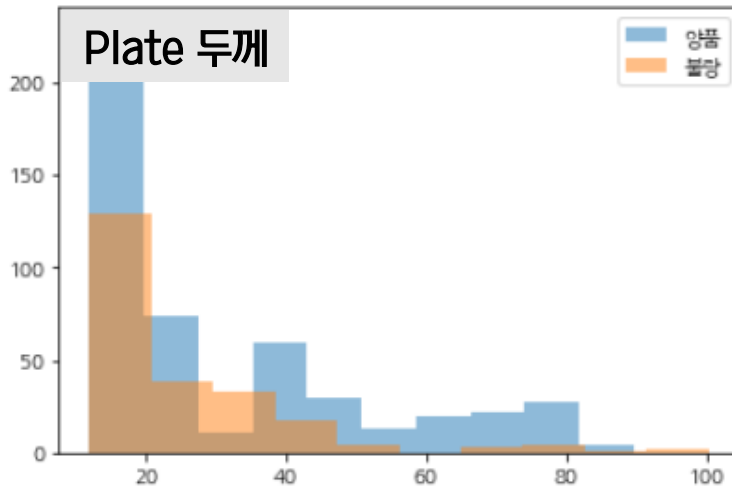


[3호기 1열과 3호기 2열 비교]



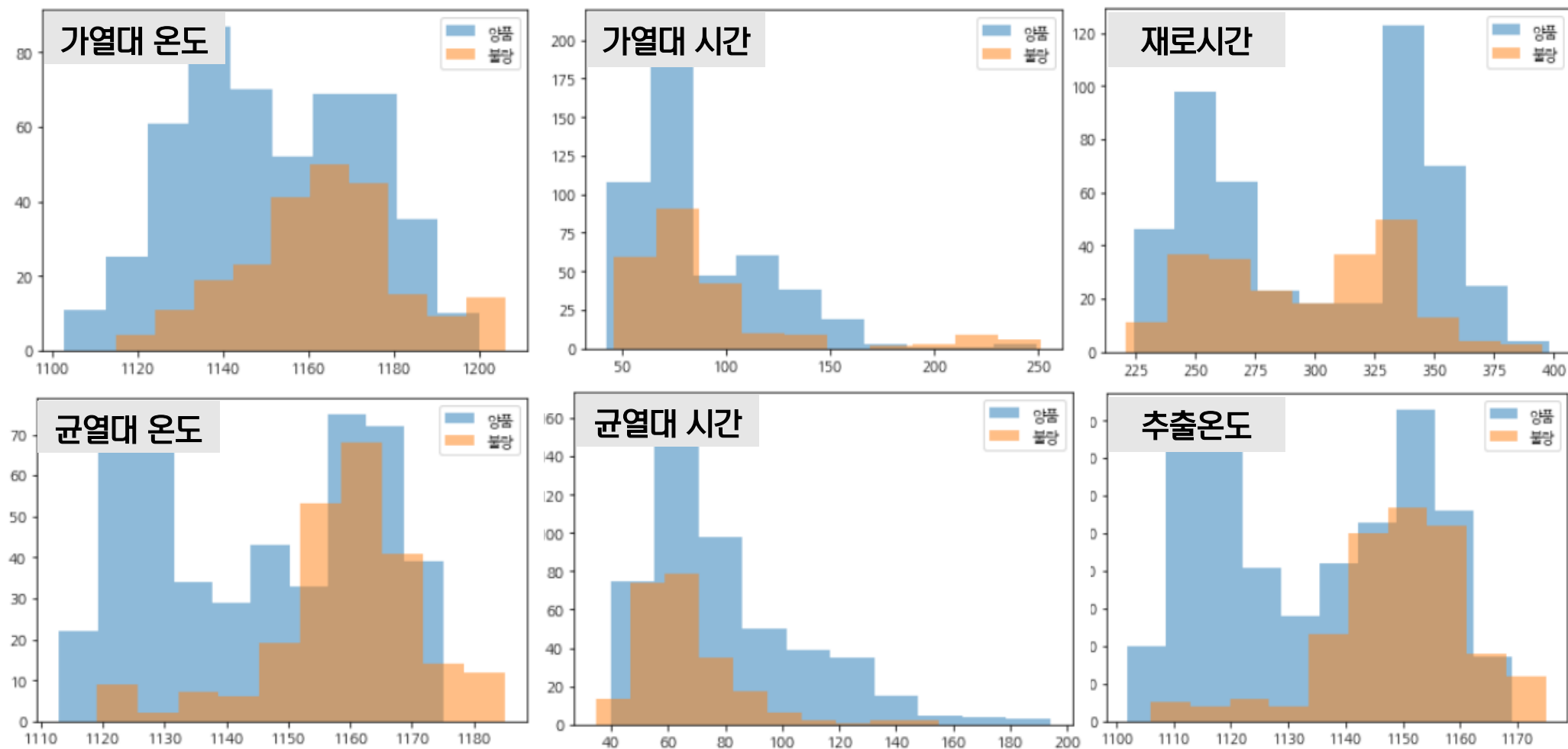
② 그래프 탐색

- Plate의 두께, 폭, 길이, 중량은 Scale 불량에 유의미한 설명변수이다.



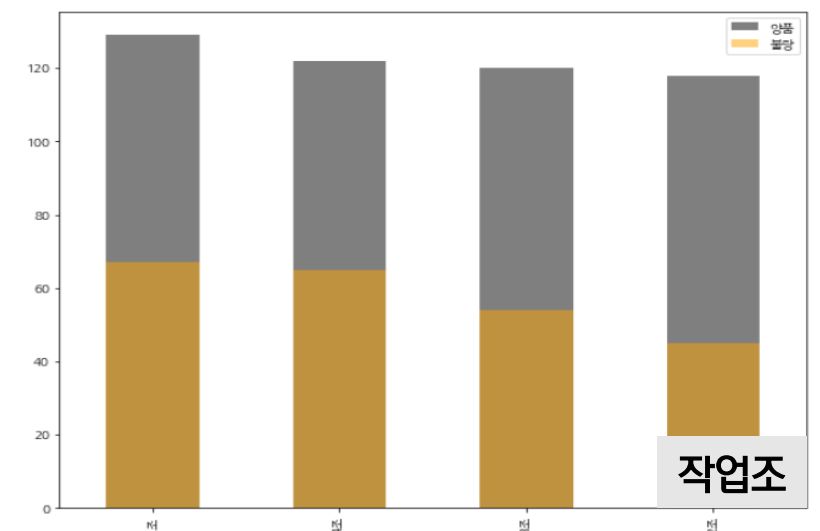
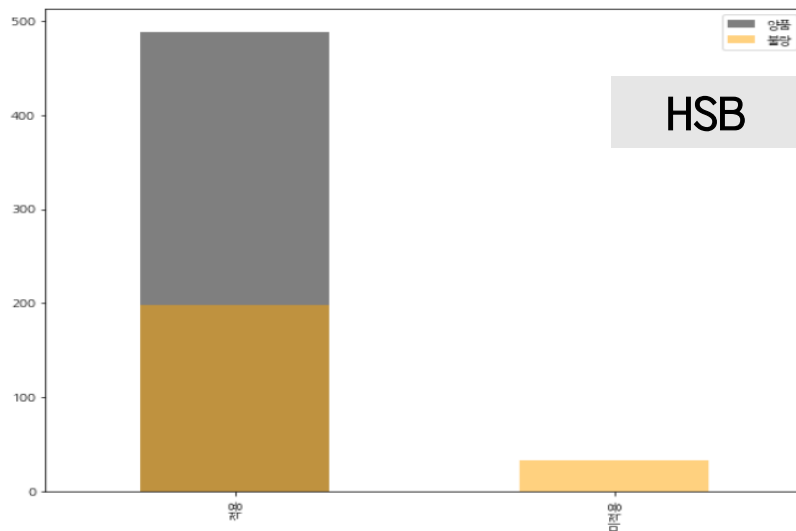
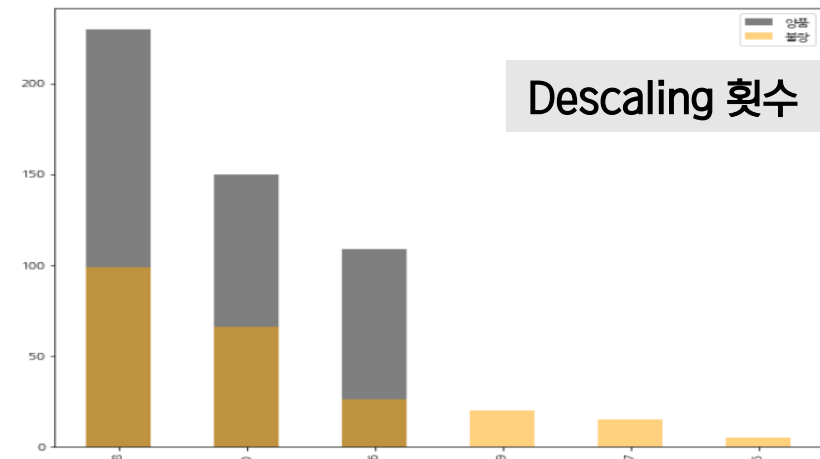
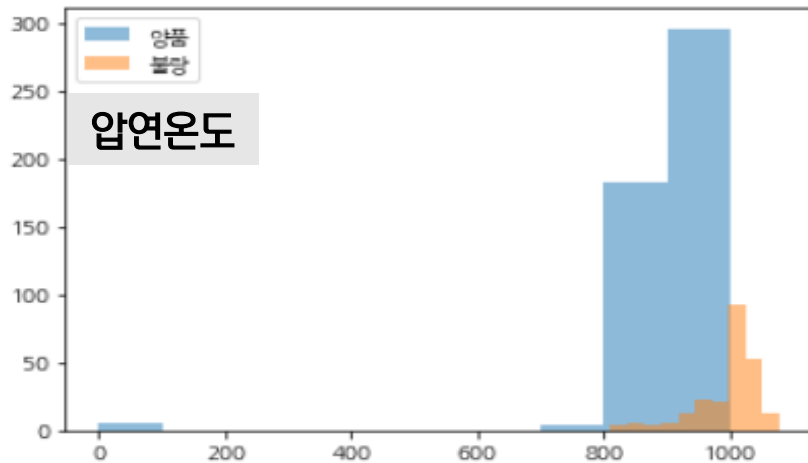
② 그래프 탐색

- 가열대, 균열대의 온도, 시간 / 재로시간, 추출온도 변수는 유의미한 설명변수이다.



② 그래프 탐색

- 압연온도, Descaling 횟수에 따라 불량/양품 비율은 상당한 차이를 보인다. 특히 HSB를 적용하지 않았을 때 모두 불량인 점은 주목할 만 하다. 그러나 작업조는 유의미한 설명변수가 아니므로 제거.



③ 설명변수 간 상관관계 확인

	PT_THICK	PT_WIDTH	PT_LENGTH	PT_WEIGHT	FUR_HZ_TEMP	FUR_HZ_TIME	FUR_SZ_TEMP	FUR_SZ_TIME	FUR_TIME	FUR_EXTEMP	ROLLING_TEMP_T5	ROLLING_DESCALING
PT_THICK	1	-0.31	-0.86	-0.39	-0.52	0.17	-0.69	0.28	0.41	-0.69	-0.5	-0.84
PT_WIDTH	-0.31	1	0.12	0.03	0.18	-0.13	0.23	-0.056	-0.39	0.22	-0.11	0.34
PT_LENGTH	-0.86	0.12	1	0.45	0.47	-0.088	0.64	-0.23	-0.24	0.64	0.43	0.81
PT_WEIGHT	-0.39	0.03	0.45	1	0.15	0.009	0.36	-0.19	-0.054	0.36	0.42	0.23
FUR_HZ_TEMP	-0.52	0.18	0.47	0.15	1	-0.11	0.77	-0.33	-0.34	0.76	0.36	0.47
FUR_HZ_TIME	0.17	-0.13	-0.088	0.009	-0.11	1	-0.21	0.067	0.48	-0.19	0.004	-0.15
FUR_SZ_TEMP	-0.69	0.23	0.64	0.36	0.77	-0.21	1	-0.57	-0.47	0.98	0.66	0.64
FUR_SZ_TIME	0.28	-0.056	-0.23	-0.19	-0.33	0.067	-0.57	1	0.5	-0.57	-0.48	-0.29
FUR_TIME	0.41	-0.39	-0.24	-0.054	-0.34	0.48	-0.47	0.5	1	-0.46	-0.21	-0.36
FUR_EXTEMP	-0.69	0.22	0.64	0.36	0.76	-0.19	0.98	-0.57	-0.46	1	0.66	0.64
ROLLING_TEMP_T5	-0.5	-0.11	0.43	0.42	0.36	0.004	0.66	-0.48	-0.21	0.66	1	0.37
ROLLING_DESCALING	-0.84	0.34	0.81	0.23	0.47	-0.15	0.64	-0.29	-0.36	0.64	0.37	1

■ 연속형 설명변수 간의 상관관계가 굉장히 높게 나타난다.

→ 핵심 영향인자만을 선정하여 원인을 분석하고 분류모델을 생성하는 것은 도리어 모델의 신뢰성을 해칠 우려가 있으므로 **주 성분 분석을 통한 차원 축소**를 진행한다.

	PrinNo	EigenValues	EigenValueRatio	CumEigenVauleRatio
0	1	5.700	0.474	0.474
1	2	1.572	0.131	0.605
2	3	1.283	0.107	0.712
3	4	0.910	0.076	0.788
4	5	0.807	0.067	0.855
5	6	0.631	0.053	0.907
6	7	0.420	0.035	0.942
7	8	0.305	0.025	0.968

- 총 12개의 연속형 설명변수를 8개의 주성분으로 차원 축소. 그 중에서도 12개 변수의 80% 이상을 설명할 수 있는 주성분1 ~ 주성분5까지를 선택.

<주성분 설명>

- Prin1
→ Plate 두께와 가열대 온도 간의 연관성
- Prin2
→ Plate 폭과 가열대 재로 시간 간의 연관성
- Prin3
→ Plate 면적과 시간, descaling 횟수 간의 연관성
- Prin4
→ Plate 중량과 가열대 시간 및 온도 간의 연관성
- Prin5
→ Plate 밀도와 가열대 공정 시간 간의 연관성

	Feature	Prin1	Prin2	Prin3	Prin4	Prin5
0	PT_THICK	-0.361	0.030	0.333	-0.084	-0.039
1	PT_WIDTH	0.126	0.465	-0.338	-0.280	-0.569
2	PT_LENGTH	0.333	-0.176	-0.366	0.133	0.143
3	PT_WEIGHT	0.186	-0.355	-0.087	0.506	-0.484
4	FUR_HZ_TEMP	0.302	0.029	0.087	-0.410	0.338
5	FUR_HZ_TIME	-0.101	-0.502	-0.101	-0.624	-0.375
6	FUR_SZ_TEMP	0.393	-0.025	0.159	-0.135	0.088
7	FUR_SZ_TIME	-0.239	-0.023	-0.538	0.086	0.329
8	FUR_TIME	-0.235	-0.503	-0.221	-0.146	0.176
9	FUR_EXTEMP	0.390	-0.035	0.161	-0.136	0.083
10	ROLLING_TEMP_T5	0.277	-0.341	0.299	0.128	-0.055
11	ROLLING_DESCALING	0.335	0.036	-0.370	-0.048	0.084

최종적으로 연속형 변수를 설명하는 주성분1 ~ 주성분5와 범주형 변수 2개(강종, HSB 적용 여부)를 더하여 총 7개의 설명변수로 구성된 데이터셋을 완성하였다. 위 데이터셋을 기반으로 로지스틱 회귀분석, 의사결정나무(DT), 랜덤포레스트(RF), 그래디언트 부스팅(GB)를 통해 변수 간 중요도 순위를 파악한 결과, 가장 많은 영향을 미치는 인자는 순서대로 Prin3, Prin2, STEEL_KIND, HSB이다.

	모델링 기법				순위
	회귀분석	DT	RF	GB	
Prin1		6	3	4	5
Prin2	4	3	2	2	2
Prin3	2	1	1	1	1
Prin4		5	6	7	7
Prin5			7	6	6
HSB	1	4	5	5	4
STEEL_KIND	3	2	4	3	3

[그래디언트 부스팅 모델 성능평가]

Train data Accuracy : 0.961924

Test data Accuracy : 0.902326

Confusion Matrix :

```
[[148   8]
 [ 16  43]]
```

	precision	recall	f1-score	support
0	0.902	0.949	0.925	156
1	0.843	0.729	0.782	59
accuracy			0.888	215
macro avg	0.873	0.839	0.853	215
weighted avg	0.886	0.888	0.886	215

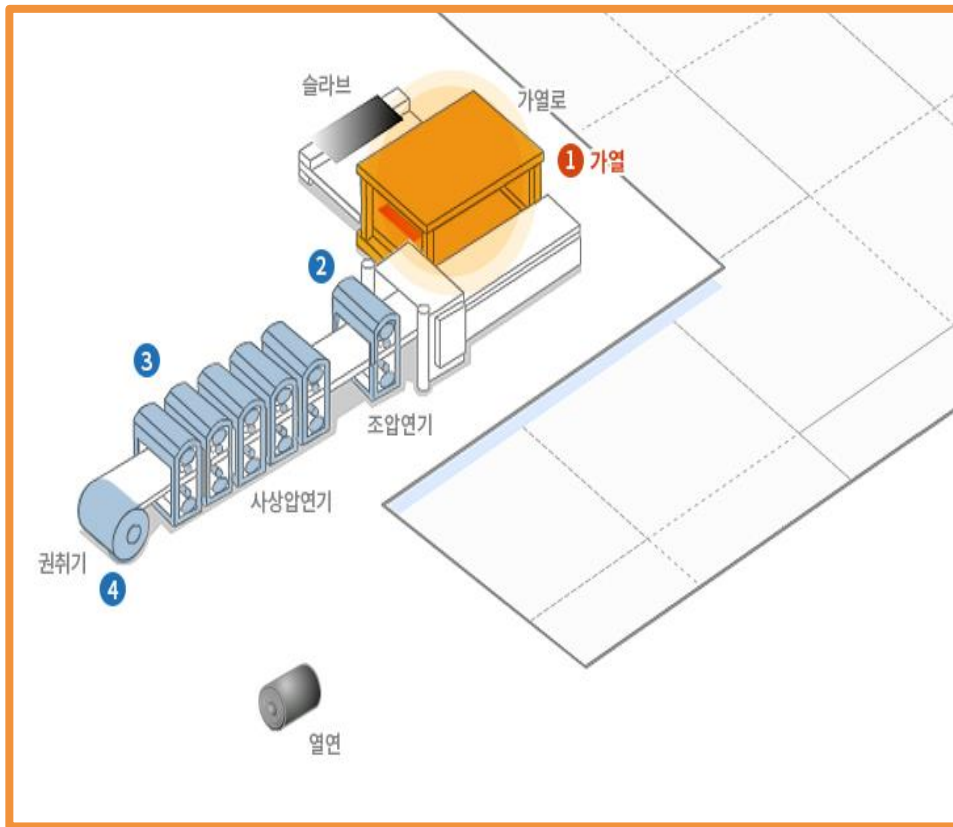
- 최종 구성된 데이터셋을 통해 Scale 불량 분류 모델을 생성한 결과, **그래디언트 부스팅 모델을 최종적으로 채택한다.**

- 최종 성능**
→ Train data : **96.1 %**
→ Test Data : **90.2%**

- 그래디언트 부스팅 모델은 불량에 대한 재현율, f1-score, 정확도에서 다른 모델보다 더 좋은 성능을 보였으므로 최종 채택되었다.

분석 결과 해석

최종모델에서의 영향인자 간 중요도 파악 결과, 주성분 중 Scale 불량에 가장 많은 영향을 미치는 것은 주성분2, 주성분3이었으며 HSB와 강종 역시 중요한 원인임을 알 수 있었다. 네 가지 원인을 구체적으로 분석하여 Scale 불량 개선 방안을 수립하고자 한다.



- 압연공정은 슬라브에 강한 열을 가하면 공기 중으로 철이 노출되면서 산화 스케일링이 형성되고 이러한 불순물을 제거하기 위해 HSB에 통과시키고 고압수를 분사하여 descaling 과정을 거친다.
- 따라서 HSB 적용 여부에 따라 Scale 불량률이 크게 달라진다. 그러나 descaling 횟수는 scaling 불량과 연관은 있으나 그 관계가 선형적인 것은 아니다.
- Plate의 두께 및 강종에 따라 적합한 descaling 패턴과 횟수가 모두 다르기 때문이다. 마찬가지로 Plate의 물리적인 성질마다 scale 불량을 최소화 하기 위한 열전달 시간 역시 상이하므로 가열대의 재로시간을 Plate의 성질에 맞춰 조정할 필요가 있다.

후판 압연 온라인 온도예측 모델

압연공정이 진행되는 중 **Plate 별로** 가열로 가열대와 균열대에서의 온도 변화를 실시간으로 모니터링하여 관리자에게 **가열로 온도 변화를 예측하여 전달**한다. 또한 Plate 별 Scale 불량 최소화를 위해 **적정 온도**를 함께 전달하여 관리자들로 하여금 불량률을 최소화 하도록 한다.

Plate 분류 시스템을 통한 스마트 Descaling 모델 개발

Plate의 두께, 강종 등 물리적인 성질에 따라 효율적인 descaling 횟수가 달라진다. 따라서 공정이 진행되는 중 **Plate의 상태와 물리적인 성질을 모니터링하고 분류함**으로써 적절한 descaling 횟수를 전달하여 자동으로 **descaling 횟수를 조절**할 수 있는 스마트 descaling 모델을 개발한다.

감사합니다.