철 스크랩 가격 예측 웹 서비스

STEEL CUT

CONTENTS

01. 프로젝트 개요

02. 프로젝트 기획

03. 프로젝트 모델링

04. 웹 서비스 개발

05. 프로젝트 결론

01. 프로젝트 개요

기획 배경

철강 제조의 핵심 매개체:
고로 (용광로)

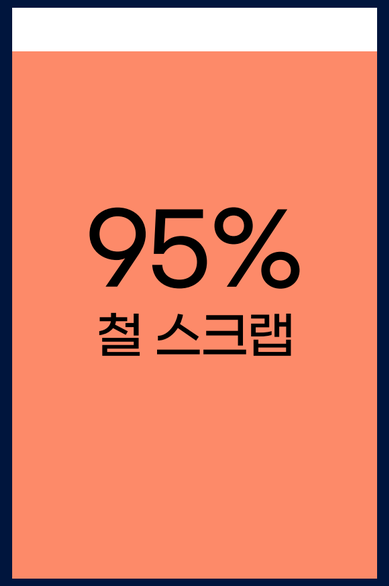
그러나 과도한 이산화탄소 배출량으로 인해
환경 파괴의 주범으로 꼽힘
친환경적 ‘ 전기로 ’

전기로의 핵심 재료:

철 스크랩

기획 배경

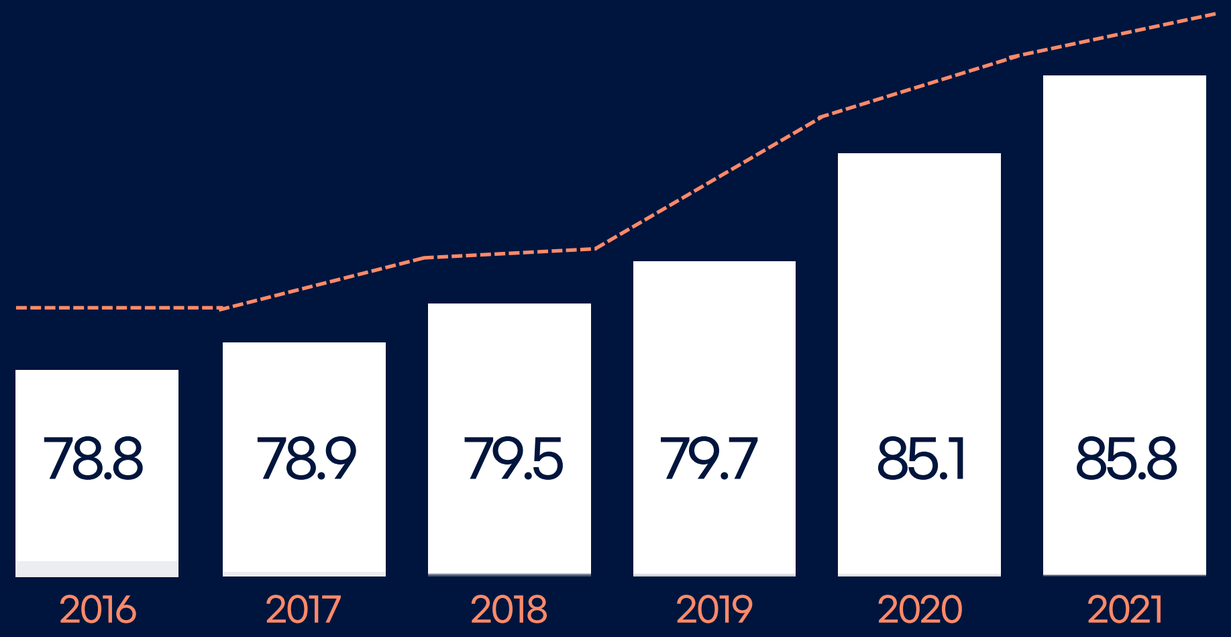
친환경 전기로 제강의 철원 비율



전기로 사용으로 이산화탄소 배출 60% 감소,
폐기물 80% 감소

→ **철 스크랩을 천연 자원으로 인식**

철 스크랩 자급률 추이 (%)



철 스크랩 수요가 계속 **증가하는 추세**

출처: FerroTimes

기획동기

PROFIT

새로운 이익 기회 창출

PURPOSE

기업과 사회를 위한 목적 달성

ENVIRONMENT

범세계적 탄소 배출 감축 패러다임 선도

개발 목표 및 방향성

웹 플랫폼

철 스크랩 가격 예측과
통계 데이터 대시보드 제공

철강 기업 & 이해관계자

가격 예측을 통한
적절한 매입 시기 확보

개발 목표

철강산업 활성화 및
기업 경쟁력 강화

개발 환경

프로그래밍 언어

python™

CSS HTML JS

머신러닝 & 딥러닝

scikit learn

Darts

Keras

AutoGluon

DB & Frameworks

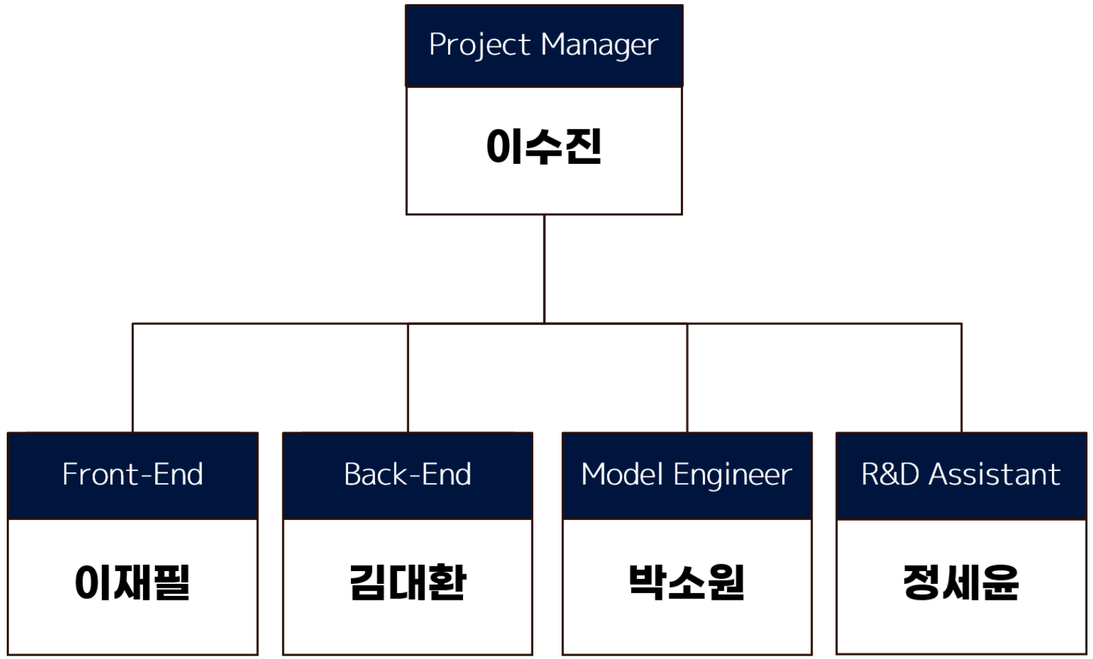
PostgreSQL

node JS

React

02. 프로젝트 기획

프로젝트 팀원 소개



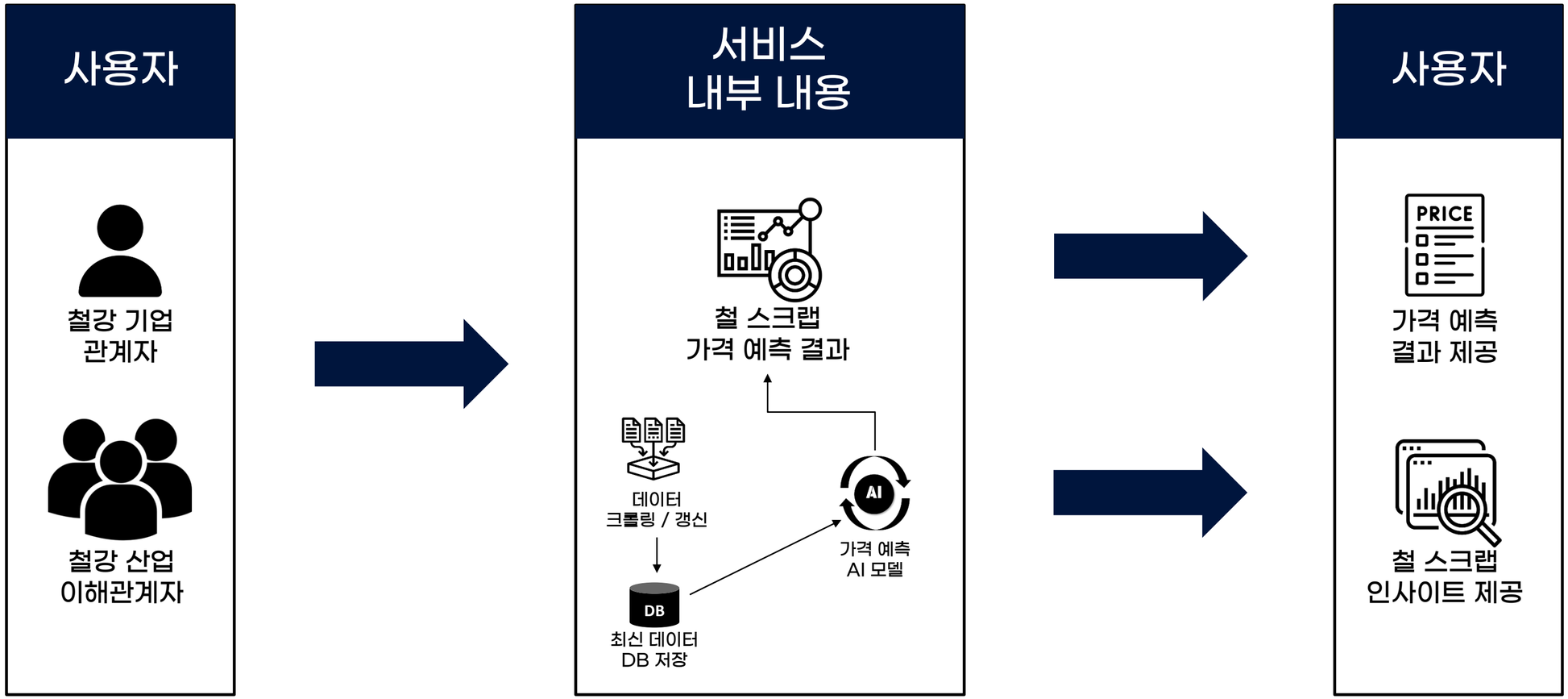
구분	인력별 역할	
	담당자	역할
프로젝트 관리	이수진	일정관리, 문서작성, 개발 보조
데이터 수집	전원	타겟 및 특성 데이터 크롤링
	이재필, 김대환, 이수진, 정세윤	데이터셋 구성
분석/설계	이재필, 김대환, 이수진	데이터 전처리 및 EDA
개발 (Server-Side)	김대환	서버사이드 웹 구축 및 DB 구축/관리
	이재필	서버사이드 웹 데이터 및 모델 적용
개발 (AI Model)	박소원, 이재필, 김대환	AI 모델 기술조사
	박소원. 이재필, 김대환, 정세윤	AI 모델링 테스트
	박소원	채택 모델 하이퍼파라미터 튜닝
개발 (Client-Side)	이재필	클라이언트사이드 웹 구축 및 웹 디자인
테스트	전원	모델 테스트 진행, 디버깅, 유지보수

개발 일정 및 진행 과정

개발 기간
2023.10.30 - 2023.12.14

1	1. 비즈니스 이해 및 범위 설정 2. 프로젝트 정의 및 계획 설정	2023년 10월 말
2	1. 수집 데이터 정의 2. 데이터 수집 및 정합성 검증 3. 모델 탐색 및 모델링 진행 4. 모델 평가 및 검증	2023년 11월
3	1. 웹 플랫폼 프로토타입 제작 및 테스트 2. 최종 웹 플랫폼 확정	2023년 12월 초
4	1. 모델 발전 계획 수립 2. 프로젝트 산출물 발표	2023년 12월 중

시스템 아키텍쳐 설계 및 구현



점검시간

매주 일요일 오후 7시

03. 프로젝트 모델링

수집 데이터 정의

종속변수

- 생철
- 중량A
- 경량A
- 선반A
- 선반C

독립변수

- 국제유가(Dubai)
- 무역수지
- 금 가격
- 수입물가지수
- 철광석 수입 금액, 수입 물량
- 금속광물 수입 금액, 수입 물량
- 환율: 달러, 엔, 유로, 위안
- 경기 선행지수: G20, 한국, 미국, 중국
- 전기요금 월별 평균단가

데이터 수집 과정 - 종속변수

Selenium 라이브러리를 활용하여 동적 크롤링

철스크랩(Scrap) ▼

ㄴ 국내(생철A)

ㄴ 국내(중량A)

ㄴ 국내(경량A)

ㄴ 국내(선반)

- 출처 : 스틸데일리(www.steeldaily.co.kr)

날짜	도매 유통가	소매 유통가	수입유통가 중국산	수입유통가 베트남	수입유통가 바레인	수입유통가 일본	비고
2023-12-08	1,130 ▼ -5	1,160 - 0	1,000 ▼ -30	1,070 ▼ -10	1,030 ▼ -20	1,030 ▼ -20	*일본산, 중국산 10m, 12m 제품 가격 상이 유의
2023-12-01	1,135 ▼ -5	1,160 - 0	1,030 - 0	1,080 ▼ -10	1,050 - 0	1,050 - 0	
2023-11-24	1,140 ▼ -5	1,160 ▼ -10	1,030 - 0	1,090 - 0	1,050 - 0	1,050 - 0	
2023-11-17	1,145 ▼ -10	1,170 - 0	1,030 - 0	1,090 - 0	1,050 - 0	1,050 - 0	
2023-11-10	1,155 ▼ -15	1,170 ▼ -10	1,030 ▼ -10	1,090 ▼ -10	1,050 ▼ -20	1,050 ▼ -10	
2023-11-03	1,170 ▼ -10	1,180 ▼ -10	1,040 - 0	1,100 ▼ -10	1,070 - 0	1,060 - 0	
2023-10-27	1,180 ▼ -10	1,190 ▼ -10	1,040 - 0	1,110 ▼ -10	1,070 - 0	1,060 - 0	
2023-10-20	1,190 ▲ 10	1,200 ▲ 10	1,040 - 0	1,120 ▲ 30	1,070 - 0	1,060 - 0	
2023-10-13	1,180 - 0	1,190 - 0	1,040 - 0	1,090 - 0	1,070 - 0	1,060 - 0	※가격인상 전후 혼조세 유의
2023-10-06	1,180 - 0	1,190 - 0	1,040 - 0	1,090 - 0	1,070 - 0	1,060 - 0	
2023-09-29	1,180 - 0	1,190 - 0	1,040 - 0	1,090 - 0	1,070 - 0	1,060 - 0	
2023-09-22	1,180 ▼ -10	1,190 - 0	1,040 - 0	1,090 - 0	1,070 - 0	1,060 - 0	
2023-09-15	1,190 ▲ 35	1,190 ▲ 20	1,040 ▼ -10	1,090 ▼ -10	1,070 ▼ -10	1,060 ▼ -20	
2023-09-08	1,155 - 0	1,170 - 0	1,050 - 0	1,100 ▼ -10	1,080 ▼ -10	1,080 ▼ -10	
2023-09-01	1,155 - 0	1,170 - 0	1,050 ▼ -10	1,110 - 0	1,090 ▼ -10	1,090 ▼ -10	
2023-08-25	1,155 ▼ -10	1,170 ▼ -10	1,060 - 0	1,110 ▼ -10	1,100 - 0	1,100 - 0	
2023-08-18	1,165 ▼ -10	1,180 - 0	1,060 ▼ -10	1,120 - 0	1,100 ▼ -10	1,100 ▼ -10	
2023-08-11	1,175 ▼ -10	1,180 ▼ -10	1,070 - 0	1,120 ▼ -10	1,110 - 0	1,110 - 0	

데이터 수집 과정 - 독립변수

소스 사이트에서 제공하는 csv , xlsx 형식의 파일 다운로드를 통해 수집

철광석 과거 데이터

기간

Daily

데이터 다운로드 2023- 11- 13 - 2023- 12- 13

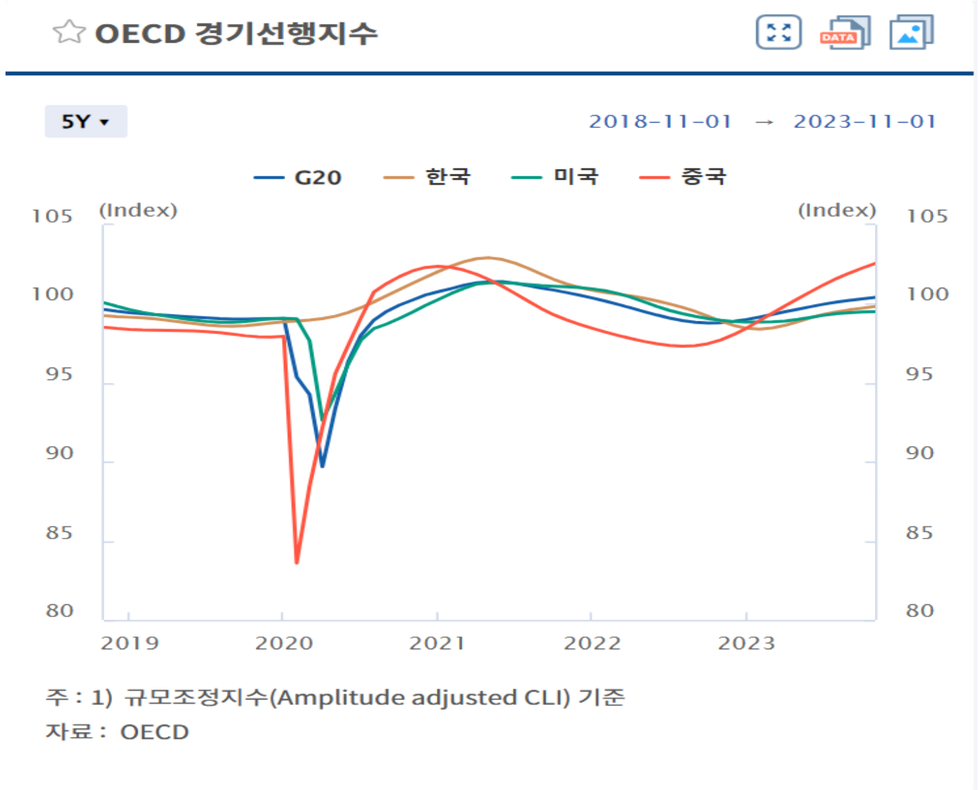
날짜	종가	시가	고가	저가	거래량	변동 %
2023- 12- 12	135.66	135.66	135.66	135.66		+0.37%
2023- 12- 11	135.16	135.16	135.16	135.16		-0.21%
2023- 12- 08	135.45	135.45	135.45	135.45		+0.78%
2023- 12- 07	134.40	134.40	134.40	134.40		+1.11%
2023- 12- 06	132.92	132.92	132.92	132.92		+1.64%
2023- 12- 05	130.77	130.77	130.77	130.77		+0.24%
2023- 12- 04	130.46	130.46	130.46	130.46		0.00%
2023- 12- 01	130.46	130.46	130.46	130.46		0.00%
2023- 11- 30	130.46	130.46	130.46	130.46		+0.06%
2023- 11- 29	130.38	130.38	130.38	130.38		-0.02%
2023- 11- 28	130.40	130.40	130.40	130.40		-0.02%
2023- 11- 27	130.42	130.42	130.42	130.42		+0.20%
2023- 11- 24	130.16	130.16	130.16	130.16		+0.25%

	A	B	C	D	E	F	G
1	날짜	종가	시가	고가	저가	거래량	변동 %
2	2023-12-12	135,66	135,66	135,66	135,66		0,37%
3	2023-12-11	135,16	135,16	135,16	135,16		-0,0021
4	2023-12-08	135,45	135,45	135,45	135,45		0,78%
5	2023-12-07	134,4	134,4	134,4	134,4		1,11%
6	2023-12-06	132,92	132,92	132,92	132,92		1,64%
7	2023-12-05	130,77	130,77	130,77	130,77		0,24%
8	2023-12-04	130,46	130,46	130,46	130,46		0,00%
9	2023-12-01	130,46	130,46	130,46	130,46		0,00%
10	2023-11-30	130,46	130,46	130,46	130,46		0,06%
11	2023-11-29	130,38	130,38	130,38	130,38		-0,0002
12	2023-11-28	130,4	130,4	130,4	130,4		-0,0002
13	2023-11-27	130,42	130,42	130,42	130,42		0,20%
14	2023-11-24	130,16	130,16	130,16	130,16		0,25%
15	2023-11-22	129,84	129,84	129,84	129,84		0,22%
16	2023-11-21	129,55	129,55	129,55	129,55		0,19%
17	2023-11-20	129,31	129,31	129,31	129,31		0,28%
18	2023-11-17	128,95	128,95	128,95	128,95		-0,0036
19	2023-11-16	129,42	129,42	129,42	129,42		0,13%
20	2023-11-15	129,25	129,25	129,25	129,25		0,36%
21	2023-11-14	128,78	128,78	128,78	128,78		0,34%
22	2023-11-13	128,34	128,34	128,34	128,34		0,21%

철광석
출처 : 주식시장(https://kr.investing.com)
csv파일

데이터 수집 과정 - 독립변수

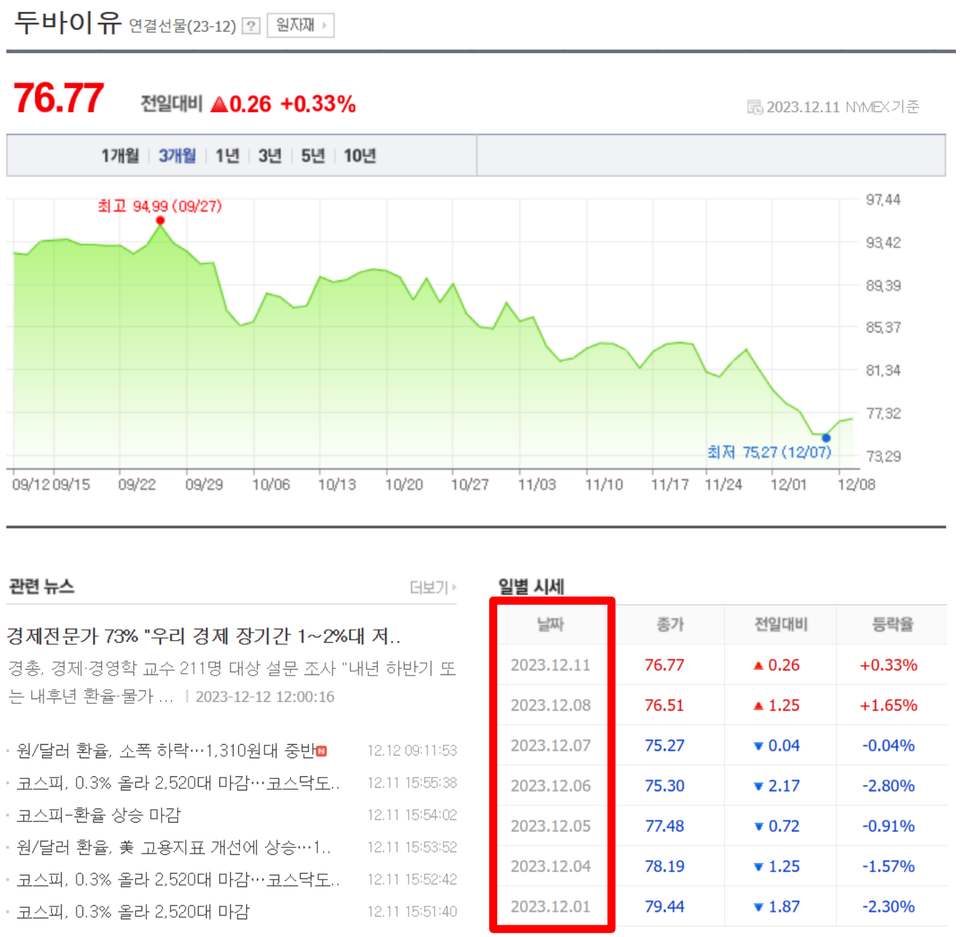
소스 사이트에서 제공하는 csv , xlsx 형식의 파일 다운로드를 통해 수집



A	B	C	D	E
OECD 경기선행지수				
주석	1) 규모조정지수(Amplitude adjusted CLI) 기준			
출처	OECD			
범례명	G20	한국	미국	중국
단위	(Index)	(Index)	(Index)	(Index)
주기	월	월	월	월
기간	1961-01 ~ 2023-11	1990-01 ~ 2023-11	1955-01 ~ 2023-11	1992-05 ~ 2023-11
2005-06	100.0649	100.2135	100.3454	100.3117
2005-07	100.087	100.4173	100.3949	100.3589
2005-08	100.1371	100.6497	100.4662	100.3733
2005-09	100.2129	100.87	100.573	100.3503
2005-10	100.317	101.0443	100.7265	100.3165
2005-11	100.4457	101.1388	100.9116	100.2973
2005-12	100.5814	101.1386	101.0878	100.3051
2006-01	100.7062	101.0435	101.2236	100.3464
2006-02	100.8109	100.885	101.3089	100.4259
2006-03	100.892	100.6971	101.3381	100.5432
2006-04	100.9464	100.508	101.3082	100.6901
2006-05	100.9756	100.3402	101.2442	100.8362
2006-06	100.9936	100.2255	101.186	100.9488
2006-07	101.02	100.195	101.1531	101.0287
2006-08	101.0784	100.2446	101.1669	101.1195
2006-09	101.1805	100.3167	101.2337	101.2772
2006-10	101.3204	100.365	101.3326	101.5313
2006-11	101.4892	100.3706	101.4495	101.89
2006-12	101.6717	100.3588	101.5743	102.3289
2007-01	101.8531	100.3643	101.6948	102.826
2007-02	102.0414	100.4258	101.8118	103.3779
2007-03	102.2351	100.5515	101.9324	104.0012

OECD 경기선행지수
출처 : 한국은행(https://www.index.go.kr)
xlsx 파일

데이터 수집 과정 - 독립변수



기존 독립변수인 환율, 금, 국제유가와 추가 독립변수인 비철금속의 정보를 네이버 증권에서 크롤링

BeautifulSoup4 라이브러리를 활용해 자동화된 데이터 수집 파이프라인의 개발 계획을 수립

데이터 전처리

리샘플링(resample) ?

- 시계열 데이터에서 일정 간격으로 데이터를 선택하거나, 선택된 데이터를 요약하여 데이터의 수를 조정하는 과정

893 rows × 6 columns → 6298 rows × 6 columns

893건의 데이터를 6298건으로 확장

선형보간 ?

- 누락된 값이 두 개의 알려진 값 사이에 있을 때, 그 누락된 값을 추정하는 방법

DB 구축

PostgreSQL

ERD(Entity Relation Diagram)

종속변수

busheling
date
ba_k
ba_k_ud
ba_y
ba_y_ud
ba_m
ba_m_ud

heavy_scrap
date
hsa_k
hsa_k_ud
hsa_y
hsa_y_ud
hsa_m
hsa_m_ud

light_scrap
date
lsa_k
lsa_k_ud
lsa_y
lsa_y_ud
lsa_m
lsa_m_ud

turning_scrap
date
tsa_k
tsa_k_ud
tsa_y
tsa_y_ud
tsa_m
tsa_m_ud
tsc_k
tsc_k_ud
tsc_y
tsc_y_ud
tsc_m
tsc_m_ud

- 생철, 중량, 경량, 선반 (원/kg)
=> 주 단위

독립변수

gold_price
date
gold_price

iron_ore
date
iron_ore

oil_price
date
oil_price

trade_balance
date
trade_balance

nickel
date
nickel

exchange_rate
date
dollar
yen
euro
yuan

aluminum_alloy
date
aluminum_alloy

plumbum
date
plumbum

stannum
date
stannum

copper
date
copper

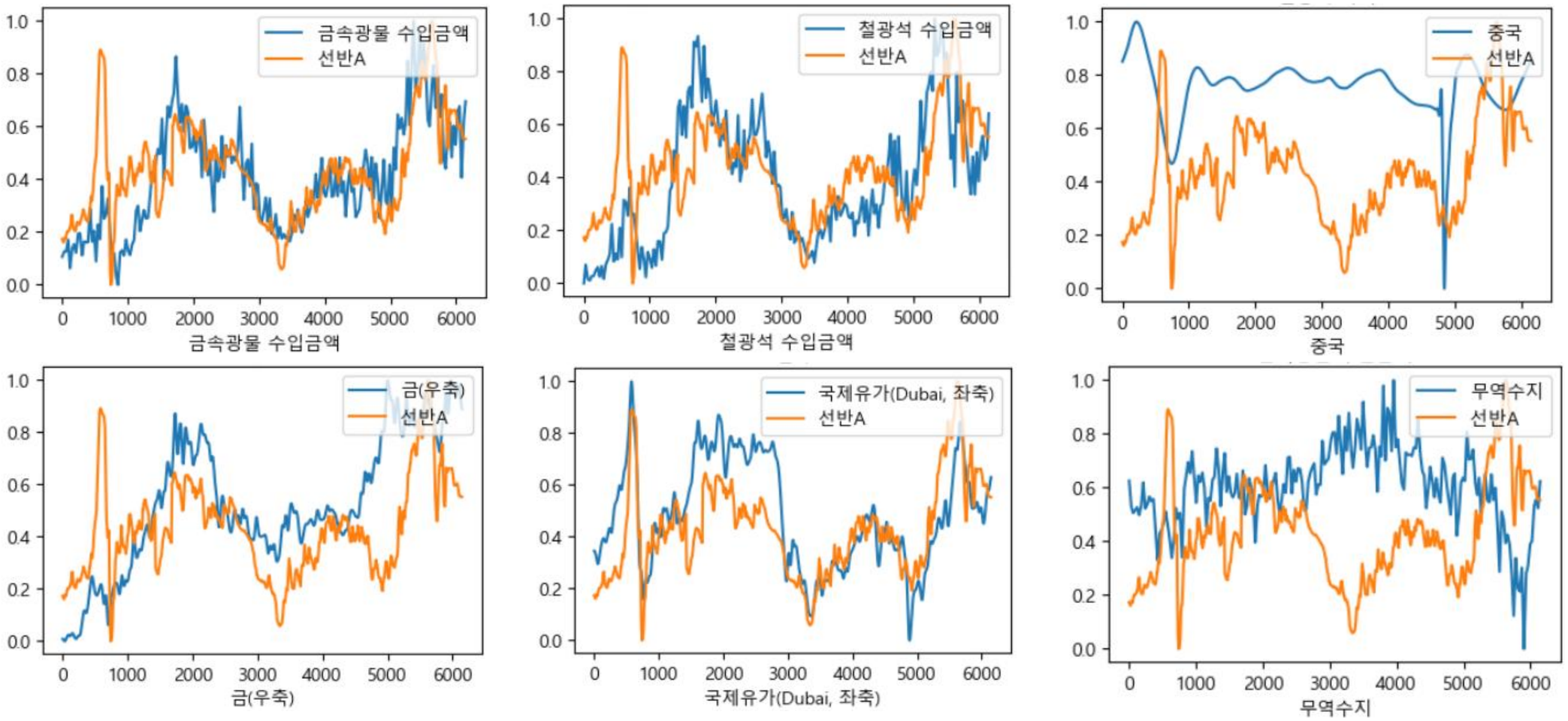
zinc
date
zinc

- 금 시세 => 일 단위
- 철광석 가격 => 일 단위
- 유가(두바이) => 일 단위
- 환율 => 일 단위
- 비철금속(구리, 납, 아연, 니켈, 알루미늄, 주석) => 일 단위
- 무역수지 => 월 단위

※ 항목별로 갱신 주기가 다름 => 통합하지 않고 각각의 테이블로 생성, PK는 date로 지정 => 데이터 최신화 시 오류발생 최소화

모델링 과정 - 앙상블

Min-Max-Scaling된 독립변수와 종속변수 그래프



※ 초기 모델링을 위한 임시 데이터 사용

모델링 과정 - 앙상블

XGBoost & RandomForest 결과

	상관계수
국제유가(Dubai, 좌축)	0.684361
금속광물 수입금액	0.654380
철광석 수입금액	0.569439
금(우축)	0.509773
수입물가지수	0.492782
무역수지	0.487305
철광석 가격	0.474755

Best correlation & Positive R2

```
RandomForest => train score : 0.9928, test score : 0.3487
** rmse : 91.1997217240516
XGBoost => train score : 0.9885, test score : 0.3589
** rmse : 90.48322060630834
독립변수 : Index(['금속광물 수입금액', '철광석 수입금액', '무역수지'], dtype='object')
```

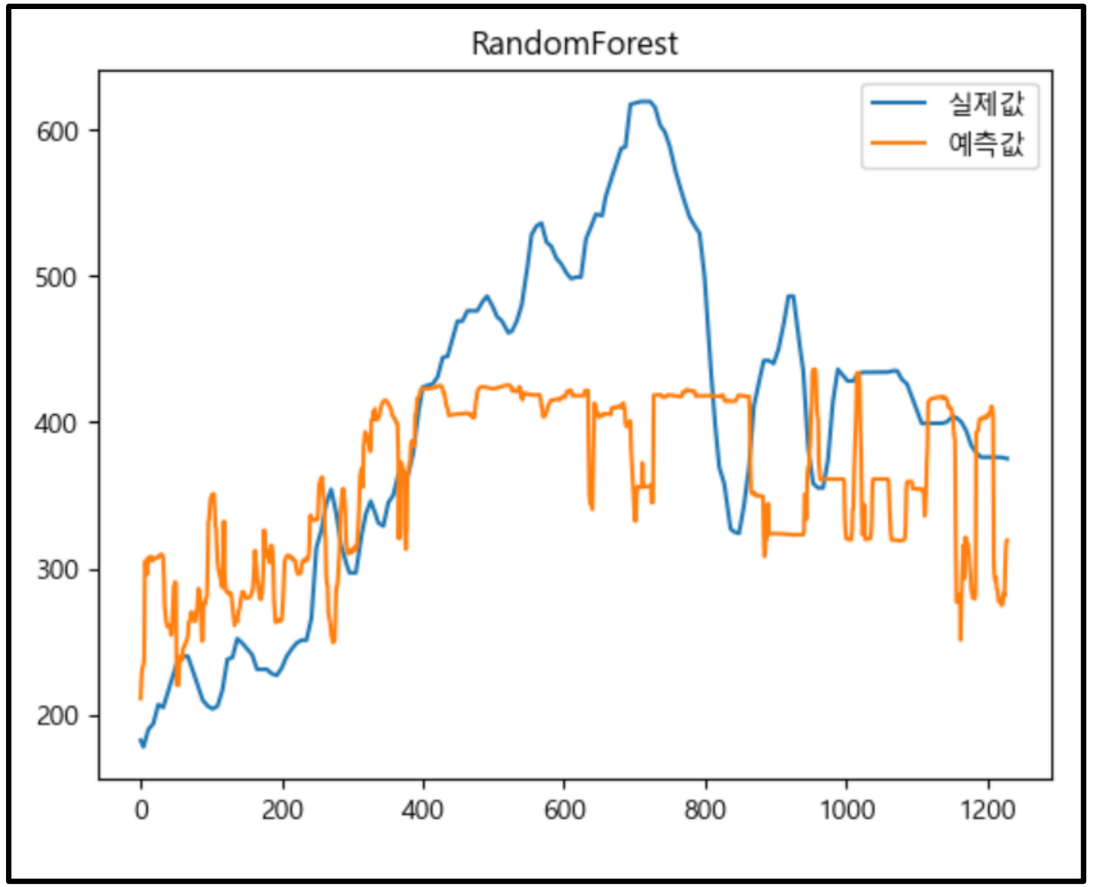
※ 상관계수가 0.45보다 높은 독립변수들의 모든 조합을 고려하여 모델링

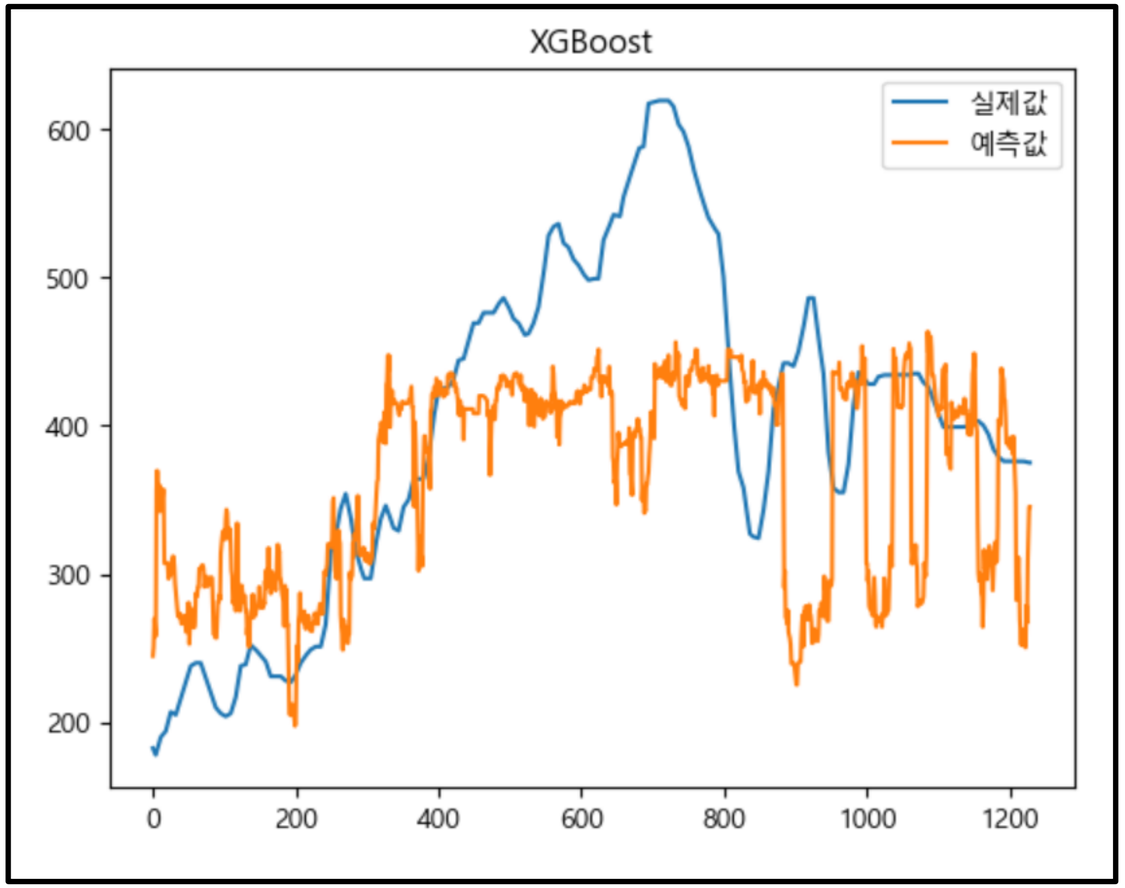
Shift 63일

```
RandomForest => train score : 0.9941, test score : 0.3300
** rmse : 91.25340289511264
XGBoost => train score : 0.9917, test score : 0.3262
** rmse : 91.51267064518305
독립변수 : Index(['금속광물 수입금액', '철광석 수입금액', '무역수지'], dtype='object')
```

모델링 과정 - 앙상블

각 모델의 예측값 & 실제값 시각화





모델링 과정 - 앙상블

모든 변수 조합으로 모델링 => 대부분 음수 결과

```
-----------------------------------------------------------------------------------------------
RandomForest => train score : 0.9981, test score : -0.9559
** rmse : 158.04872668947507
XGBoost => train score : 0.9966, test score : -0.9512
** rmse : 157.85636311674062
독립변수 : Index(['국제유가(Dubai, 좌축)', '금속광물 수입금액', '철광석 수입금액'], dtype='object')
-----------------------------------------------------------------------------------------------
RandomForest => train score : 0.9995, test score : -1.5015
** rmse : 178.73802218869628
XGBoost => train score : 0.9990, test score : -1.2980
** rmse : 171.31221884056933
독립변수 : Index(['국제유가(Dubai, 좌축)', '금속광물 수입금액', '금(우축)'], dtype='object')
-----------------------------------------------------------------------------------------------
RandomForest => train score : 0.9994, test score : -0.9436
** rmse : 157.55137594935223
XGBoost => train score : 0.9987, test score : -1.0235
** rmse : 160.75798431758602
독립변수 : Index(['국제유가(Dubai, 좌축)', '금속광물 수입금액', '수입물가지수'], dtype='object')
-----------------------------------------------------------------------------------------------
```

$$R^2 = 1 - \frac{\sum(\text{타겟} - \text{예측})^2}{\sum(\text{타겟} - \text{평균})^2}$$

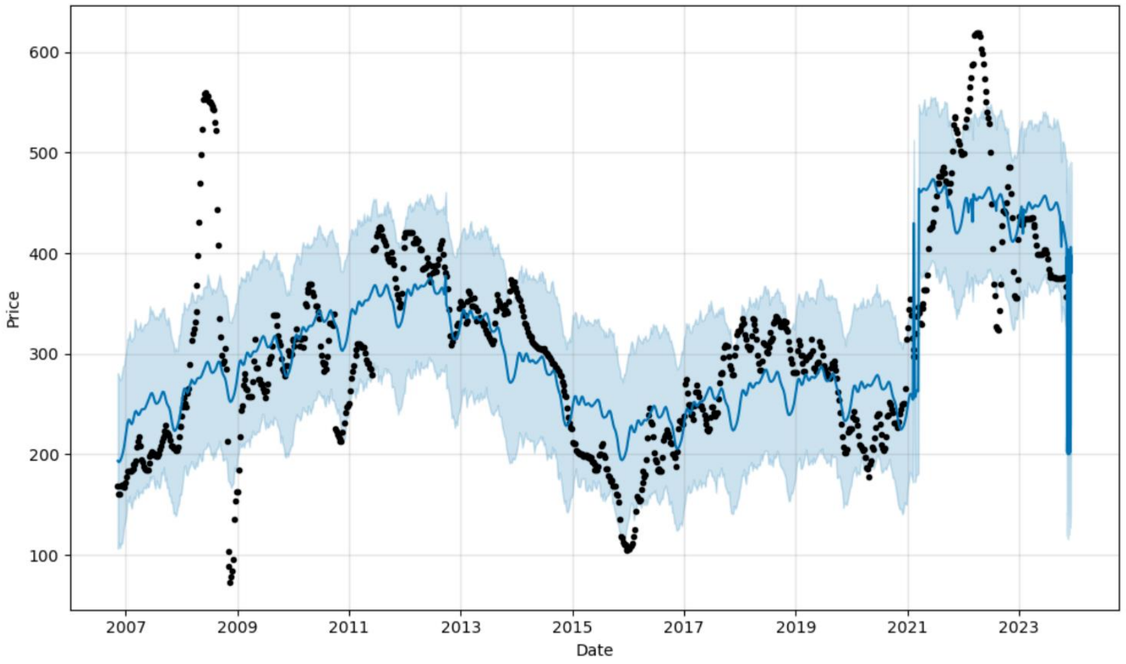
※ 앙상블 모델을 이용하여 더 이상 좋은 결과를 도출할 수 없다고 판단 => 머신러닝 모델 중 시계열 모델인 Prophet을 사용

모델링 과정 - Prophet

Prophet을 이용한 모델 학습 결과

* 선형보간 전

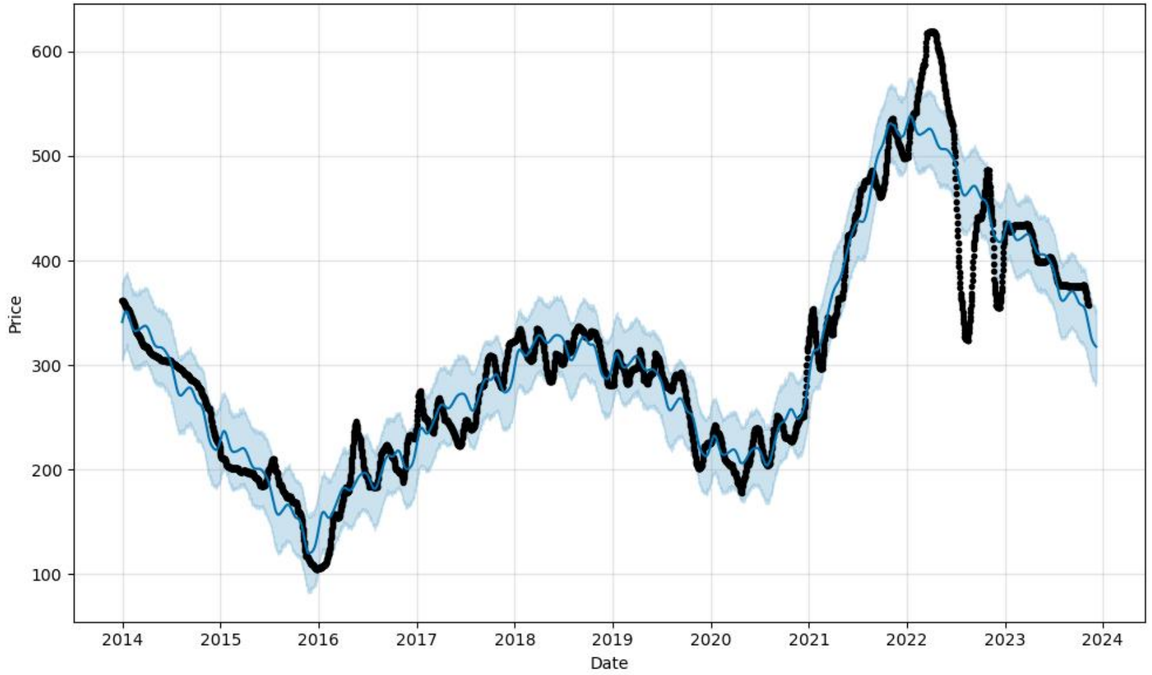
· 종속변수 => 가격
· 독립변수 => 날짜



** rmse : 63.915

* 선형보간 후

· 종속변수 => 가격
· 독립변수 => 날짜



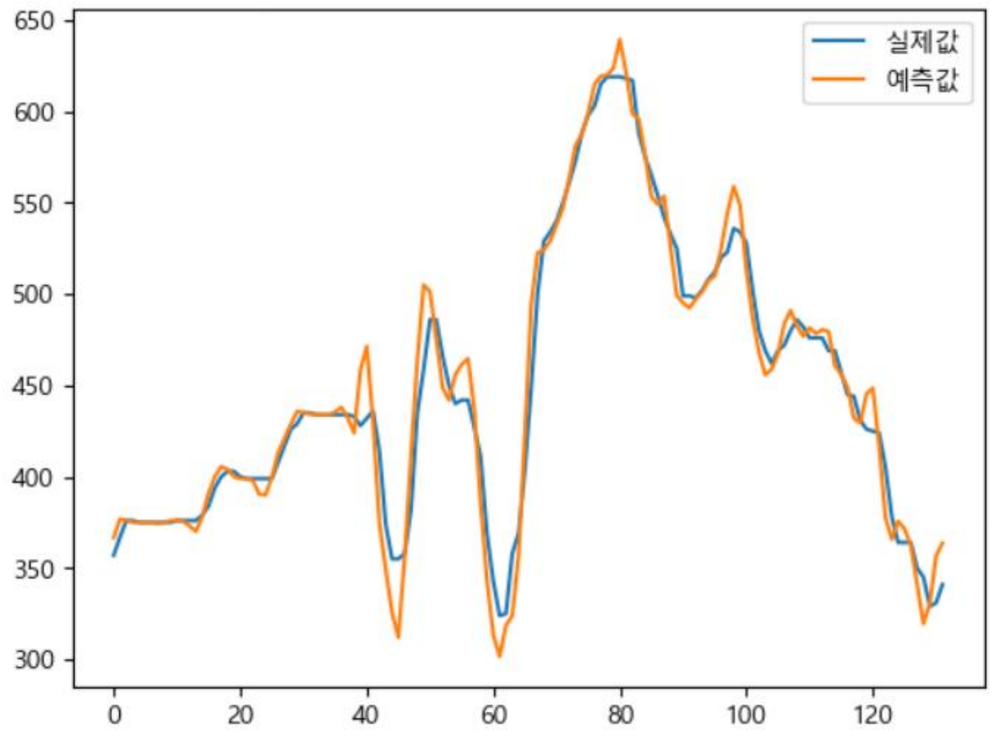
** rmse : 27.742

모델링 과정 - prophet

최적의 학습일 선정

```
-----------------------------------------
** rmse : 7.944200665955963
간격 70일 : 오차가 10 미만인 확률은 86%입니다.
-----------------------------------------
** rmse : 8.768273591682112
간격 80일 : 오차가 10 미만인 확률은 86%입니다.
-----------------------------------------
** rmse : 9.690294779400858
간격 90일 : 오차가 10 미만인 확률은 84%입니다.
-----------------------------------------
** rmse : 10.447234969140652
간격 100일 : 오차가 10 미만인 확률은 82%입니다.
-----------------------------------------
** rmse : 11.518562187698672
간격 150일 : 오차가 10 미만인 확률은 80%입니다.
-----------------------------------------
** rmse : 19.253499452897387
간격 200일 : 오차가 10 미만인 확률은 71%입니다.
-----------------------------------------
** rmse : 25.709561631835086
간격 300일 : 오차가 10 미만인 확률은 69%입니다.
-----------------------------------------
```

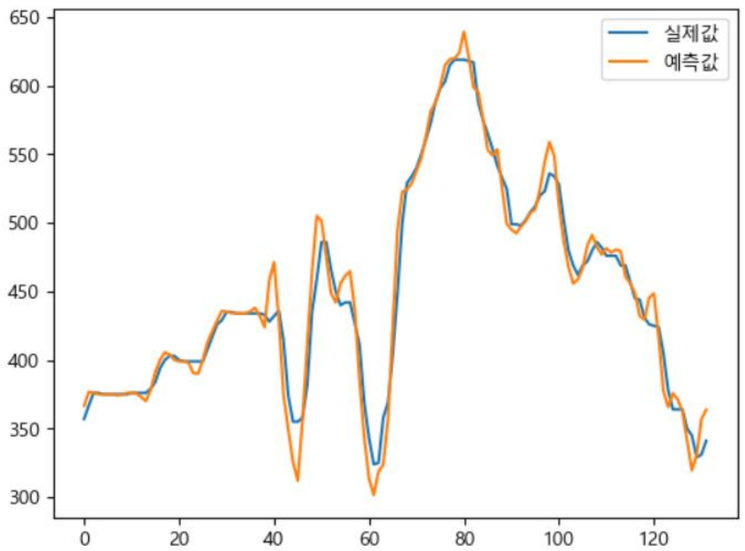
*** 80일 간격 학습 => 7일 이후 가격 예측 그래프**



모델링 과정 - prophet

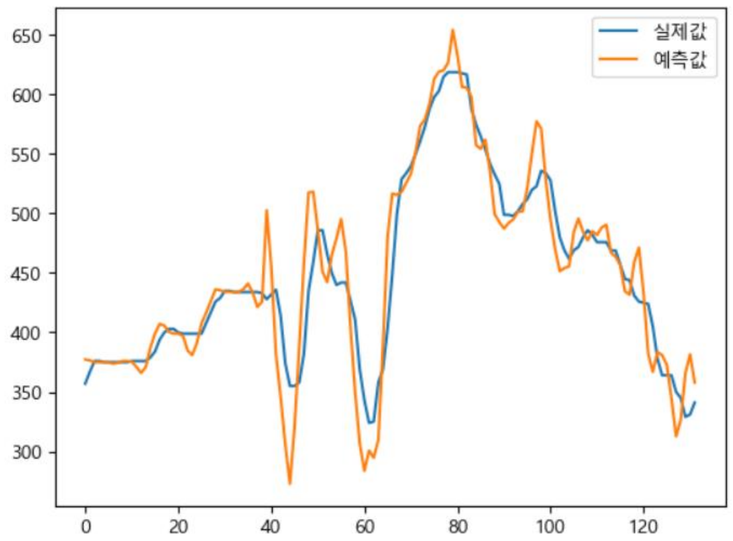
80일 학습일로 7, 14, 28일 후 데이터 예측

* 7일



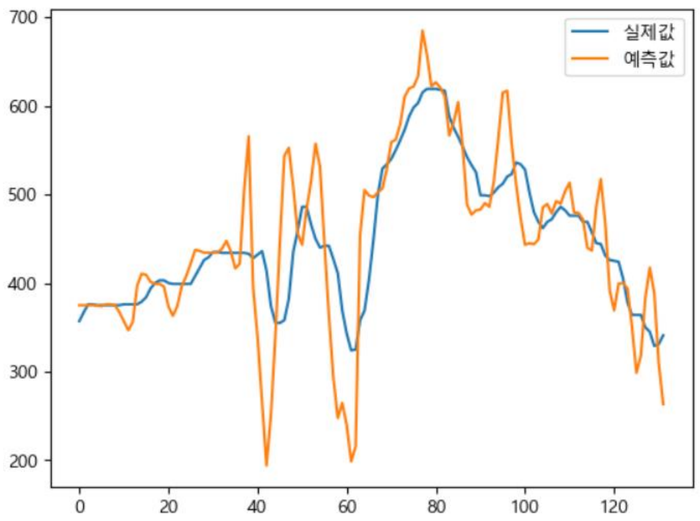
** rmse : 15
등락 예측률 : 79%

* 14일



** rmse : 28.6
등락 예측률 : 65%

* 28일



** rmse : 60.1
등락 예측률 : 50%

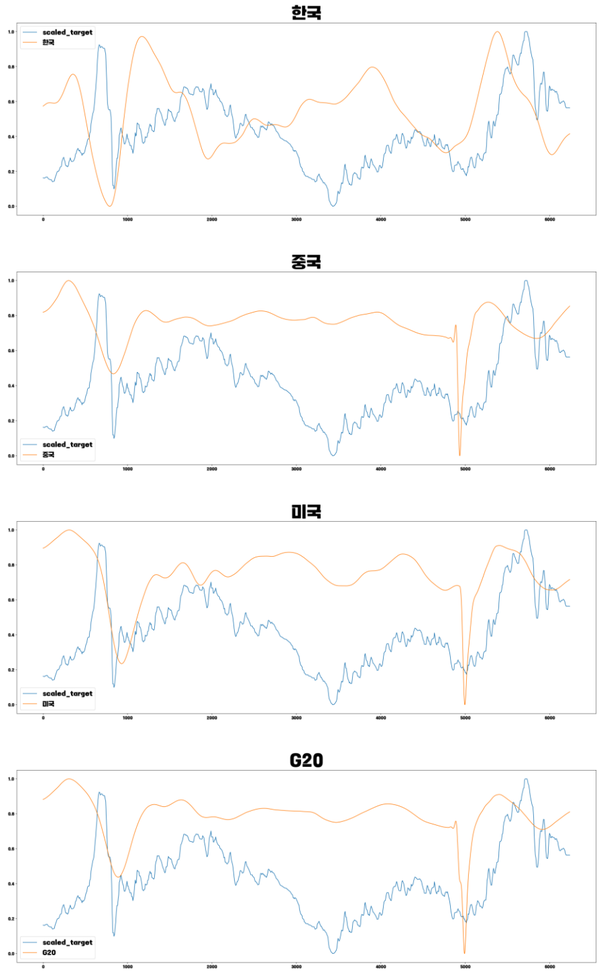
※ rmse가 목표기준인 10 보다 높아 모델 채택 불가

모델링 과정 - LSTM

각 변수의 상관계수(절댓값) 및 시각화 그래프 확인

	target
target	1.000000
국제유가(Dubai, 좌축)	0.732490
금속광물 수입금액	0.653596
철광석 수입금액	0.616530
수입물가지수	0.568250
무역수지	0.564204
철광석 가격	0.560864
금(우축)	0.469257
유로 환율	0.366996
위안 환율	0.256886
엔 환율	0.256255
금속광물 수입물량	0.209399
철광석 수입물량	0.191791
달러 환율	0.189431
한국	0.171710
중국	0.120211
미국	0.099229
G20	0.074829

[상위 4개 변수]



[하위 4개 변수]

상관성이 낮은
하위 4개
변수 제거

모델링 과정 - LSTM

모델링에 사용할 독립변수 탐색

feature	rmse	r2
('달러 환율', '금속광물 수입금액', '국제유가(Dubai, 좌축)', '금(우...	93.070935	-0.169168
('달러 환율', '금속광물 수입금액', '철광석 가격', '금(우축)')	96.365388	-0.253403
('달러 환율', '금속광물 수입금액', '철광석 수입금액', '금(우축)')	97.061768	-0.271584
('위안 환율', '금속광물 수입금액', '철광석 가격', '금(우축)')	99.931453	-0.347886
('달러 환율', '금속광물 수입금액', '금속광물 수입물량', '금(우축)')	101.285405	-0.384658
('금속광물 수입금액', '철광석 수입물량', '금(우축)')	101.581438	-0.392764
('위안 환율', '금속광물 수입금액', '철광석 수입물량', '금(우축)')	101.681247	-0.395502
('금속광물 수입금액', '철광석 가격', '금(우축)')	102.445463	-0.416558
('금속광물 수입금액', '국제유가(Dubai, 좌축)', '금(우축)')	102.656509	-0.422400
('달러 환율', '금속광물 수입금액', '수입물가지수', '금(우축)')	102.679299	-0.423032
('위안 환율', '금속광물 수입금액', '금(우축)')	102.789973	-0.426101
('달러 환율', '철광석 가격', '국제유가(Dubai, 좌축)', '금(우축)')	104.051413	-0.461318
('금속광물 수입금액', '철광석 수입금액', '철광석 수입물량', '금(우축)')	104.210777	-0.465798
('달러 환율', '철광석 수입물량', '국제유가(Dubai, 좌축)', '금(우축)')	104.771281	-0.481608
('달러 환율', '철광석 수입금액', '국제유가(Dubai, 좌축)', '금(우축)')	105.325508	-0.497324

가능한 모든 독립변수들의

조합을 생성

달러 환율, 금속광물 수입금액,

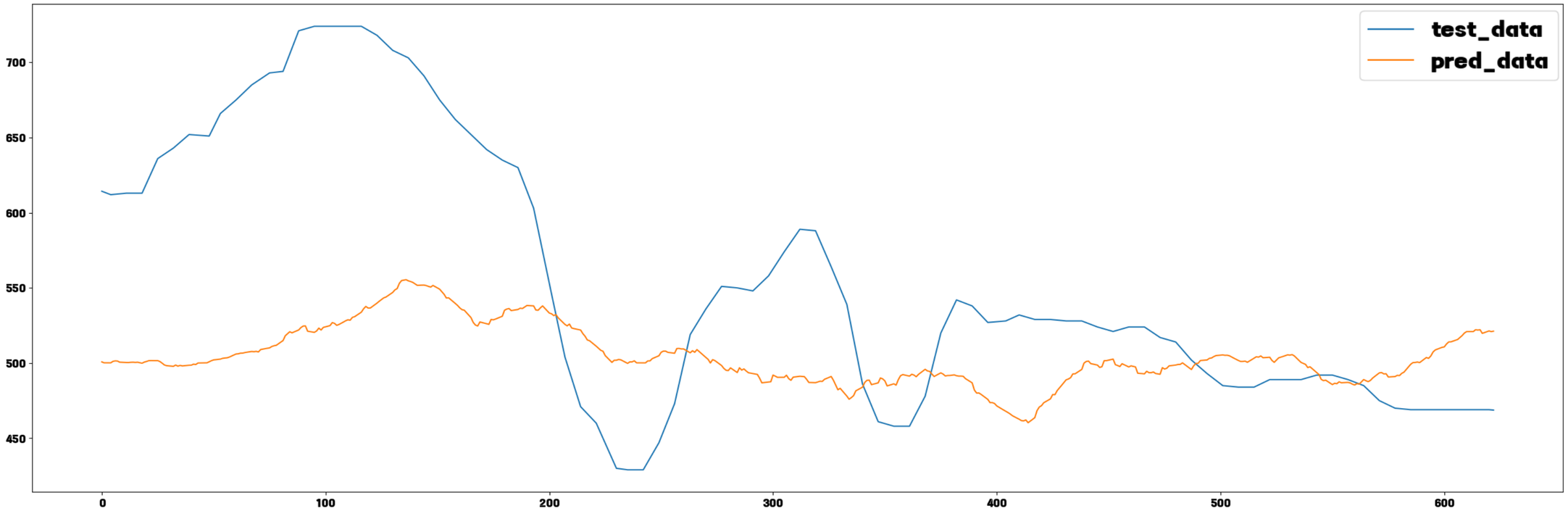
국제유가(Dubai), 금

4개의 독립변수를 적용한 모델이

상대적으로 가장 좋은 성능을 보임

모델링 과정 - LSTM

4개의 독립변수를 이용한 예측

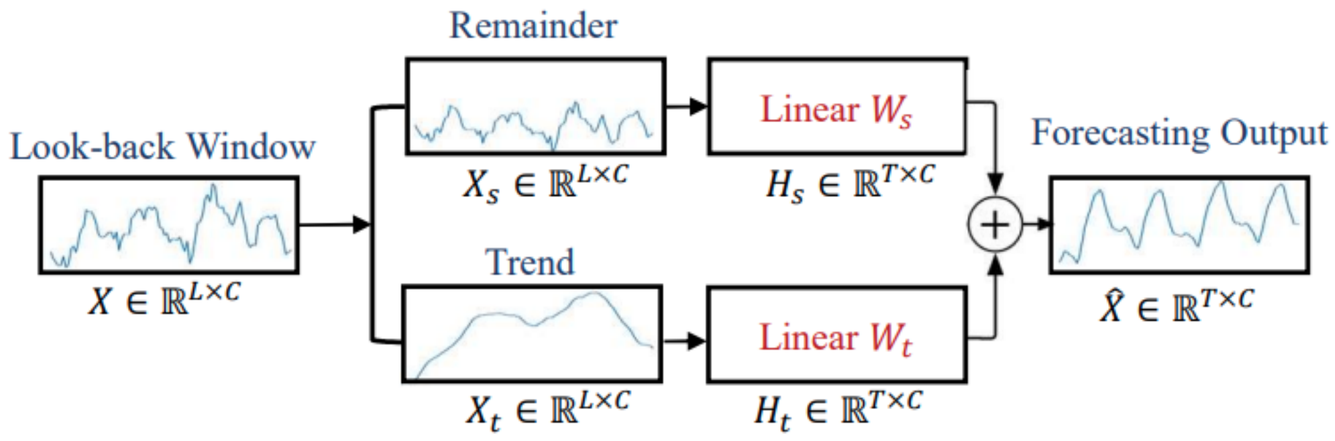



```
RMSE: 93.127
R2: -0.171
```

높은 RMSE와 음수값의 R2 도출로 **모델 사용 불가**

모델링 추가 탐색 - Darts _ DLinear

DLinear Model ?



(a) The whole structure of *DLinear*

SOTA?

- State-of-the-Art의 약어
- 현재까지 가장 우수한 결과를 달성한 모델을 말함

- 2022년 6월에 발표된 시계열 예측 모델
- 일부 LTSF(Long-term Time Series Forecasting)에서 SOTA를 달성함
- Autoformer에서 사용된 분해 방식에 선형 레이어를 결합한 단순한 구조
- Darts 라이브러리에서 DLinear 모델 사용 시 날짜와 목표값 만으로 시계열 데이터 예측 가능

Ailing Zeng, Muxi Chen, Lei Zhang, Qiang Xu
「Are Transformers Effective for Time Series Forecasting?」(26 May 2022)

모델링 추가 탐색 - Darts _ DLinear

학습 기간 탐색

days	rmse
5000	2.679566
4990	3.741673
4980	1.330035
4970	2.450540
4960	3.731692
4950	1.753705
4940	4.892329
4930	4.139029
4920	1.350448
4910	3.171271
4900	3.048857
4890	4.023536
4880	3.191058
4870	1.914212
4860	3.564494
4850	2.125740
4840	2.477774

정렬

days	rmse
2850	0.310579
620	0.454816
3770	0.475781
2990	0.789749
2980	1.071399
4230	1.073235
2620	1.079578
2550	1.082247
2920	1.088284
2780	1.096574
3140	1.134913
3600	1.138123
2860	1.140836
3650	1.165015
2540	1.175737
3450	1.215187
3940	1.235877

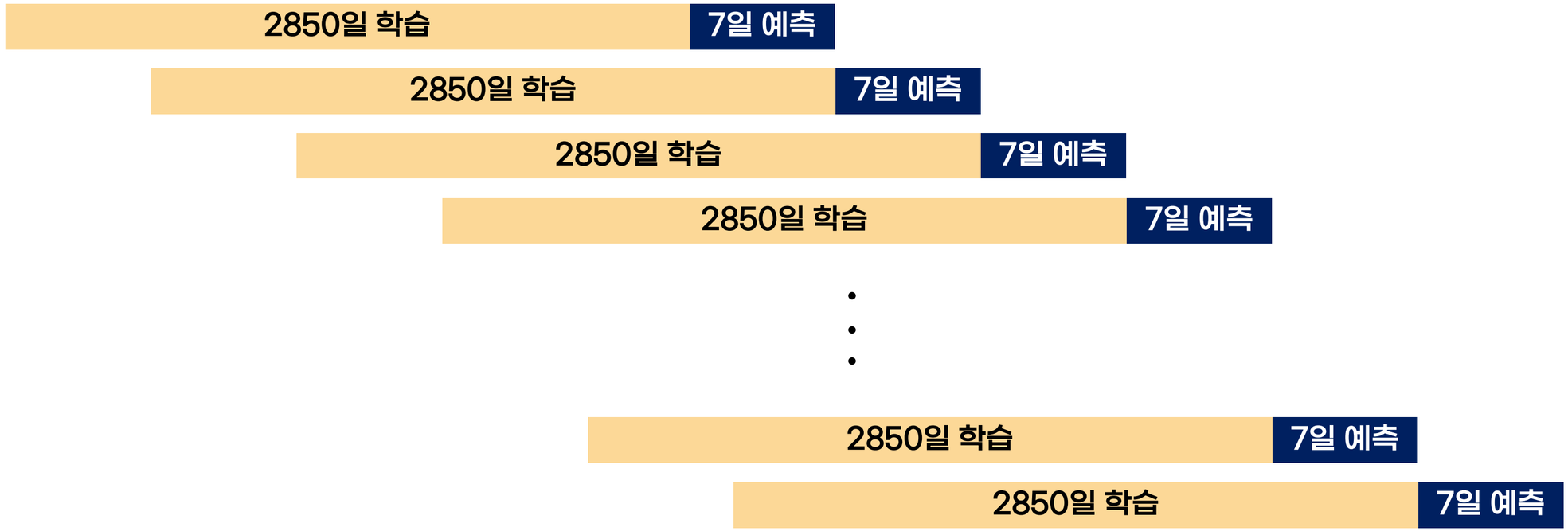
5000일 ~ 200일까지
50일 단위로 예측 모델링 진행

RMSE 확인 시 2850일이
가장 좋은 성능을 보임

학습일자 : 2850일로 설정

모델링 추가 탐색 - Darts _ DLinear

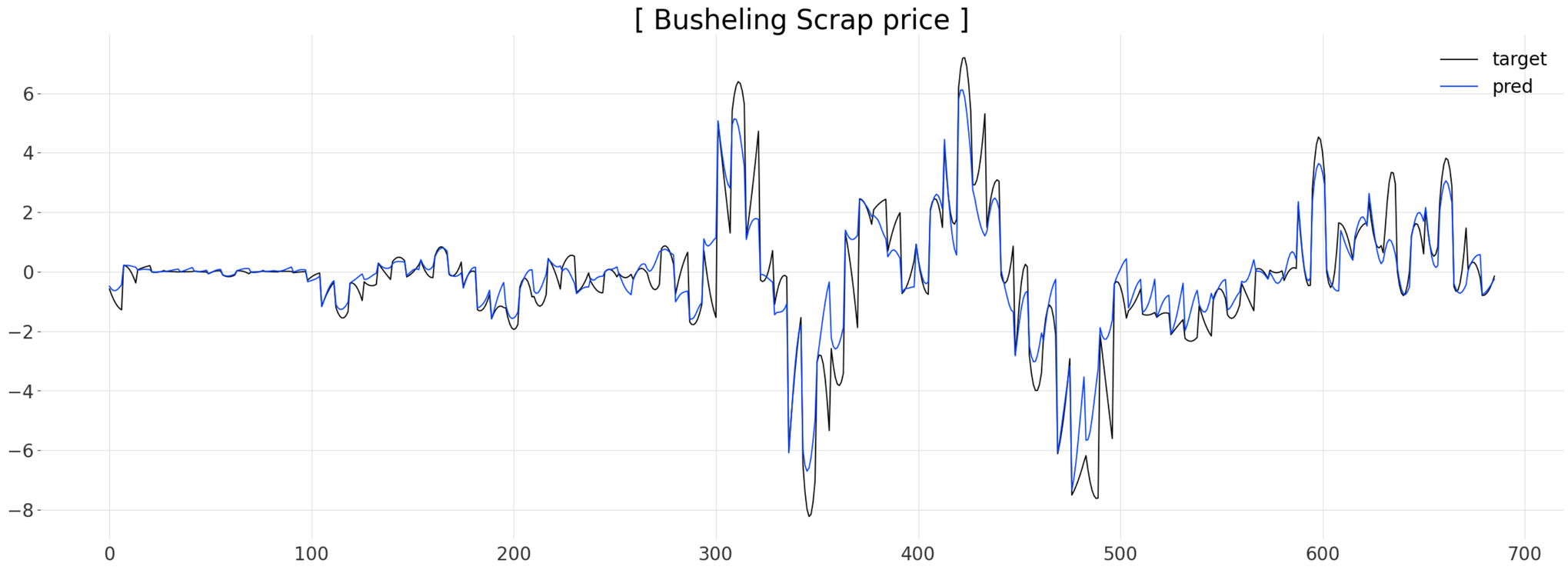
2850일 학습을 통한 7일 예측



2850일 학습 / 7일 예측을 100회 반복 » 총 700일을 예측

모델링 추가 탐색 - Darts _ DLinear

2850일 학습을 통한 7일 예측




```
RMSE : 8.250889878
R2 : 0.989435
```

이전 모델이 비해 매우 높은 성능을 보임

DLinear 모델 채택

모델링 추가 탐색 - Darts _ DLinear

문제점

Darts 라이브러리의
Dlinear 모델 사용 시
쉽게 시계열 데이터를 예측할 수 있지만,
독립변수 적용 **불가**
하이퍼 파라미터 튜닝 **어려움**

문제 발생

Time Series Made Easy in Python

Darts

AutoGluon 라이브러리의
Dlinear 모델 사용 시
쉽게 시계열 데이터를 예측할 수 있으며
독립변수 적용 **가능**
하이퍼 파라미터 튜닝 **용이**

문제 해결

AutoGluon

모델 발전 수립 - Autogluon _ DLinear

독립변수 적용한 DLinear 모델링

데이터 변수명	컬럼명	데이터 종류
구리 가격	copper	정형 데이터
주석 가격	stannum	정형 데이터
니켈 가격	nickel	정형 데이터
철광석 가격	iron ore	정형 데이터
국제 유가	oil price	정형 데이터
알루미늄 합금 가격	aluminum alloy	정형 데이터
금 가격	gold price	정형 데이터
납 가격	plumbum	정형 데이터
유로 환율	euro	정형 데이터
년도	year	파생 변수
월	month	파생 변수
일	day	파생 변수
요일	week_day	파생 변수
해당 년도의 월	year_month	파생 변수
해당 년도의 주	week	파생 변수
누적주차	week_num	파생 변수
휴일	holiday	파생 변수

9개의 정형데이터, 8개의 파생변수
총 17개의 변수를 적용하여
AutoGluon 라이브러리를 사용한
Dlinear 모델링

모델 발전 수립 - Autogluon _ DLinear

독립변수 적용한 DLinear 모델링

```
Provided dataset contains following columns:
  target:           'target'
  known covariates: ['year', 'month', 'day', 'week_day', 'year_month', 'week', 'week_num', 'holiday']

AutoGluon will gauge predictive performance using evaluation metric: 'RMSE'
  This metric's sign has been flipped to adhere to being higher_is_better. The metric score can be multiplied by -1 to get the metric value.
===================================================

Starting training. Start time is 2023-12-10 14:32:54
Models that will be trained: ['DLinear']
Training timeseries model DLinear. Training for up to 72000.0s of the 72000.0s of remaining time.
  -5.0519       = Validation score (-RMSE)
  76.03    s    = Training runtime
  0.01     s    = Validation (prediction) runtime
Not fitting ensemble as only 1 model was trained.
Training complete. Models trained: ['DLinear']
Total runtime: 76.06 s
Best model: DLinear
Best model score: -5.0519
```

```
Provided dataset contains following columns:
  target:           'target'
  known covariates: ['year', 'month', 'day', 'week_day', 'year_month', 'week', 'week_num', 'holiday']
  past covariates:  ['copper', 'stannum', 'nickel', 'iron_ore', 'oil_price', 'aluminum_alloy', 'gold_price', 'euro', 'plumbum']

AutoGluon will gauge predictive performance using evaluation metric: 'RMSE'
  This metric's sign has been flipped to adhere to being higher_is_better. The metric score can be multiplied by -1 to get the metric value.
===================================================

Starting training. Start time is 2023-12-10 17:22:03
Models that will be trained: ['DLinear']
Training timeseries model DLinear. Training for up to 72000.0s of the 72000.0s of remaining time.
  -5.0519       = Validation score (-RMSE)
  54.77    s    = Training runtime
  0.01     s    = Validation (prediction) runtime
Not fitting ensemble as only 1 model was trained.
Training complete. Models trained: ['DLinear']
Total runtime: 54.80 s
Best model: DLinear
Best model score: -5.0519
```

모델 발전 수립 - Autogluon _ DLinear

독립변수 적용한 DLinear 모델링

```
['year', 'month', 'day', 'week_day', 'year_month', 'week', 'week_num', 'holiday']
```

```
Best model score: -5.0519
```

```
['year', 'month', 'day', 'week_day', 'year_month', 'week', 'week_num', 'holiday']
['copper', 'stannum', 'nickel', 'iron_ore', 'oil_price', 'aluminum_alloy', 'gold_price', 'euro', 'plumbum']
```

```
Best model score: -5.0519
```

- Autogluon 내부 모델링 과정에서 피쳐 중요도가 낮은 변수를 자동으로 제외함
- 날짜 외 독립변수 추가 유무에 따른 score(-RMSE) 차이 없음

» 독립변수가 모델의 성능 개선에 영향을 주지 않음

모델 발전 수립 - Autogluon _ DLinear

하이퍼 파라미터 튜닝한 DLinear 모델링

```
predictor.fit(
    train_set_list[j][i],
    time_limit=1200*60,
    hyperparameters={
        "DLinear": [
            {"hidden_dimension":25, "learning_rate ":0.0004, "epochs":600, "scaling":"std",},
            ],
    },
    hyperparameter_tune_kwargs=None,
)
```

- Hidden_dimension : 25
- Learning_rate : 0.0004
- Epochs : 600

하이퍼 파라미터 튜닝 전

RMSE : 8.25

등락 예측률 : 61%

생철, 7일 예측 기준

하이퍼 파라미터 튜닝 후

RMSE : 3.81

등락 예측률 : 84%

생철, 7일 예측 기준

철 스크랩 가격 예측 최종 모델로 채택

모델 발전 수립 - Autogluon _ DLinear

DLinear 모델링 결과

[RMSE]

종속변수	생철	중량A	경량A	선반A	선반C
7일	3.81	3.69	3.75	3.6	3.49
14일	7.99	7.78	7.82	7.7	7.17
28일	20.13	21.8	21.58	20.54	20.04

[등락 예측률]

종속변수	생철	중량A	경량A	선반A	선반C
7일	84%	81%	80%	81%	82%
14일	73%	77%	70%	75%	72%
28일	62%	54%	51%	57%	57%

04. 웹 서비스 개발

웹 개발 과정

Node.js

- JavaScript 기반의 웹 프로그램 코드를 작성
- JavaScript 프로그램을 실행하는 런타임

- 다양한 JavaScript 프로그램이 등록된 Node Pakage Manager - npm 사용
- 시간 대비 효율적으로, 쉽게 프로그램을 구성

React

- Meta에서 개발한 JavaScript기반의 라이브러리
- 사용자와 상호작용할 수 있는 동적인 UI 개발

- 웹페이지의 적절한 UI icon사용과 깔끔한 디자인을 가미하기 위해 Material-UI 라이브러리 사용

시연

http://172.20.20.78:3000

05. 프로젝트 결론

기대효과

산업의 탄소중립, ESG경영 등 친환경 추세

- 지속 가능성을 추구하는 기업 입장에서는 저비용과 고효율 전략을 동시에 고려해야 합니다.
- 본 프로젝트는 철 스크랩 가격 예측을 통해 기업 및 이해관계자들에게 저비용으로 철 스크랩을 매입할 수 있는 기회를 제공합니다.
- 이를 통해, 기업은 비용을 절감하고 매입한 철 스크랩으로부터 다양한 부가가치를 창출할 수 있습니다.
- 결과적으로, 이는 철 스크랩의 원활한 수요와 공급을 촉진하고, 궁극적으로는 탄소 중립을 달성하여 환경 보호에 크게 이바지하는 결과를 가져올 것입니다.

개선사항

데이터 보수, 서버 및 AI모델링 개선

- 서버의 안정성, 사용자 편의성 및 인적 자원의 효율적인 배분을 보장하기 위해 Cron을 활용한 파이프라인 개발이 필요함을 인식하였습니다.
- 향후 서버에 대한 기술조사를 진행하여 클라이언트사 서버와 연결 프로세스를 개선하는 것을 본 프로젝트의 개선 사항으로 제시합니다.
- 본 프로젝트에서는 시계열 SOTA 모델인 Dlinear 모델을 활용하였으나, 기술 발전과 함께 효과적인 모델이 출시되므로, 경쟁력 유지를 위해서 지속적으로 모델에 대한 조사 및 정확도 개선이 요구됩니다.

감사합니다

STEEL CUT