

# 국제 유가 변동에 따른 전력시장 가격 예측

Prediction of Electricity  
Market Prices by  
International Oil Prices

01

계획 수립

02

데이터 수집 및 가공

03

데이터 마이닝 시스템 구축

04

데이터 분석 결과 평가

05

웨카 장단점

## 계획 수립

### 주제 선정 배경

- ✓ **우크라이나-러시아 전쟁 → 원유가격 상승 → 휘발유, 경유 가격 상승**
  - » 특히 경유가 휘발유 가격보다 더 상승하여 **상승 원인**에 대한 궁금증 생김
  - » 전쟁으로 인한 **수출 제한** → 자연스러운 가격 상승
  - » **국제 에너지 수급 불균형** 지속됨에 따라 국제 유가와 LNG 가격 상승 지속 예측
- ✓ **우리나라 도매 전력 시장 : 공급입찰에 참여한 발전기 발전비용에 따라 가격 결정**
  - » SMP 는 원유, LNG 등 연료가격과 연동
  - » 원유가격 = LNG 도입단가 , 다른 연료가격 연관성 높음  
→ 전력시장 분석 중요 요인
  - » 국제 원유시장과 우리나라 전력시장 관계 : **장기 균형 관계**

## 계획 수립

### 문제 정의

- ✓ 국제 원유시장 기준이 되는 두바이유 가격과 우리나라 전력시장 가격(SMP)의 **연계성**에 대해 분석
  - » 원유시장과 국내 전력시장이 연계
    - 국제 유가 변동에 대한 **전력공급과 가격 예측 가능**
    - 전력 공급 계획 수립에 **불확실성 줄어듦**
  - » 따라서 국제 원유가격과 국내 전력가격의 연관계성 분석
    - 에너지 산업의 의사 결정자나 연구자에게 **중요한 정보 제공** ○
- ✓ **연료가격과 전력 가격 연계성 국내외 다수 연구 존재 하지만**
  - » 관련 연구는 LNG에 국한됨 (예 : Chae et al, 2012; 박민혁 외, 2014)
  - » 국제 원유가격과 전력가격의 연계성을 분석한 연구는 부족한 설정

# 데이터 수집 및 가공

## 수집 데이터

	날짜	종가	오픈	고가	저가	거래량	변동 %
0	2022- 11- 01	91.22	91.22	91.22	91.22	NaN	0.10%
1	2022- 10- 31	91.13	91.13	91.13	91.13	NaN	0.02%
2	2022- 10- 28	91.11	91.11	91.11	91.11	NaN	-0.18%
3	2022- 10- 27	91.27	91.27	91.27	91.27	NaN	0.12%
4	2022- 10- 26	91.16	91.16	91.16	91.16	NaN	0.35%
...	...	...	...	...	...	...	...
205	2022- 01- 07	79.53	79.53	79.53	79.53	NaN	-0.20%
206	2022- 01- 06	79.69	79.69	79.69	79.69	NaN	1.12%
207	2022- 01- 05	78.81	78.81	78.81	78.81	NaN	1.00%
208	2022- 01- 04	78.03	78.03	78.03	78.03	NaN	1.30%
209	2022- 01- 03	77.03	77.03	77.03	77.03	NaN	5.25%

## 두바이유 가격

	날짜	최대가격	최소가격	평균가격
0	2022-01-01	최대 160.07 원/kWh	최소 87.39 원/kWh	평균 125.10 원/kWh
1	2022-01-02	최대 156.17 원/kWh	최소 87.96 원/kWh	평균 114.68 원/kWh
2	2022-01-03	최대 172.74 원/kWh	최소 88.64 원/kWh	평균 151.21 원/kWh
3	2022-01-04	최대 196.49 원/kWh	최소 156.26 원/kWh	평균 162.99 원/kWh
4	2022-01-05	최대 166.88 원/kWh	최소 156.58 원/kWh	평균 161.84 원/kWh
...	...	...	...	...
150	2022-10-29	최대 263.54 원/kWh	최소 207.63 원/kWh	평균 242.85 원/kWh
151	2022-10-30	최대 283.37 원/kWh	최소 204.70 원/kWh	평균 229.65 원/kWh
152	2022-10-31	최대 296.14 원/kWh	최소 140.37 원/kWh	평균 244.80 원/kWh
153	2022-11-01	최대 292.12 원/kWh	최소 206.60 원/kWh	평균 255.26 원/kWh
154	2022-11-02	최대 284.16 원/kWh	최소 202.55 원/kWh	평균 255.84 원/kWh

## 전력 가격

	일자	송금	현찰	매매기준율	기준환율	대미 환산율
0	2022.11.23	1339.5	1376.27	1,352.60	₩100=2.40	1355.7
1	2022.11.22	1341.9	1378.71	1,355.00	₩100=5.00	1352.3
2	2022.11.21	1346.9	1383.80	1,360.00	₩100=18.50	1340.0
3	2022.11.18	1328.5	1364.97	1,341.50	₩100=10.50	1338.4
4	2022.11.17	1338.9	1375.66	1,352.00	₩100=24.00	1325.6
...	...	...	...	...	...	...
215	2022.01.07	1192.4	1225.07	1,204.00	₩100=1.00	1199.6
216	2022.01.06	1193.4	1226.08	1,205.00	₩100=7.00	1198.2
217	2022.01.05	1186.4	1218.96	1,198.00	₩100=2.00	1194.3
218	2022.01.04	1184.4	1216.93	1,196.00	₩100=3.00	1190.8
219	2022.01.03	1181.5	1213.87		1193	1185.5

## 환율

## 전처리

### 두바이유 가격

### 전력 가격

### 환율

필요한 컬럼만 뽑아내기

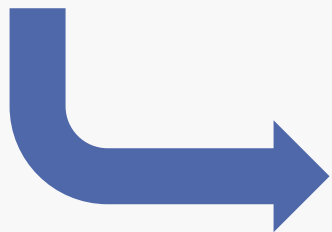
주말 데이터 결측치값 채워주기

날짜 오름차순 정렬

전력 평균 가격에서 숫자만 추출

Datetime 타입 변경

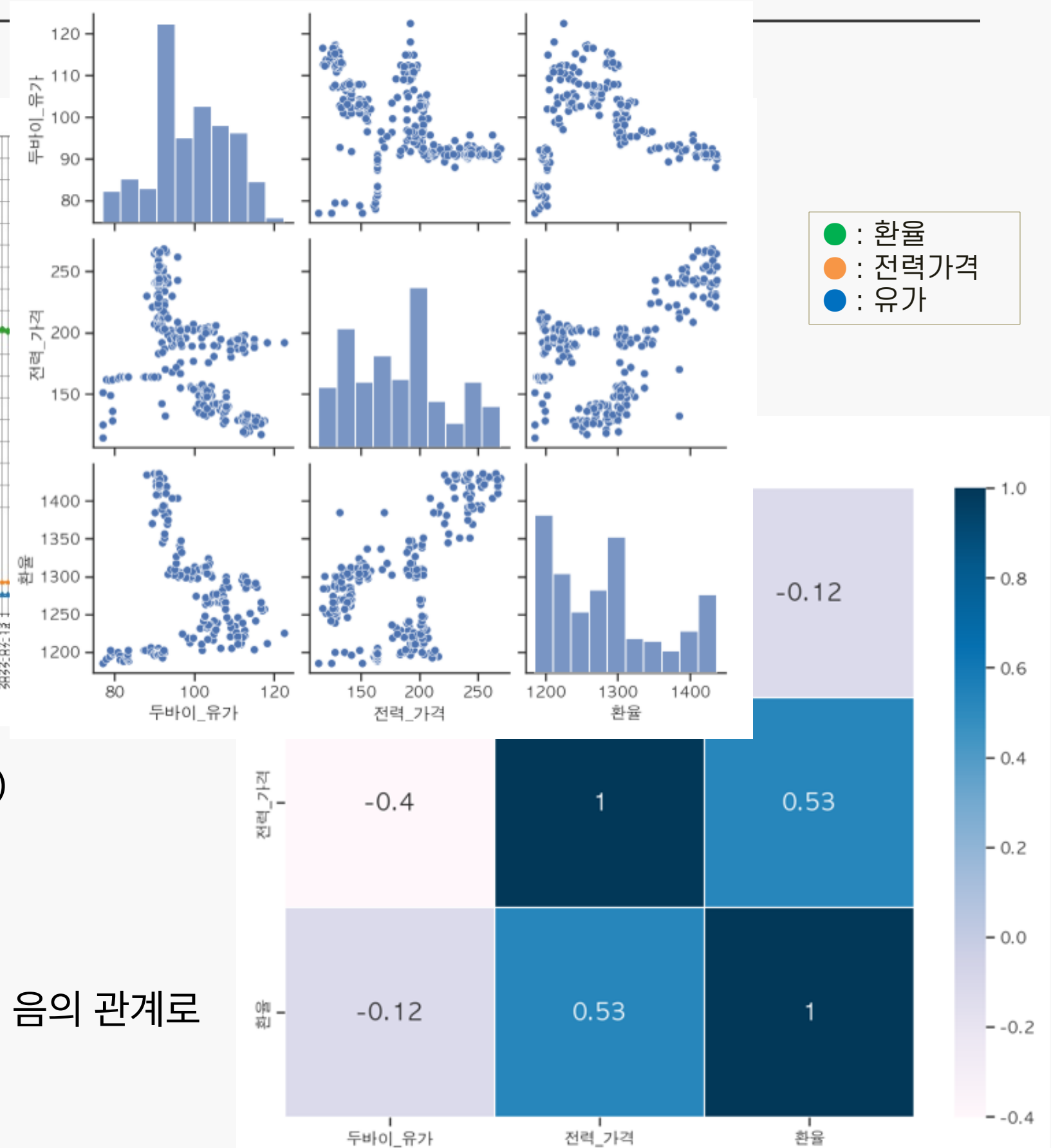
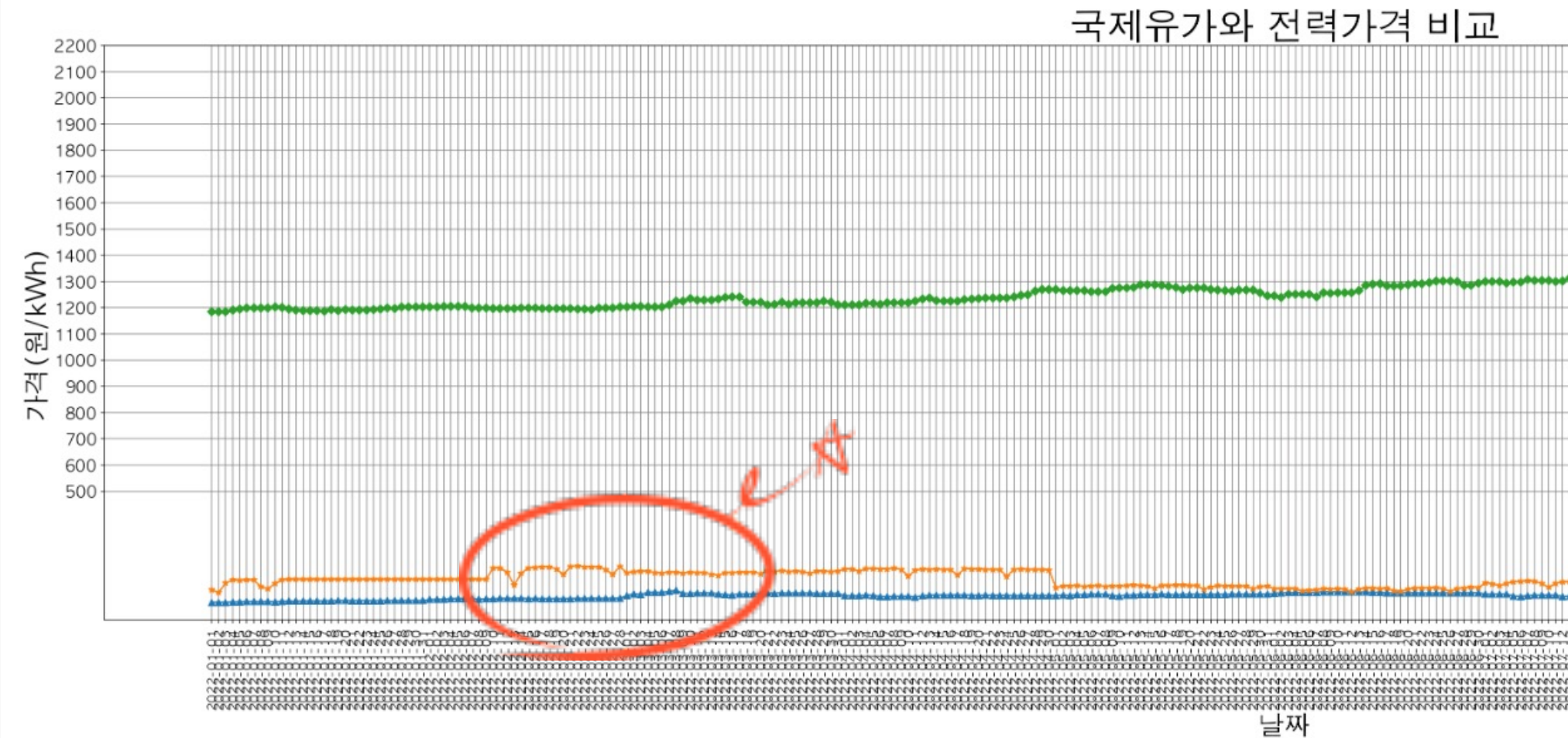
유가 전력 환율 데이터 합치기



## 최종 데이터

	두바이_유가	전력_가격	환율
날짜			
2022-01-01	77.03	125.0	1185.5
2022-01-02	77.03	114.0	1185.5
2022-01-03	77.03	151.0	1185.5
2022-01-04	78.03	162.0	1190.8
2022-01-05	78.81	161.0	1194.3
...	...	...	...
2022-10-29	91.11	242.0	1418.1
2022-10-30	91.11	229.0	1418.1
2022-10-31	91.13	244.0	1419.3
2022-11-01	91.22	255.0	1422.2
2022-11-02	91.22	255.0	1421.7

## 데이터 마이닝 시스템 구축 | 상관관계 분석



- 전력 가격 / 환율 가격은 대체적으로 양의 관계가 보인다(상관계수로 설명)
- 두바이\_유가 / 전력\_가격 은 대체적으로 음의 관계가 보인다
- 전력가격-환율:0.53, 유가-전력 : -0.4 => 다소 높음
- 전력 가격과 원유 가격이 양의관계로 올라가다, **우크라이나 전쟁**으로 인해 음의 관계로 바뀌었다.



## 데이터 마이닝 시스템 구축 | 웨카 사용 함수

### 선형회귀 분석

- 지도학습
- 입력과 결과값을 이용한 학습
- 다수의 설명 변수들과 하나의 종속 변수  
사이 관계 → 선형 방정식 모델링

+

### IBk

- Weka 의 대표적인 k-NN 알고리즘
- 데이터를 가장 가까운 속성에 따라 분류  
→ 레이블링 (k개의 가까운 이웃 속성)
- 게으른 학습 / 거리기반 분류 알고리즘



### 랜덤 포레스트

- 예측 알고리즘을 조합해 예측 성능 향상
- 정확성, 단순성, 유연성 높음
- 의사 결정 트리 숲 형성 → 데이터 병합  
→ 정확한 예측

+

### 배깅

- 샘플을 여러번 뽑아 각 모델 학습  
→ 결과 집계
- 동일한 크기의 다수 훈련 데이터 세트  
→ 중복 데이터 허용

## 데이터 마이닝 시스템 구축 | 웨카

### LinearRegression

=== Cross-validation ===

=== Summary ===

Correlation coefficient

Mean absolute error

Root mean squared error

Relative absolute error

Root relative squared error

Total Number of Instances

0.6214

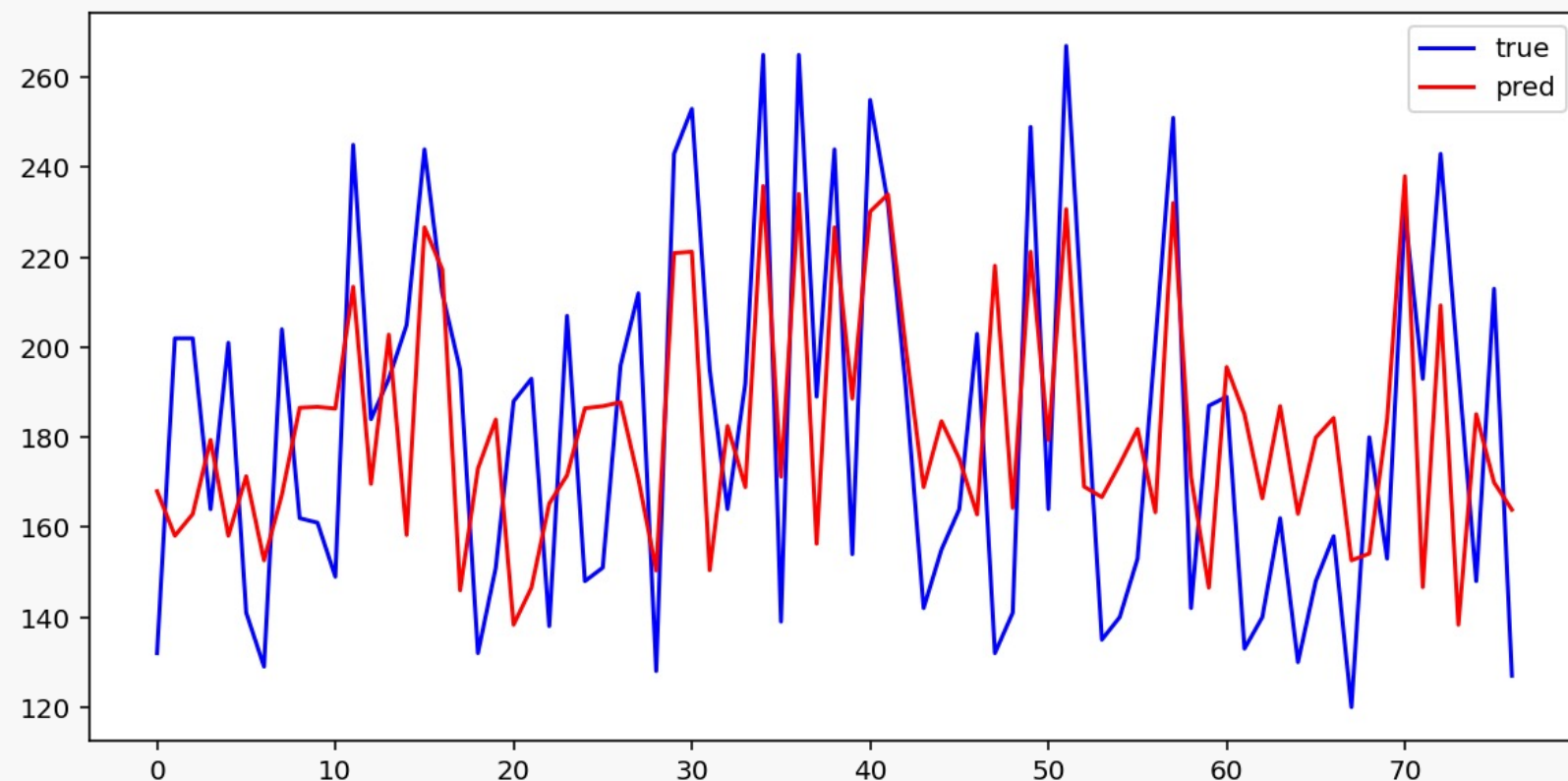
28.1355

31.5256

81.892 %

78.2267 %

306



### IBk

=== Classifier model (full training set) ===

IB1 instance-based classifier

using 3 similarity-weighted nearest neighbour(s) for classification

Time taken to build model: 0 seconds

=== Cross-validation ===

=== Summary ===

Correlation coefficient

Mean absolute error

Root mean squared error

Relative absolute error

Root relative squared error

Total Number of Instances

0.9187

10.6522

15.9313

31.0046 %

39.5313 %

306

[k-최근접 이웃 알고리즘]

KNN : 1 → 3

distanceWeighting : No → Weight by 1

windowSize : 0 → 1



## 데이터 마이닝 시스템 구축 | 웨카

### RandomForest

=== Classifier model (full training set) ===

RandomForest

Bagging with 150 iterations and base learner

weka.classifiers.trees.RandomTree -K 0 -M 1.0 -V 0.001 -S 1 -do-not-check-capabilities

Time taken to build model: 0.08 seconds

=== Cross-validation ===

=== Summary ===

Correlation coefficient	0.9194
Mean absolute error	10.6197
Root mean squared error	15.8914
Relative absolute error	30.9101 %
Root relative squared error	39.4324 %
Total Number of Instances	306

[랜덤 포레스트]

calcOutOfBag : True → False

numIteration : 100 → 150

Seed : 1 → 4

### Bagging

\*\*\* Out-of-bag estimates \*\*\*

Correlation coefficient	0.9074
Mean absolute error	10.9467
Root mean squared error	16.9899
Relative absolute error	31.7706 %
Root relative squared error	42.0689 %
Total Number of Instances	303

Time taken to build model: 0.64 seconds

=== Cross-validation ===

=== Summary ===

Correlation coefficient	0.9205
Mean absolute error	10.701
Root mean squared error	15.7493
Relative absolute error	31.1467 %
Root relative squared error	39.0798 %
Total Number of Instances	306

[배깅]

calcOutOfBag : True → False

Classifier : RandomForest

numDecimalPlaces : 2 → 7

### 선형 회귀 | 릿지 | 라쏘

- 선형 회귀 : 단 하나의 특징을 가지고 라벨값 혹은 타겟 예측
- 라쏘 : 제약 조건 통해 일반화된 모형 찾음
- 릿지 : 가중치 0에 가까워질 뿐 (더 좋음)

+

### 그리드 탐색

- 격자 탐색
- 모델에 가장 적합한 하이퍼 파라미터 찾기

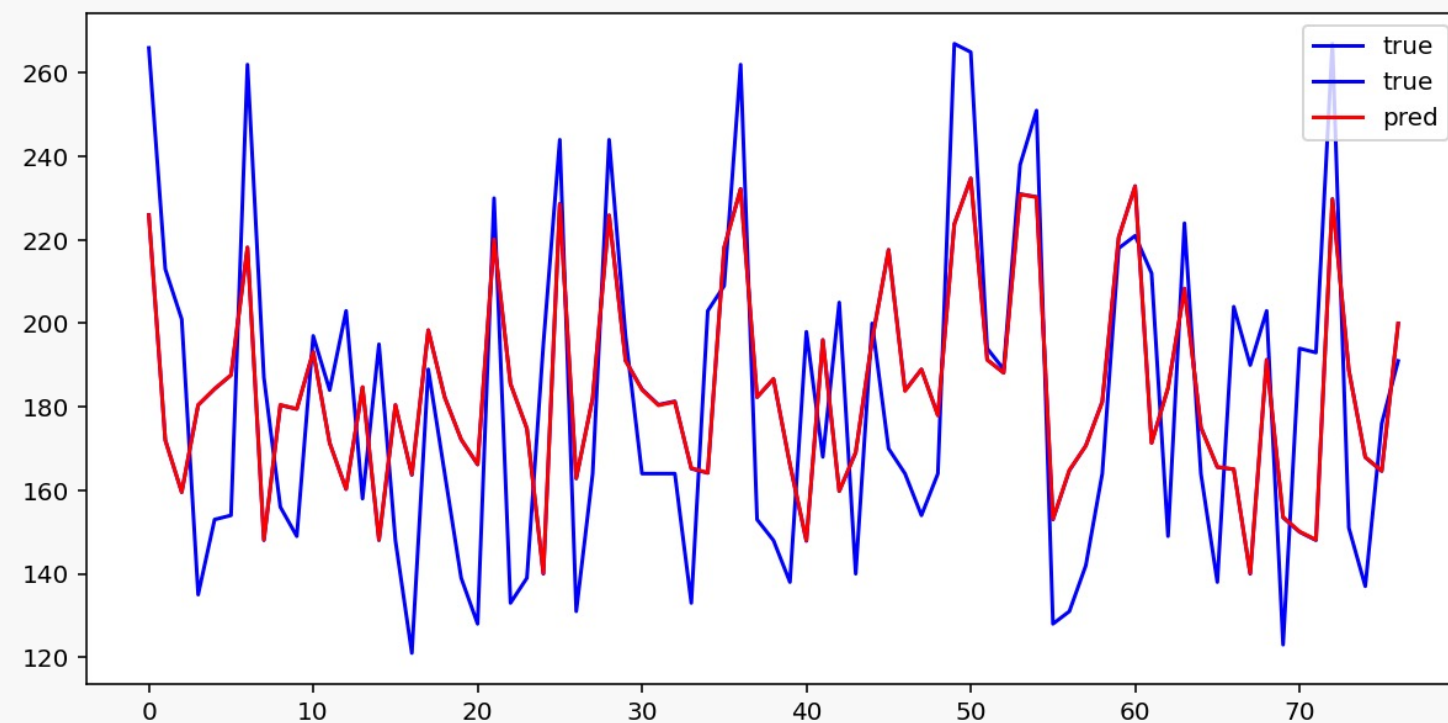
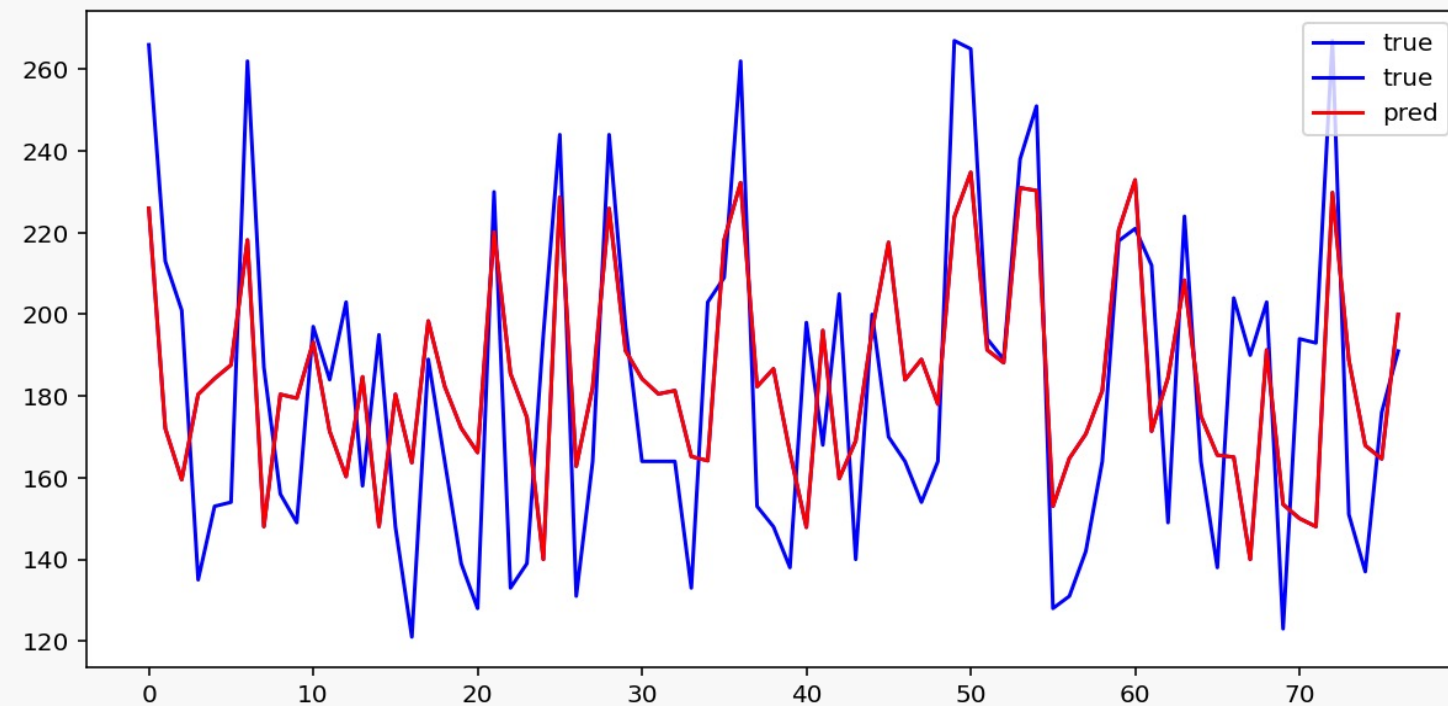


### 시계열 예측

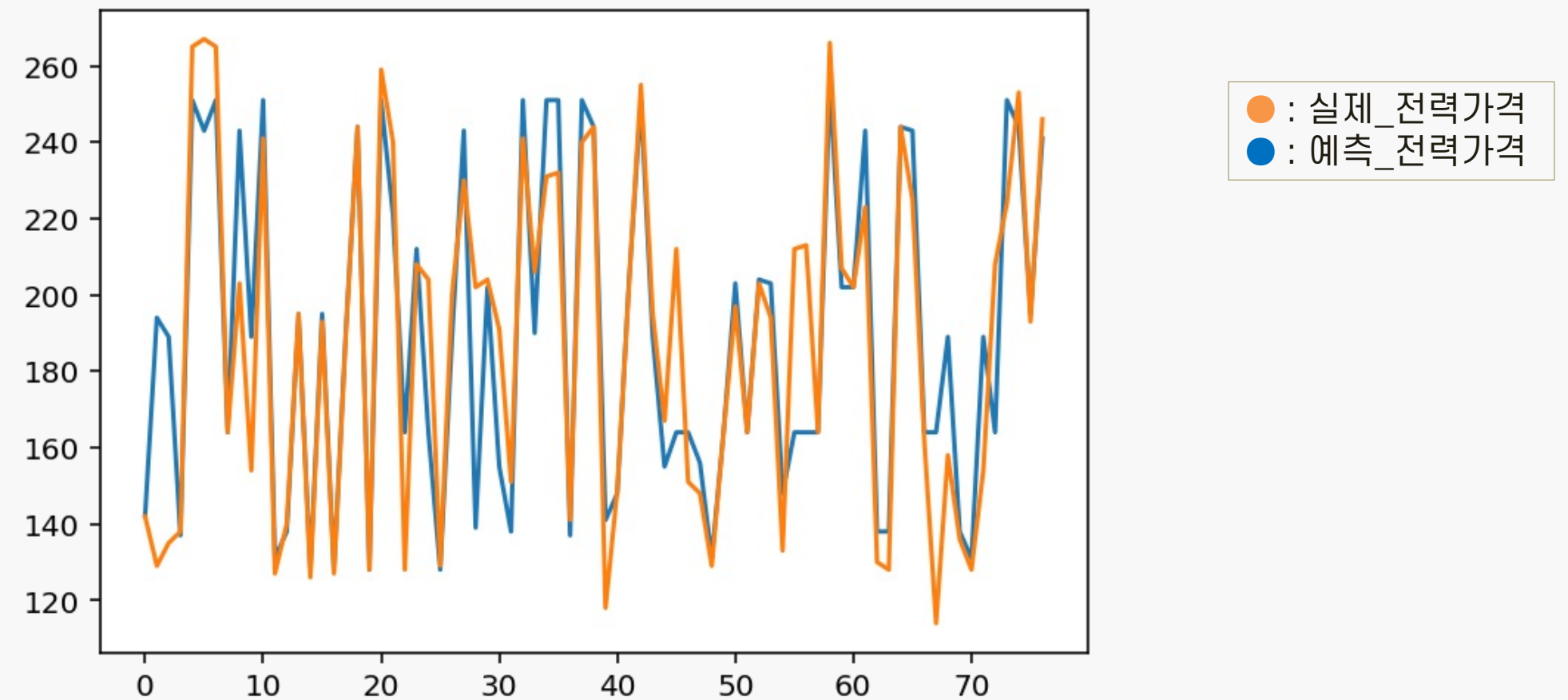
- 예측 알고리즘을 조합해 예측 성능 향상
- 정확성, 단순성, 유연성 높음
- 의사 결정 트리 숲 형성 → 데이터 병합 → 정확한 예측

# 데이터 마이닝 시스템 구축 | 파이썬

## Ridge | Lasso



## GridSearchCV



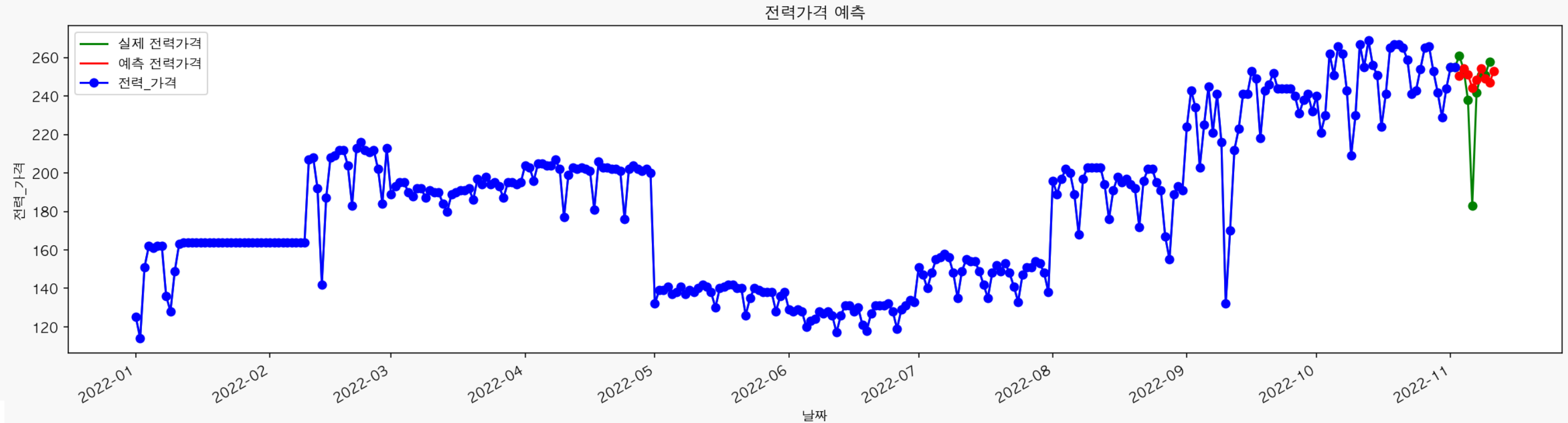
### [그리드 탐색]

: 각각의 파라미터로 모델을 여러개 만들어 최적의 파라미터를 찾아주는 함수

1. 최적의 파라미터를 찾아주고,
2. 교차검증도 해준다!

# 데이터 마이닝 시스템 구축 | 파이썬

## ARIMA



	실제 전력가격	예측 전력가격
2022-11-02	255.0	NaN
2022-11-03	261.0	250.672936
2022-11-04	254.0	254.513769
2022-11-07	242.0	248.585941
2022-11-08	251.0	254.202898
2022-11-09	251.0	249.264228
2022-11-10	258.0	247.013353

### [시계열 분석]

: 불규칙적인 시계열 데이터에 규칙성을 부여하는 방식  
시간에 대해 순차적 관측되는 데이터 집합

1. AR 과 MA를 합친 모델 + '차분' → 좀 더 나은 예측
2. 현재 시점의 상태를 파악하는 과거 시점의 자신과 추세까지 전부 반영하겠다는 의미

## 데이터 분석 결과 평가

### 원유가격과 우리나라 전력시장 가격 간 연계성

#### ✓ 전력 산업은 대규모 장치 산업

진입 장벽이 높은 독과점 형태를 나타냄  
발전회사들이 규가가 상승할 때 혹은 하락할 때  
→ 자신들에게 유리한 방향으로 공급량 조정할  
유인 존재

#### ✓ 조정과정 대칭적

국제유가 변동에 따라 인위적 전력가격 조정을  
통한 차익 거래가 어려울 것으로 예상

### 참고 문헌 및 분석툴

#### 논문

- 박민혁·문양택·박중구. 2014. “전력계통한계가격(SMP)과 기저발전비율, LNG 도입가격, 환율 간 인과관계 분석”, 「조명전기설비학회논문지」, 28(7), 97-105.
- 박진호·김철현·김형건. 2016. 「국제 가스시장 구조적 변화에 따른 우리나라 LNG 도입 대응방안 연구」, 에너지경제연구원 기본연구보고서, 16-16.
- 박해선·이상직. 2015. “주요 원유 현물가격간의 비선형 동적조정에 관한 연구”, 「자원 환경경제연구」, 24(4), 657-677.
- 정수관. 2018. “구조변화를 고려한 국제원유시장의 통합 연구”, 「에너지경제연구」, 17(1), 153-177.
- 조홍중·한원희. 2015. “구조변화를 고려한 한국의 LNG 가격 추정”, 「자원환경경제연구」,

#### 분석툴



### 활용 데이터

분석 데이터	기간	제공기관
두바이유 가격 데이터	2022	인베스팅.com
전력 일별 가격 데이터	2022	GEMS
기간별 환율 데이터	2022	우리은행 외환센터



## 웨카 장단점

### 장점

Knowledge folw 기능으로  
기계학습을 설계/시뮬레이션 할 수  
있다.

### 기계학습을 GUI로 구현/검증 가능

- 클릭만으로 모델평가까지 가능하다.
- 1회성이나 검증용

### java 코딩이 가능

- GUI, KnowledgeFlow와 같은 과정을 코딩으로 구현 가능

### 단점

### 메모리 부족을 피할 수 없음

- java로 구현 → 자연어 처리와 같은 대용량 데이터를 GPU까지 사용해야할 경우 여지 없이 메모리 부족으로 kf나 GUI가 꺼진다.

### 무료 소프트웨어의 한계가 보임

- kf 경우 시뮬레이션 결과를 json 형태로 파일저장이 되는데 가끔 다시 불러오기가 안 된다.

### 한글 인코딩 처리 문제

- 영어를 기반으로 제작된 weka는 한글인식이 안 된다.

---

마무리

---

감사합니다 :)