

# 하이트 진로 주가 예측

함쳐서 8병 먹조: 강지원 노하늘 박희연 이윤지



# INDEX

- 1.Topic 주제 선정
- 2.Data 수집 데이터
- 3.Model 적용 모델
4. Result 결과 및 느낀점

# 합쳐서 8명 먹조

**강지원**

데이터 수집  
데이터 전처리  
중간발표  
기사 크롤링  
워드클라우드  
감성분석  
DNN

조장  
**노하늘**

데이터 전처리  
중간 ppt  
VAR  
기사 크롤링  
Prophet  
LSTM  
GRU

**박희연**

데이터 전처리  
일지 작성  
결정트리  
랜덤포레스트  
XGB  
최종 발표

**이운지**

데이터 수집  
데이터 전처리  
결정트리  
랜덤포레스트  
XGB  
최종 ppt

## 주류 업계 선정 이유: 코로나 이후 변화

 더팩트 PICK | 1일 전 | 네이버뉴스

**거리두기 해제** 후 첫 주말, 활기 되찾은 을지로 노가리 골목 [TF사...

정부가 지난 15일 발표한 거리두기 해제 방침에 따라 사적 모임 인원 제한과 영업 시간 제한 등 사회적 거리두기의 규제 조치가 모두 해제됐다. 다만 영화관, 종교...



 세계일보 PICK | A11면 1단 | 12시간 전 | 네이버뉴스

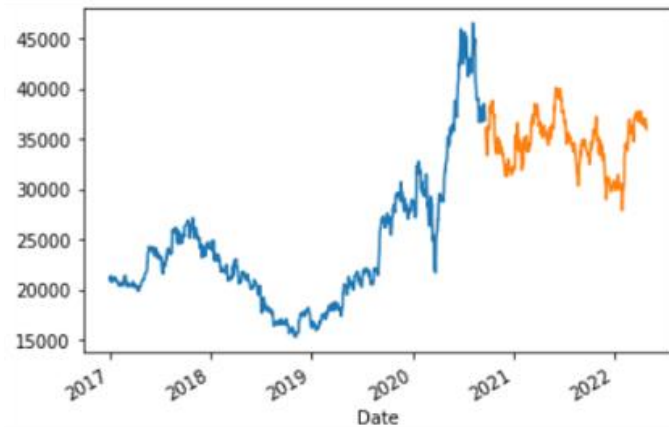
**거리두기 해제** 첫 주말...도심은 집회로 '북새통' 술집·식당은 '만석'

연합뉴스 ◆도심 집회로 '북새통'... 변화가 술집·식당은 '만석' 사회적 거리두기가 해제된 뒤 첫... 연합뉴스 강남·홍대·을지로 등 변화가는 쇼핑이나 모임을 나온 시...

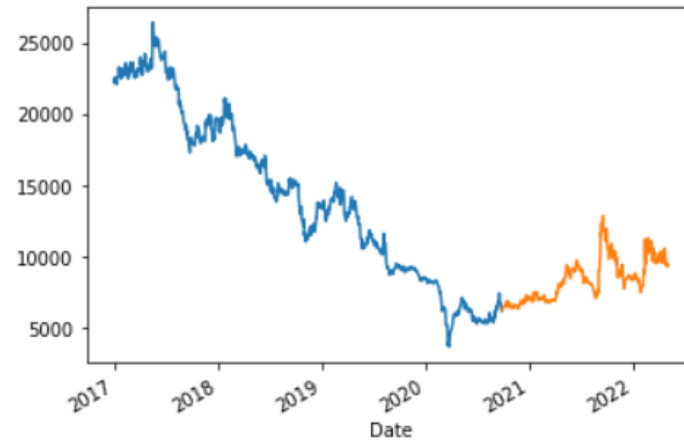


## 주류 업계 주가

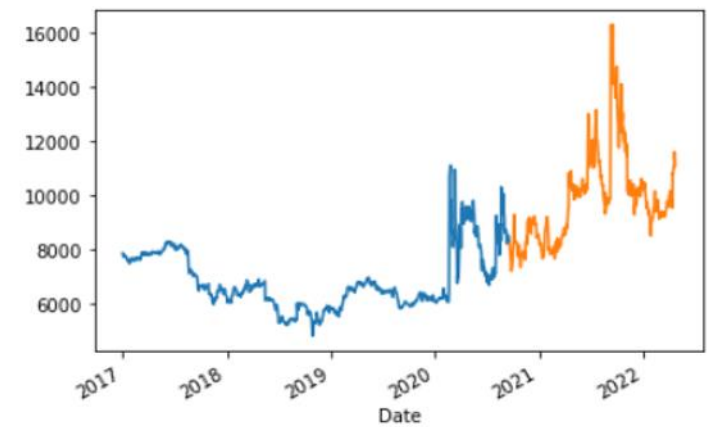
(2017년 1월 ~ 2022년 4월)



하이트진로



무학



MH에탄올

## 주류 관련 기업

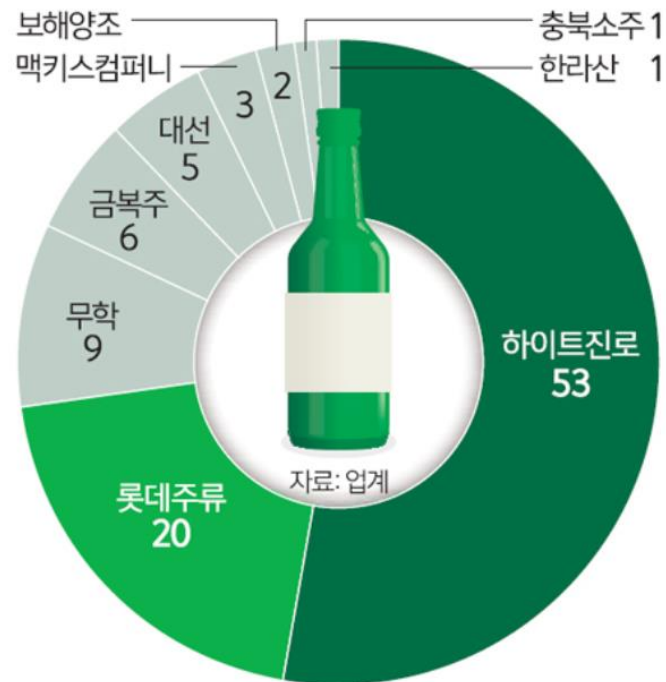


## 주류 관련 기업



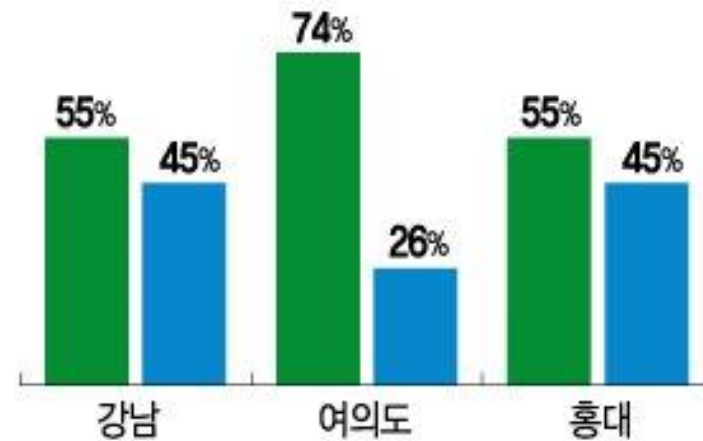
# Why 하이트 진로?

국내 소주사 전국시장 점유율 (단위: %, 2월 기준)



강남·여의도·홍대 지역별 맥주 점유율

■ 테라(하이트진로) ■ 카스(오비맥주)



\*자료: 메리츠증권증권 리서치센터  
그래픽: 김지영 디자인가자





# 워드클라우드





소주의 원조  
**진-로**



## 하이트 진로?

맥주- TERRA, Hite, 드라이피니시 d

소주- 참이슬, 진로



## 수집 데이터

(기간: 2017.01.02 ~ 2022.04.21)

데이터	단위	출처
하이트진로 주가	원	
밀 수급동향	달러	KREI 해외곡물 시장정보
맥류 수입금액 지수	2015년 100기준	한국은행 국제수지통계
WTI유 선물가격	달러	Investing.com
하이트진로 외국인 지분율	%	KRX 정보데이터시스템
하이트진로 월별 광고비	천 원	한국광고총연합회 광고정보센터
Kospi지수		
무학 주가	원	
미국,일본 환율	원	
뉴스제목		

## 수집 데이터

(기간: 2017.01.02~ 2022.04.21)

데이터	단위	출처
하이트진로 주가	원	
밀 수급동향	달러	KREI 해외곡물 시장정보
맥류 수입금액 지수	2015년 100기준	한국은행 국제수지통계
WTI유 선물가격	달러	Investing.com
하이트진로 외국인 지분율	%	KRX 정보데이터시스템
하이트진로 월별 광고비	천 원	한국광고총연합회 광고정보센터
Kospi지수		
무학 주가	원	
미국,일본 환율	원	
뉴스제목		

## 수집 데이터

(기간: 2017.01.02 ~ 2022.04.21)

데이터	단위	출처
하이트진로 주가	원	
밀 수급동향	달러	KREI 해외곡물 시장정보
맥류 수입금액 지수	2015년 100기준	한국은행 국제수지통계
WTI유 선물가격	달러	Investing.com
하이트진로 외국인 지분율	%	KRX 정보데이터시스템
하이트진로 월별 광고비	천 원	한국광고총연합회 광고정보센터
Kospi지수		
무학 주가	원	
미국,일본 환율	원	
뉴스제목		

## 수집 데이터

(기간: 2017.01.02 ~ 2022.04.21)

데이터	단위	출처
하이트진로 주가	원	
밀 수급동향	달러	KREI 해외곡물 시장정보
맥류 수입금액 지수	2015년 100기준	한국은행 국제수지통계
WTI유 선물가격	달러	Investing.com
하이트진로 외국인 지분율	%	KRX 정보데이터시스템
하이트진로 월별 광고비	천 원	한국광고총연합회 광고정보센터
Kospi지수		
무학 주가	원	
미국,일본 환율	원	
뉴스제목		

## 수집 데이터

(기간: 2017.01.02 ~ 2022.04.21)

데이터	단위	출처
하이트진로 주가	원	
밀 수급동향	달러	KREI 해외곡물 시장정보
맥류 수입금액 지수	2015년 100기준	한국은행 국제수지통계
WTI유 선물가격	달러	Investing.com
하이트진로 외국인 지분율	%	KRX 정보데이터시스템
하이트진로 월별 광고비	천 원	한국광고총연합회 광고정보센터
Kospi지수		
무학 주가	원	
미국,일본 환율	원	
뉴스제목		



## 수집 데이터

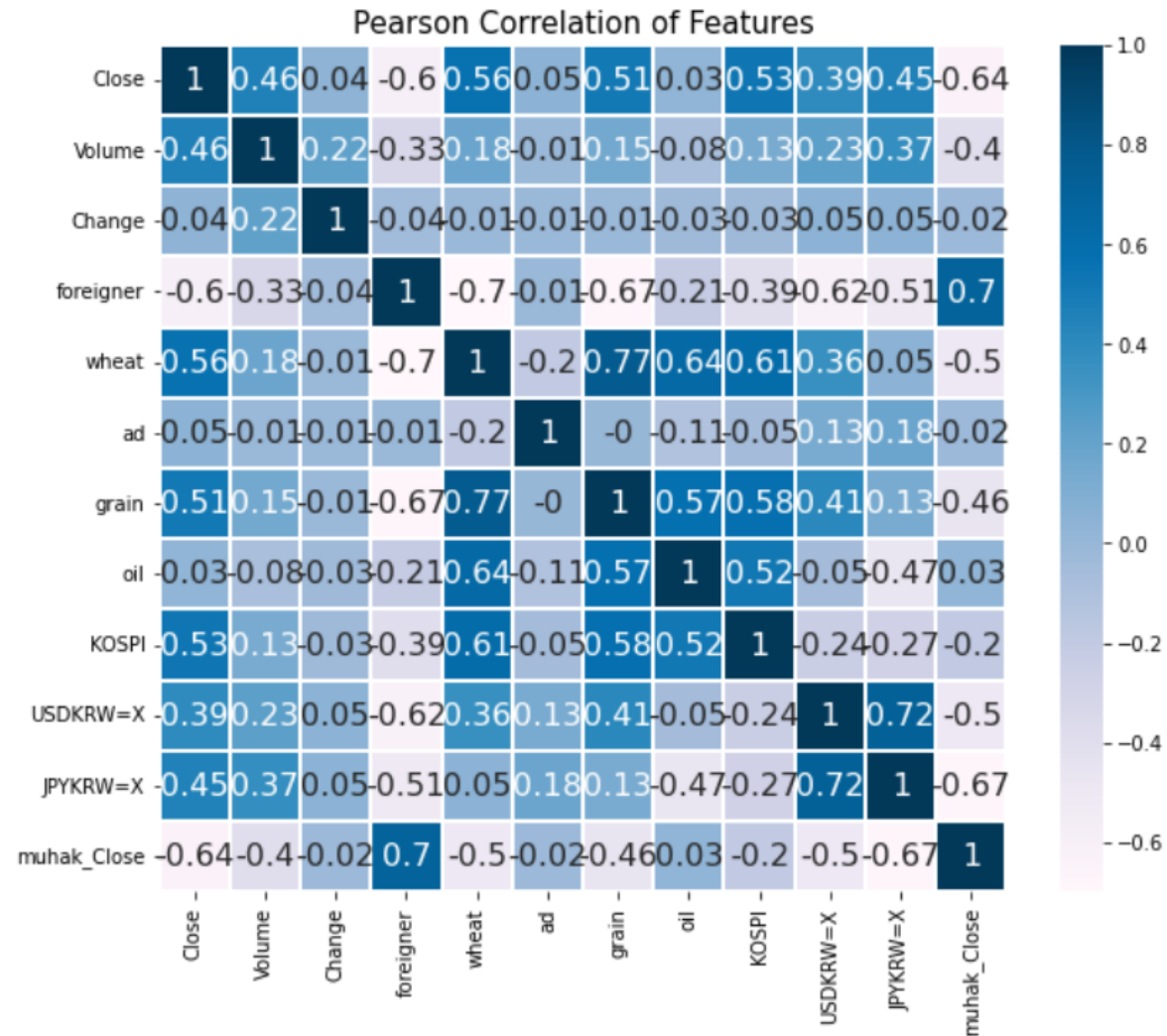
(기간: 2017.01.02~ 2022.04.21)

데이터	단위	출처
하이트진로 주가	원	
밀 수급동향	달러	KREI 해외곡물 시장정보
맥류 수입금액 지수	2015년 100기준	한국은행 국제수지통계
WTI유 선물가격	달러	Investing.com
하이트진로 외국인 지분율	%	KRX 정보데이터시스템
하이트진로 월별 광고비	천 원	한국광고총연합회 광고정보센터
Kospi지수		
무학 주가	원	
미국,일본 환율	원	
뉴스제목		

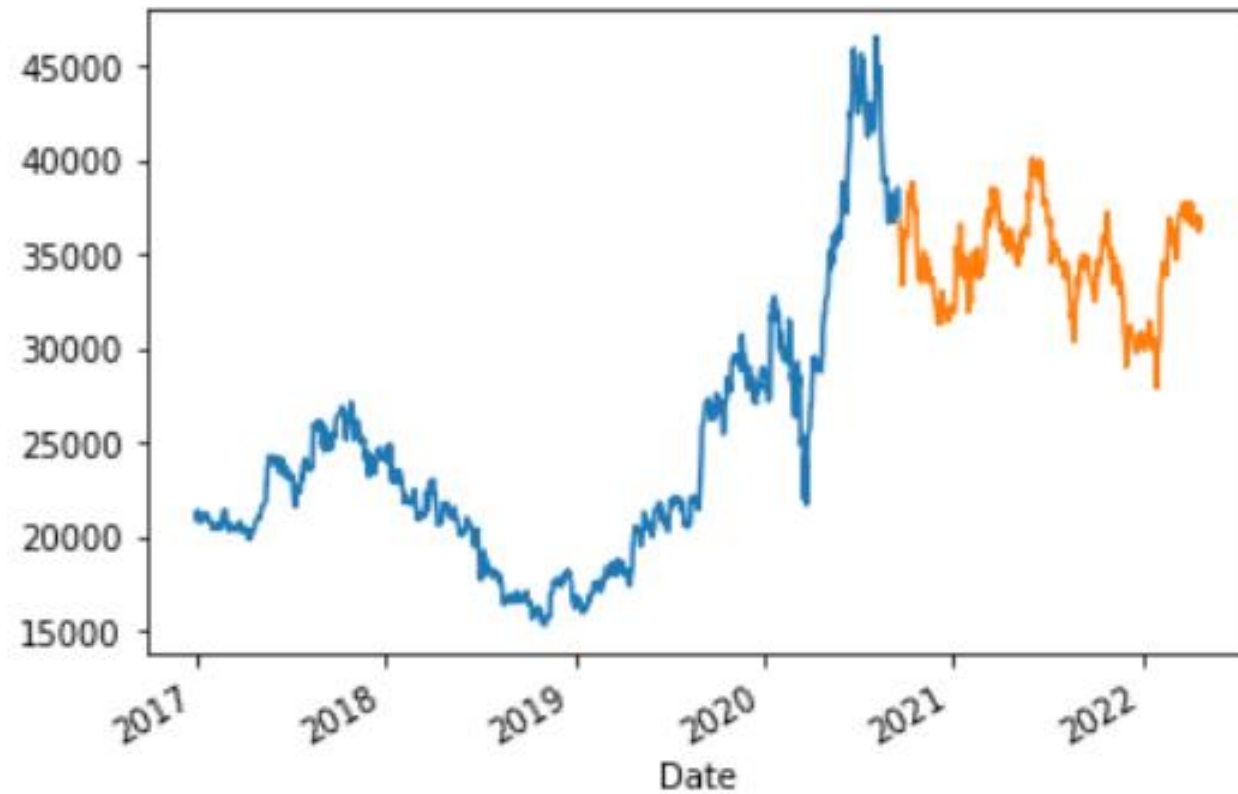
## 기초통계량

	Close	Volume	Change	foreigner	wheat	ad	grain	oil	KOSPI	USDKRW=X	JPYKRW=X	muhak_Close
count	1303.00	1303.00	1303.00	1303.00	1303.00	1303.00	1303.00	1303.00	1303.00	1303.00	1303.00	1303.00
mean	27303.76	367973.39	0.00	11.39	205.66	3098842.03	103.72	58.36	2435.14	1147.11	10.42	12480.34
std	7542.88	290121.83	0.02	4.30	51.09	712604.76	25.31	15.58	393.26	43.32	0.47	5683.50
min	15300.00	62630.00	-0.11	4.04	146.97	1525627.00	70.49	-37.63	1457.64	1053.73	9.29	3660.00
25%	20825.00	184274.50	-0.01	8.34	171.08	2977655.00	88.06	49.58	2129.70	1117.94	10.08	8070.00
50%	25950.00	277512.00	0.00	10.60	190.05	3098842.00	93.38	57.23	2348.26	1136.77	10.33	10600.00
75%	34300.00	461737.50	0.01	13.65	226.66	3098842.00	112.05	66.77	2693.16	1183.89	10.75	17000.00
max	46500.00	2779460.00	0.10	20.63	475.46	5361161.00	167.58	123.70	3305.21	1262.93	11.63	26450.00

# 피어슨 상관계수

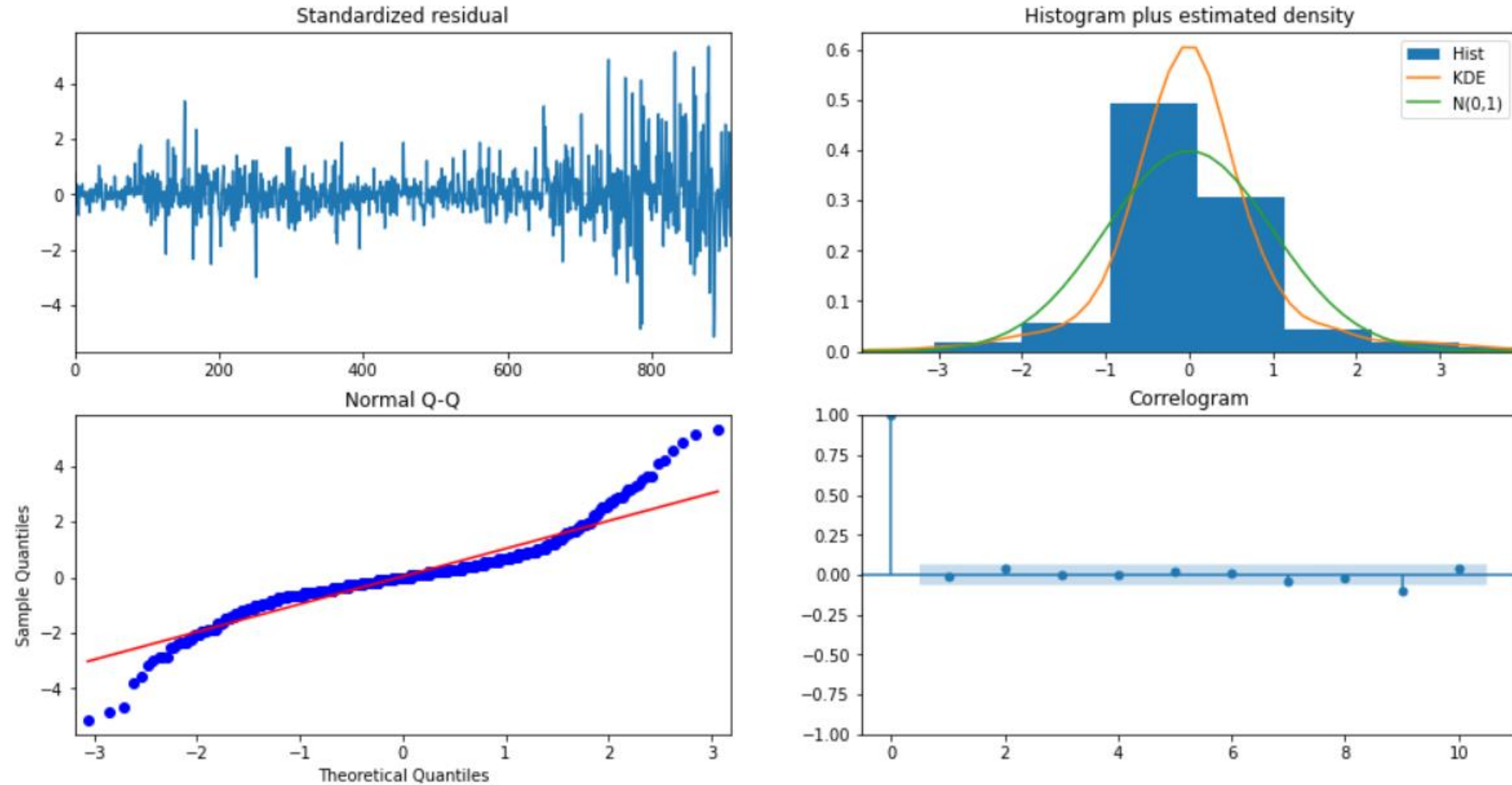


# ARIMA



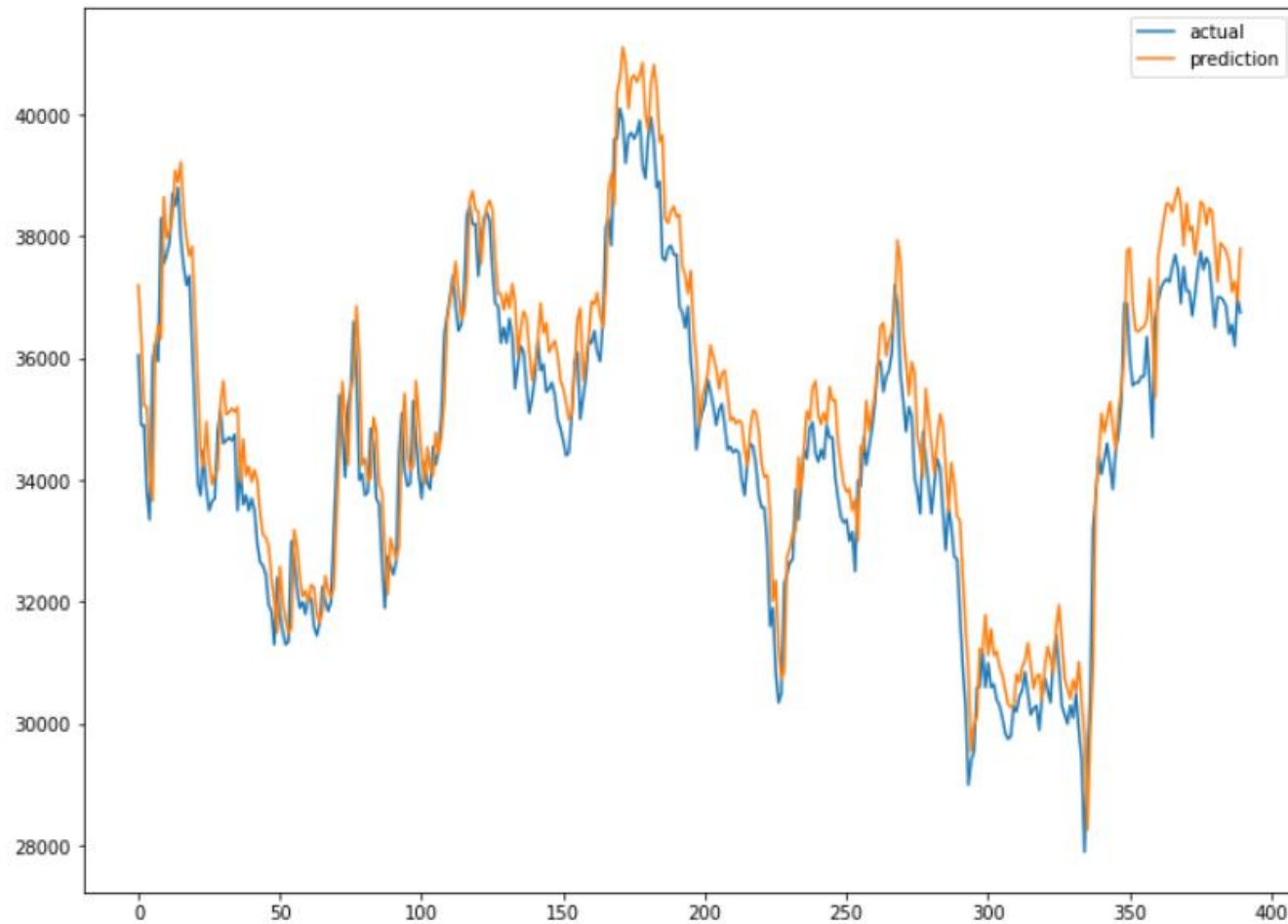
▲하이트 진로 5년 동안의 주가

# ARIMA





# LSTM

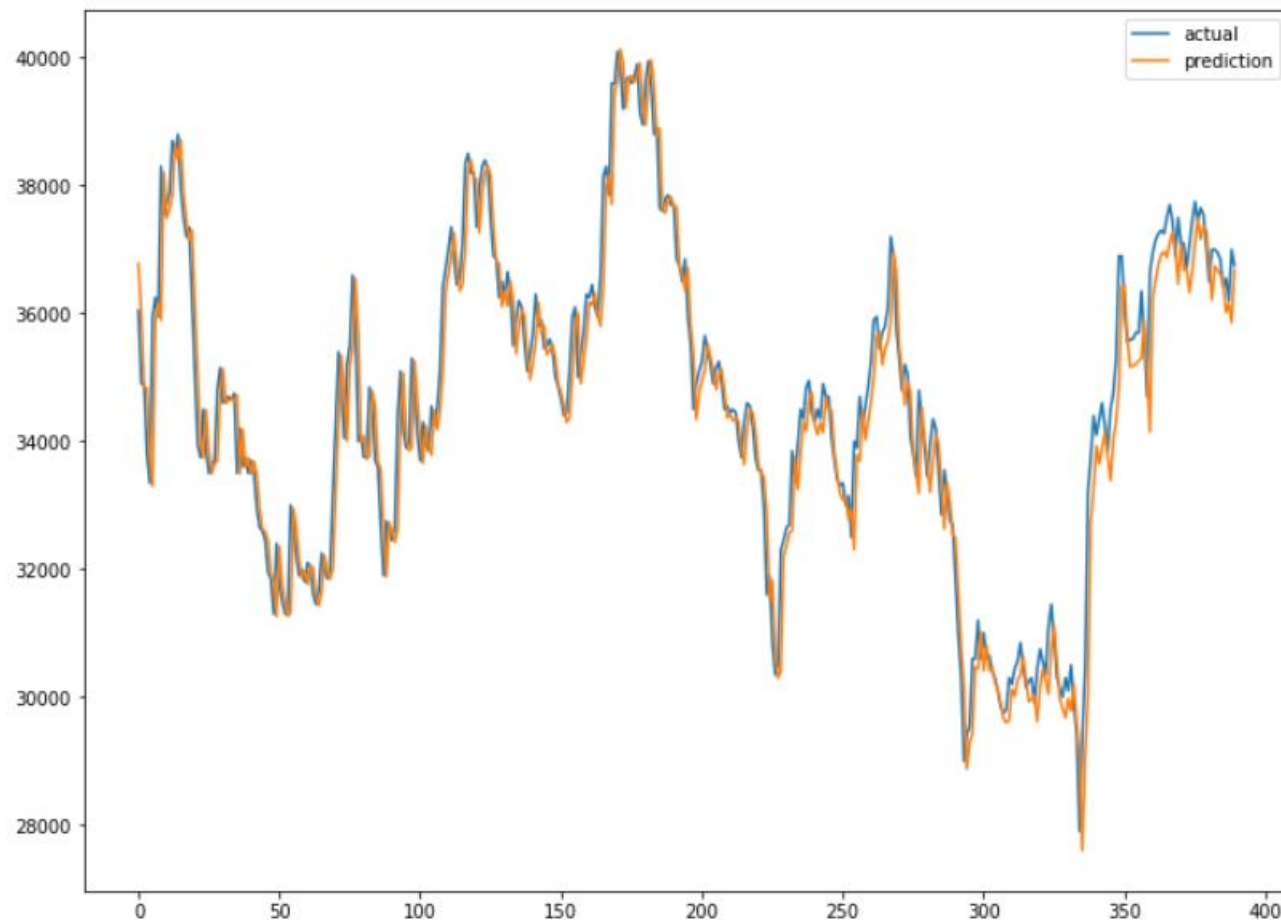


**LSTM 으로 예측한 하이트 진로 주가**





# GRU



GRU로 예측한 하이트 진로 주가

## 분류분석

**Close\_change: 전날 종가 대비 상승 시 1, 그 외 0**

	Date	Open	High	Low	Close	Volume	Change	외국인 지분율	밀	광고비	맥류 및 잡곡	원유	KOSPI	Close_change
0	2017-01-02	21100	21150	20650	20900	200092	-0.009479	11.66	149.32	2989798	79.9	54.03	2026.16	0
1	2017-01-03	20900	21300	20750	21300	154765	0.019139	11.62	149.32	2989798	79.9	52.33	2043.97	1
2	2017-01-04	21300	21300	21000	21100	126132	-0.009390	11.65	153.73	2989798	79.9	53.26	2045.64	0
3	2017-01-05	21250	21400	21100	21300	187376	0.009479	11.69	156.60	2989798	79.9	53.76	2041.95	1
4	2017-01-06	21300	21350	21200	21350	80523	0.002347	11.78	155.50	2989798	79.9	53.99	2049.12	1

# 결정트리

사용변수: 외국인 지분율, 밀, 광고비, 맥류 및 잡곡, 원유, KOSPI지수

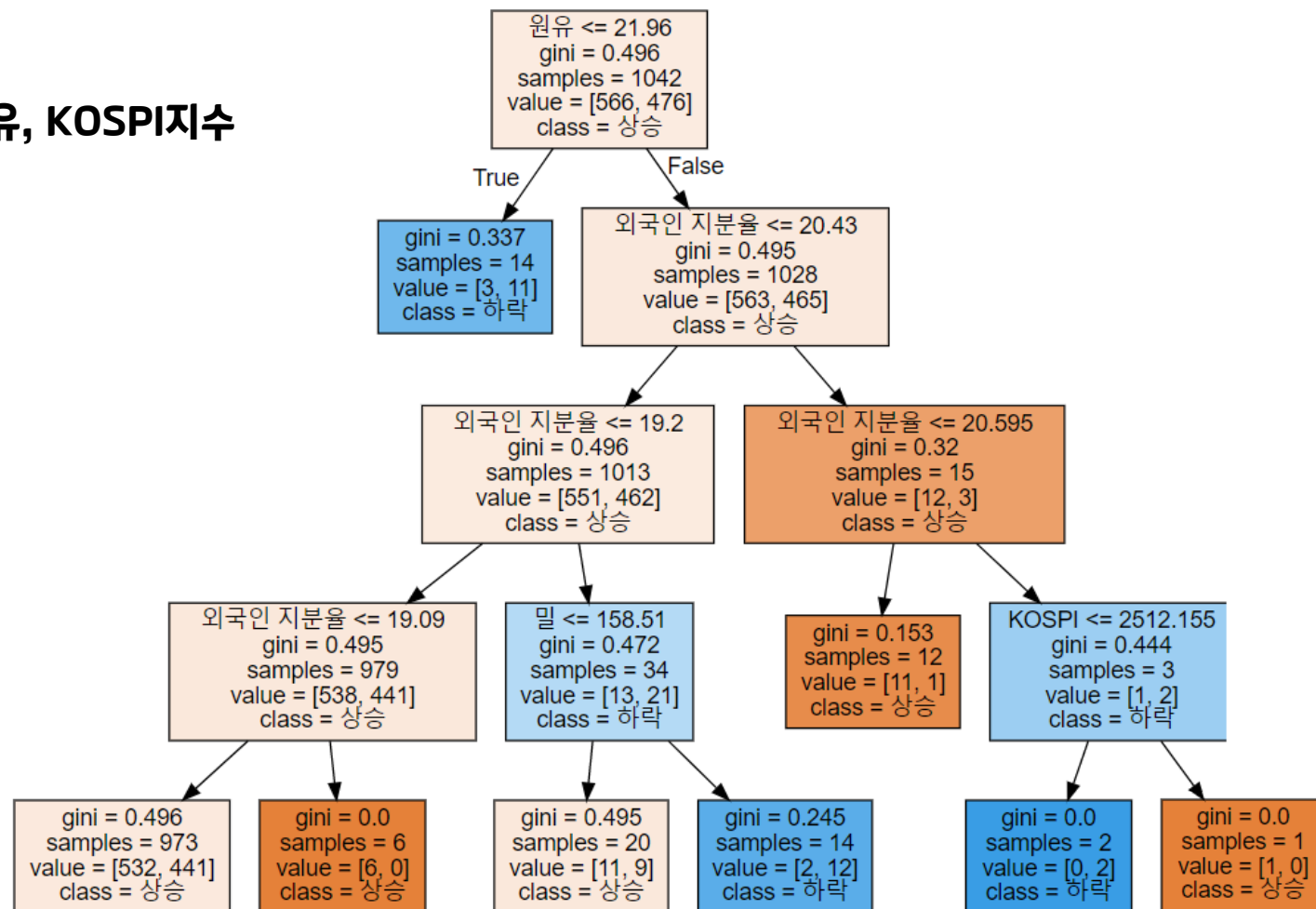
Max\_depth: 4

Max\_leaf\_nodes: 8

Min\_samples\_split: 3

훈련 데이터 정확도: 0.562

테스트 데이터 정확도: 0.513



# XGB Classifier

사용변수: 외국인 지분율, 밀, 광고비, 맥류 및 잡곡, 원유, KOSPI지수

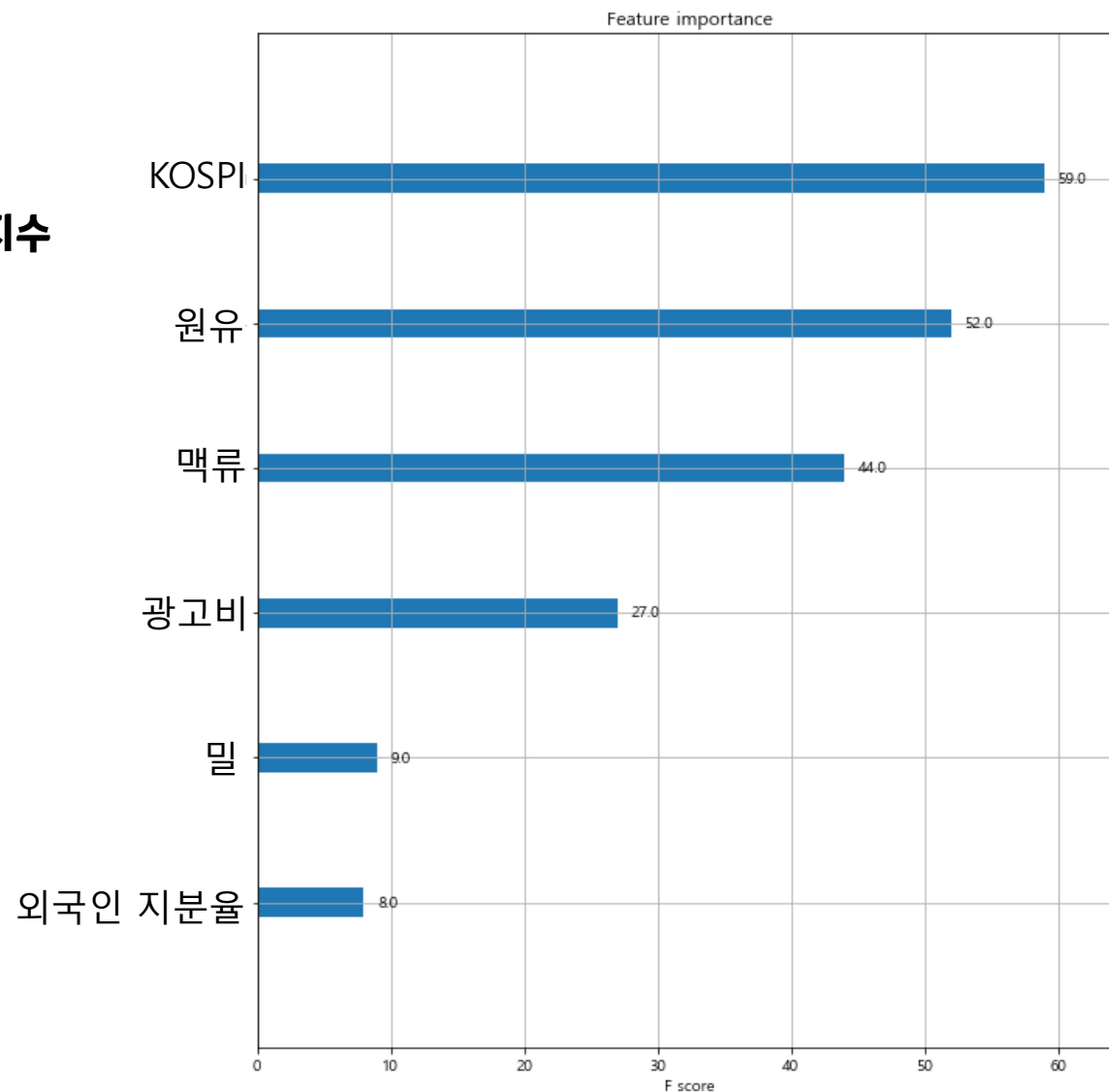
**max\_depth: 5**

**min\_child\_weight: 10**

**gamma: 1**

**정확도: 0.54    정밀도: 0.50    재현율: 0.44**

**F1: 0.47    AUC: 0.49**



## 랜덤포레스트

StandardScaler로 데이터 스케일링 후 랜덤포레스트를 실행한 결과 정확도 0.54

## 로지스틱 회귀분석

MAE 0.45, MSLE 0.21이 나오고 정확도는 0.55

# DNN

사용변수: 외국인 지분율, 밀, 광고비, 맥류 및 잡곡, 원유, KOSPI지수, 미국 환율, 일본 환율

Activation 활성화 함수: relu, sigmoid

Loss: Binary cross entropy

Epochs: 200

Batch size: 32

Loss: 0.6988      acc: 0.6411

Val\_loss: 0.7693    Val\_acc: 0.5933

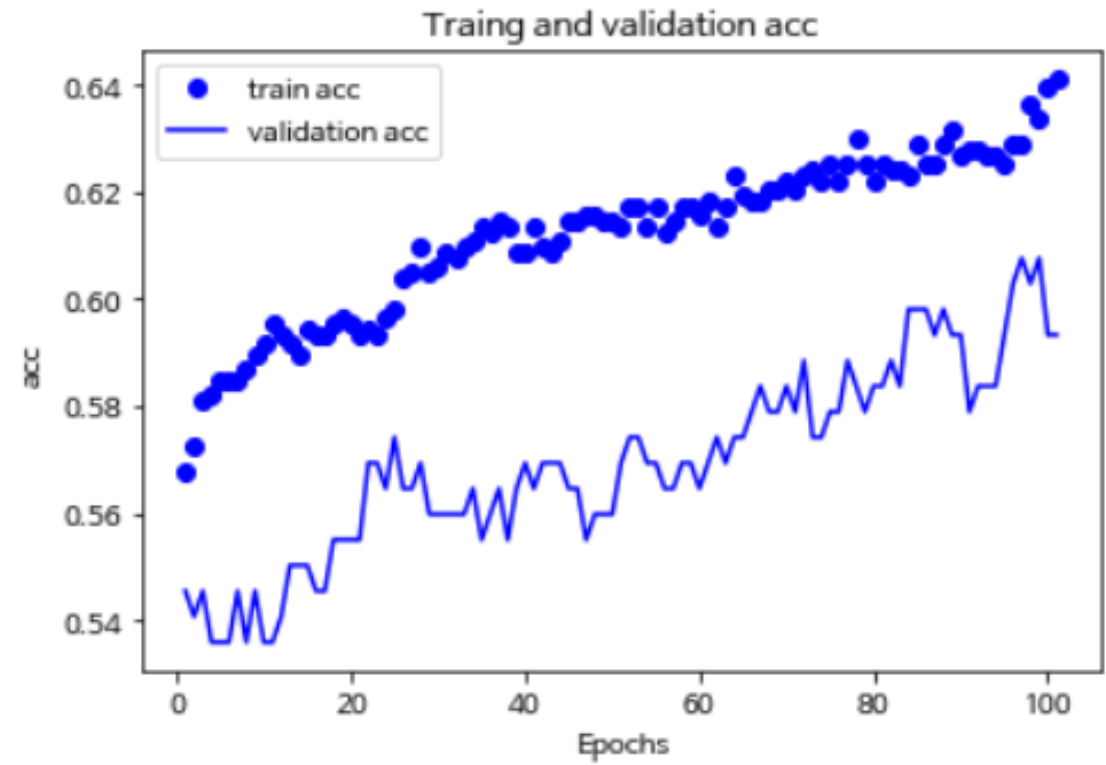
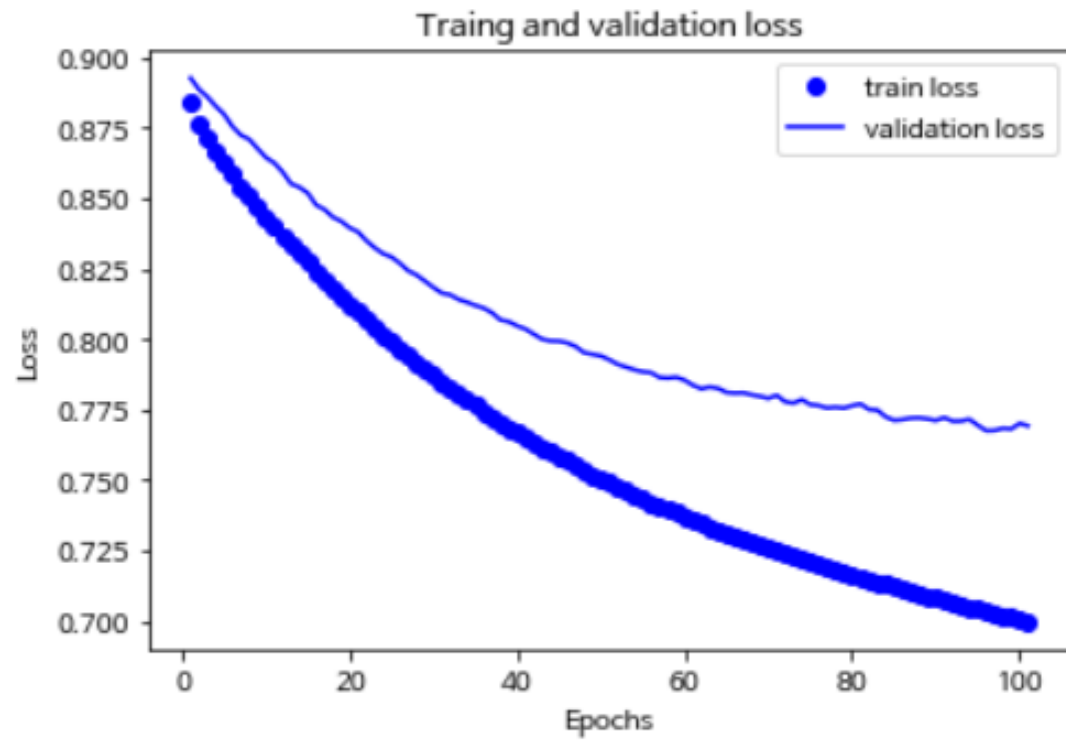
## ▼사용한 DNN 코드

```
model4 = models.Sequential()
model4.add(layers.Dense(128, activation='relu', input_shape=(max_features,),
                        kernel_regularizer=regularizers.l2(0.01)))
model4.add(layers.Dense(64, activation='relu'))
model4.add(layers.Dense(32, activation='relu'))
model4.add(layers.Dense(16, activation='relu'))
model4.add(layers.Dense(1, activation='sigmoid'))

model4.summary()

# batch_size도 성능에 영향을 미친다.
model4.compile(optimizer=RMSprop(lr=1e-4), loss='binary_crossentropy', metrics=['acc'])
history = model4.fit(X_train_scaled, y_train, epochs=200, batch_size=32,
                    validation_split=0.2, callbacks=[EarlyStopping(patience=5)])
```

# DNN



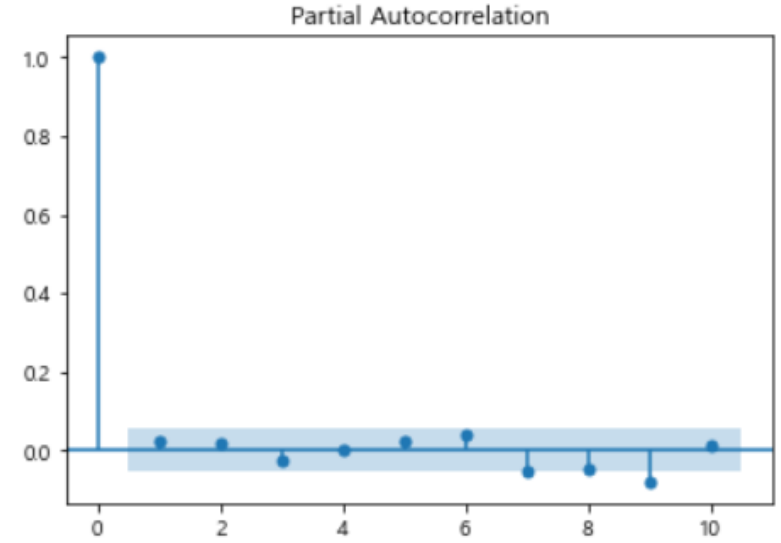
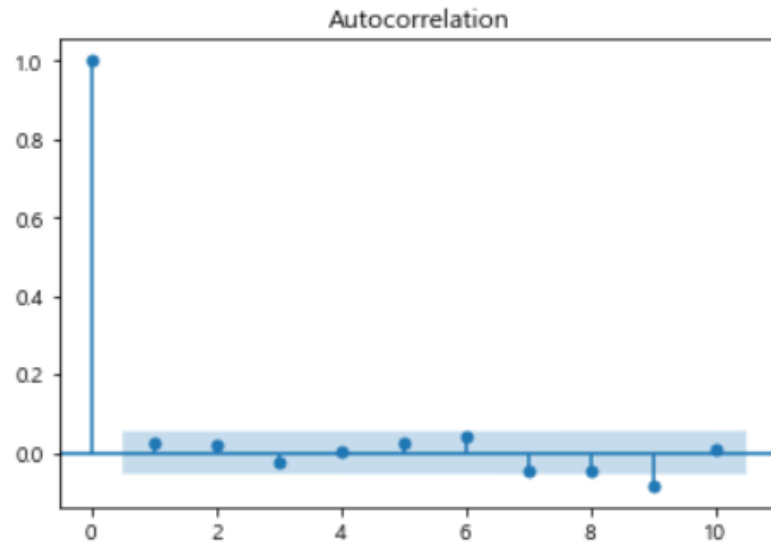
# VAR

벡터 자기회귀 모형. 일변량 ARIMA 모형과 달리 변수들 간의 상호 작용을 고려

KPSS 안정성 테스트 결과  $p\text{-value} = 0.01 < 0.05$  로 시계열이 안정적이라는 귀무가설 기각

-> 광고비를 제외한 증가, 외국인 지분율, 밀, 맥류 및 잡곡, 원유, KOSPI 1차 차분

AIC 기준으로 VAR 최적 모델 순서 2로 결정

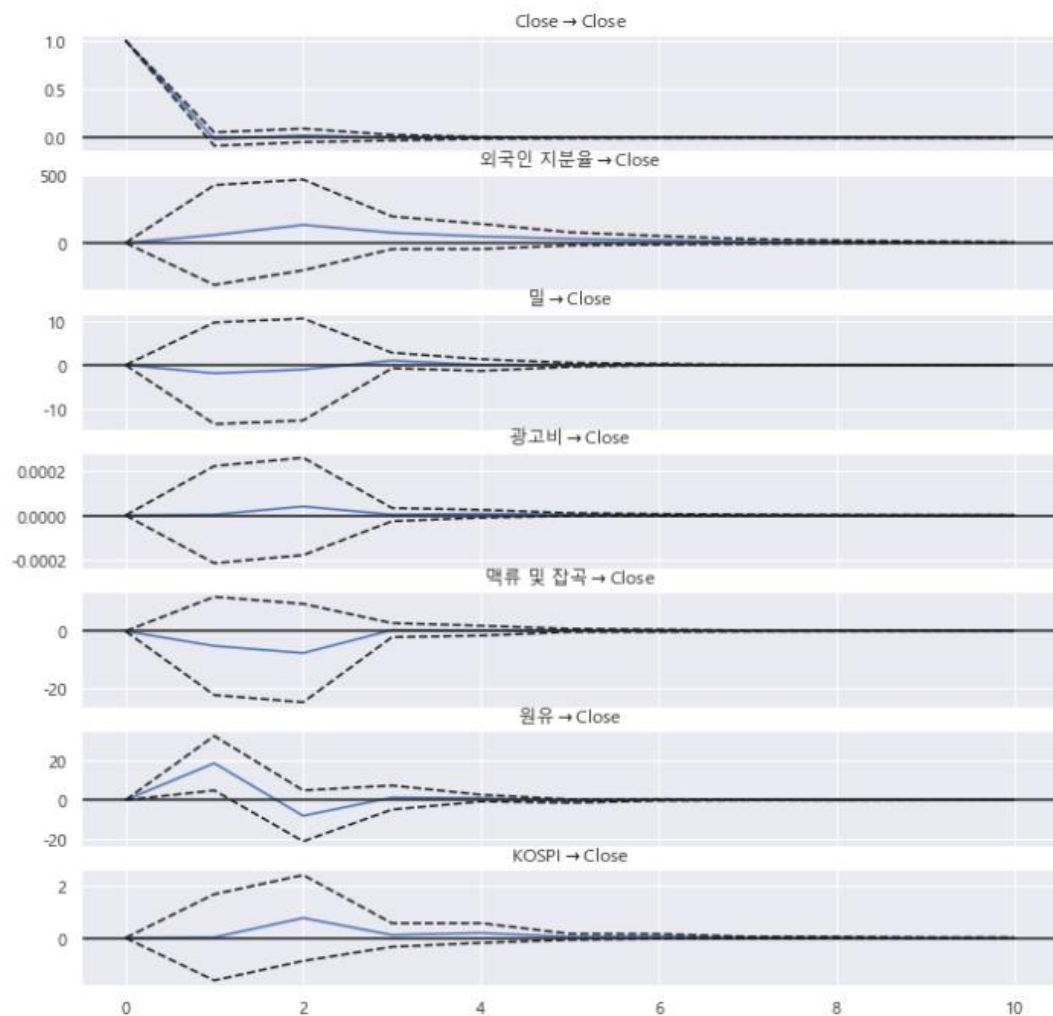




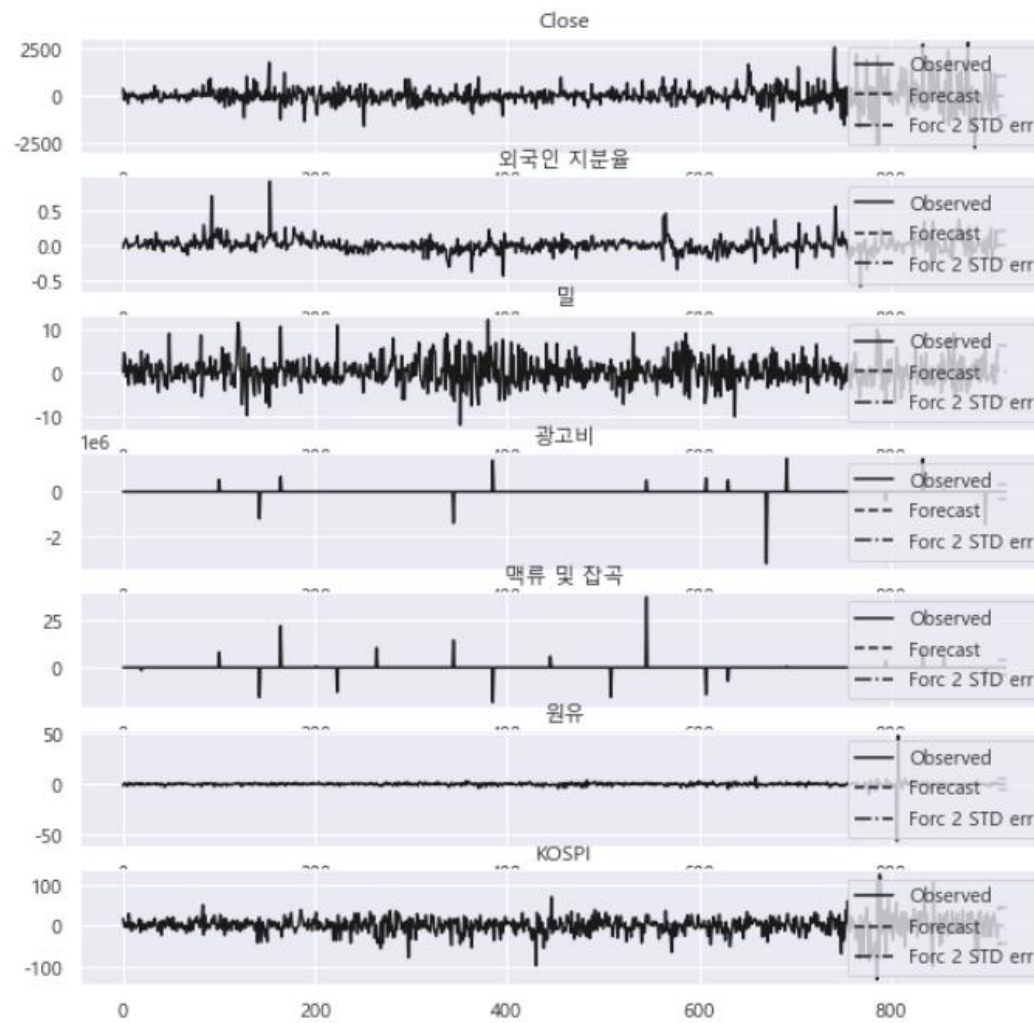
# VAR

## 충격반응함수

Impulse responses



## 예측 그래프



# 감성분석



# 감성분석

```
token_weight.sort_values('가중치').head(10)
```

	토큰	가중치	
395	참이슬	-0.172604	
480	흙	-0.148935	
305	원	-0.143432	
169	보해양조	-0.136632	
32	공략	-0.134073	
235	스마트	-0.124879	
229	수출	-0.116581	
40	국민	-0.114652	
337	인하	-0.110104	
458	필	-0.104825	

```
token_weight.sort_values('가중치').tail(10)
```

	토큰	가중치	
93	도수	0.120271	
433	투자	0.128792	
203	설	0.130540	
270	업	0.140517	
236	스타트	0.144383	
426	테라	0.146516	
319	음료	0.164098	
117	롯데	0.169339	
462	하이트진로	0.222850	
333	인상	0.313029	

## 감성분석

**Y: Close\_change (전날 종가 대비 등락)**

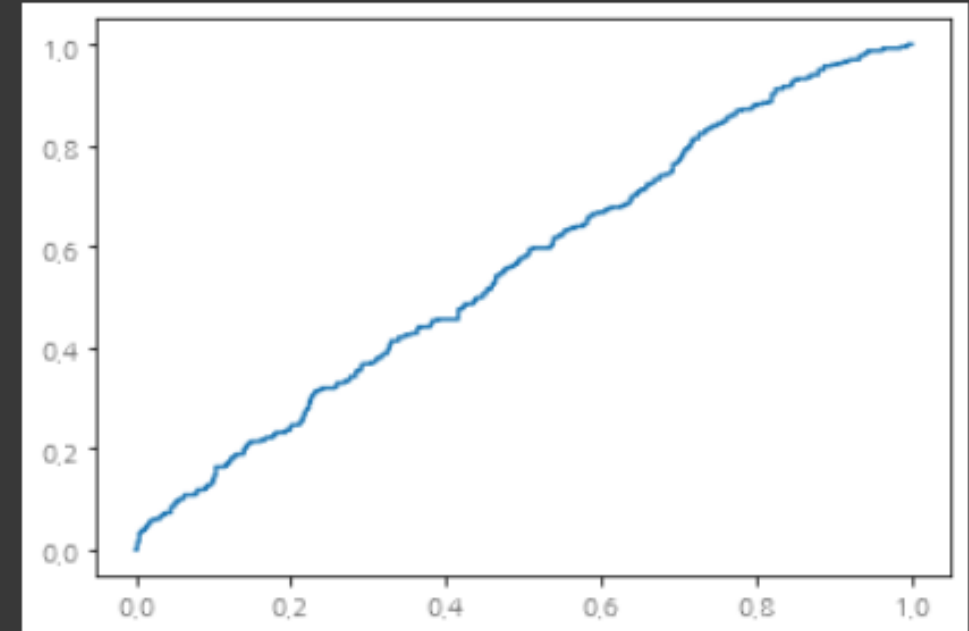
**0, 1 의 비율이 비슷하여 학습하기**

**좋은 것으로 예상했으나**

**AUC는 약 0.56 으로 다소 낮음**

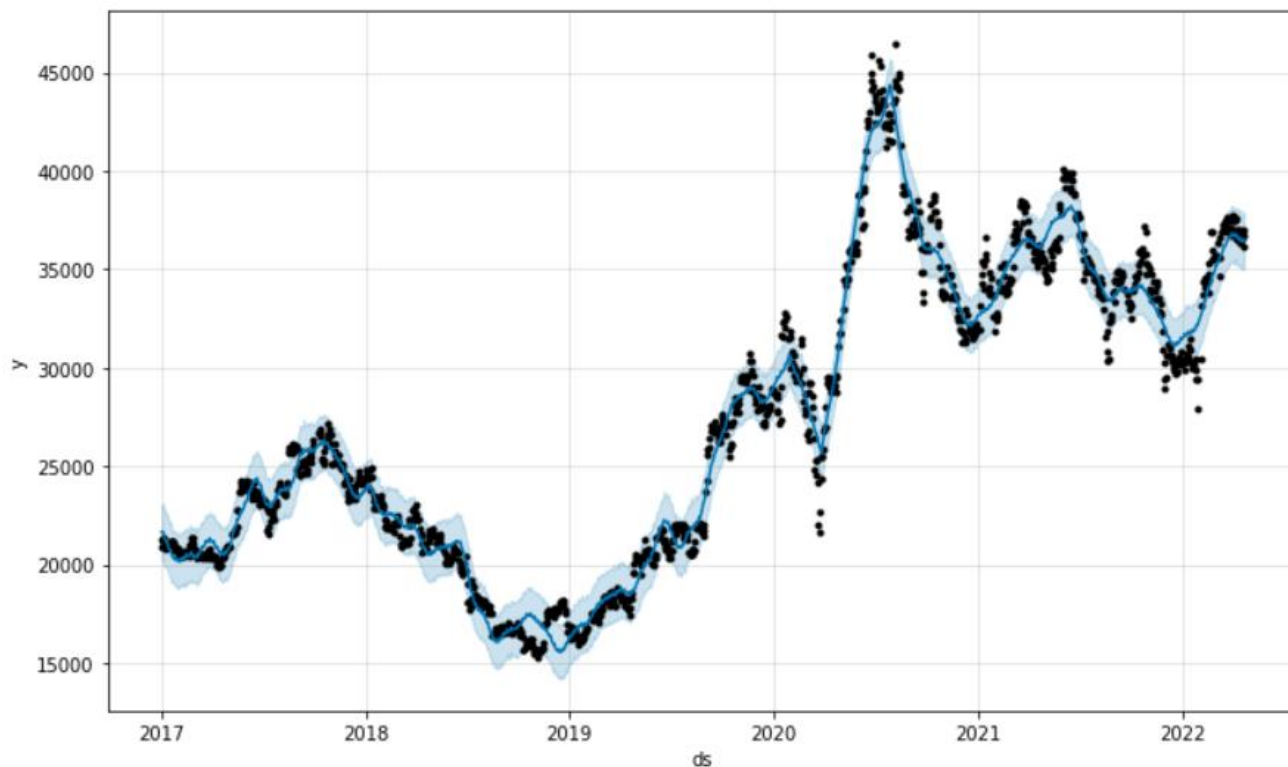
```
plt.plot(fpr, tpr)  
roc_auc_score(y_test, probs)
```

0.5607561519085449



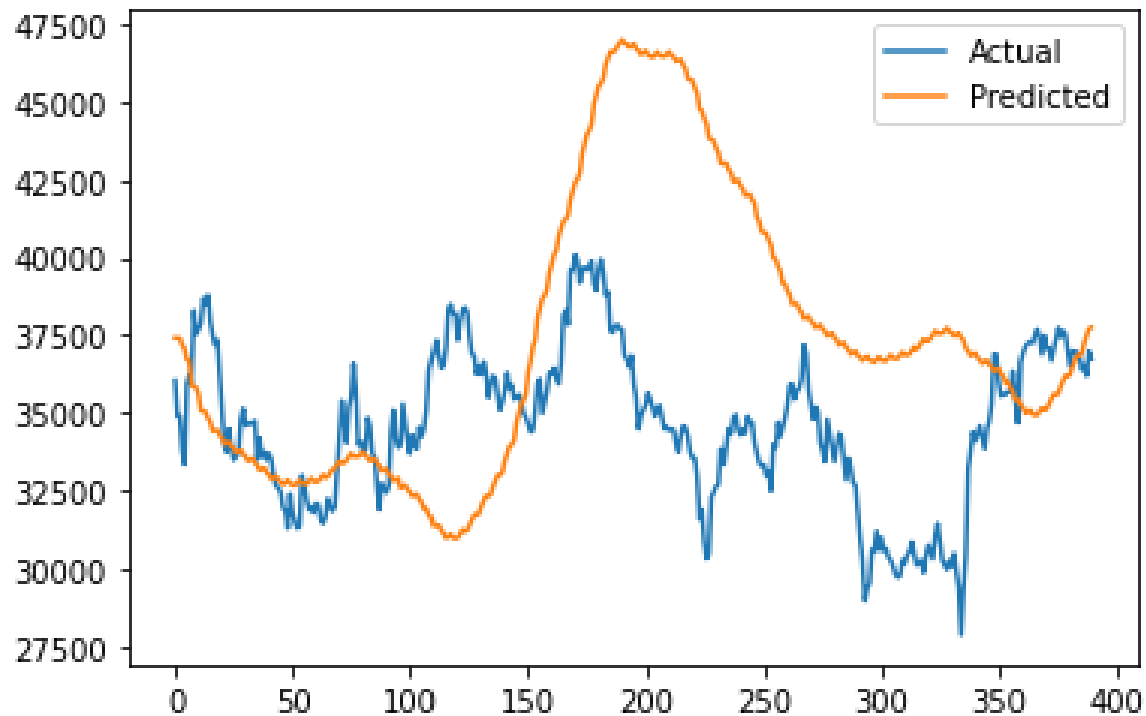
# Prophet

페이스북에서 공개한 시계열 예측 라이브러리  
추세, 계절성, 휴일과 같은 불규칙성 강한 이벤트가 있을 때 예측하기 좋음

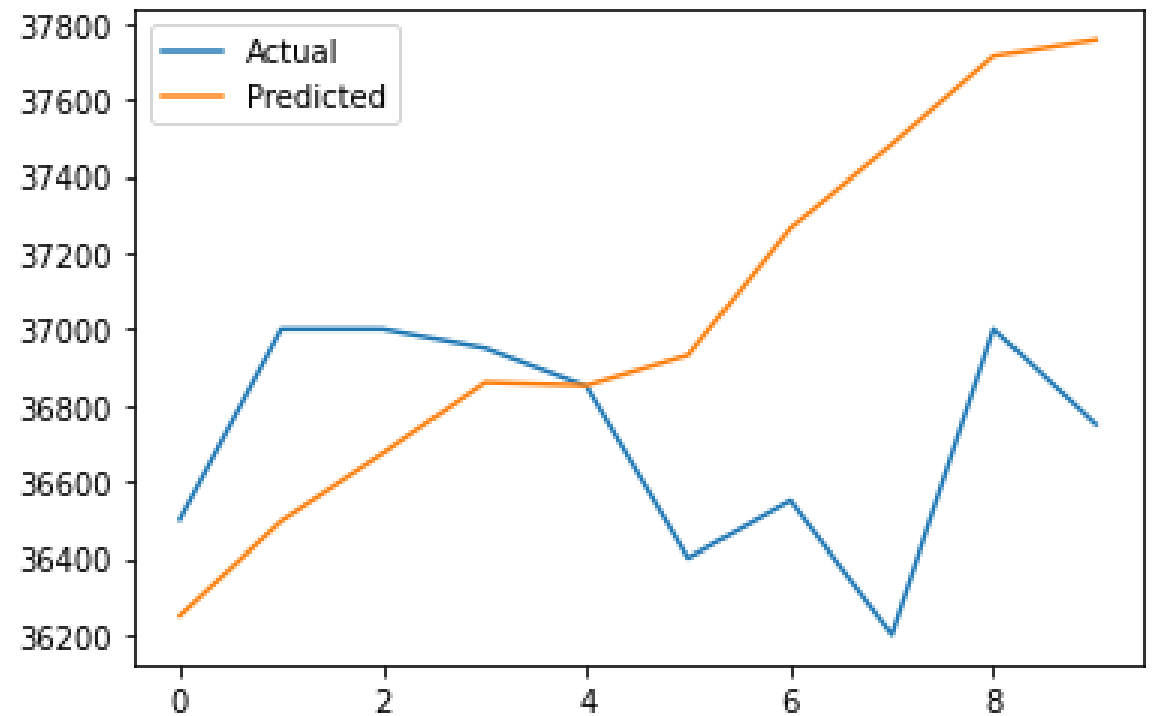


전체기간 예측

# Prophet



390일 예측



10일 예측

## 결과 및 느낀점

목표 달성 80%

- 도메인 지식 부족으로 변수 선별에 어려움을 겪음
- 회귀분석 전 선형성, 등분산성, 정규성, 독립성, 비상관성 확인 미흡
- 시계열 데이터에 분류분석이 부적합

Out [26] :

OLS Regression Results

Dep. Variable:	Close	R-squared (uncentered):	0.976
Model:	OLS	Adj. R-squared (uncentered):	0.976
Method:	Least Squares	F-statistic:	6490.
Date:	Tue, 03 May 2022	Prob (F-statistic):	0.00
Time:	14:58:17	Log-Likelihood:	-12786.
No. Observations:	1303	AIC:	2.559e+04
Df Residuals:	1295	BIC:	2.563e+04
Df Model:	8		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
외국인 지분율	-221.0174	44.076	-5.014	0.000	-307.486	-134.548
밀	76.1931	5.511	13.826	0.000	65.382	87.004
광고비	0.0007	0.000	3.968	0.000	0.000	0.001
맥류 및 잡곡	34.3150	8.838	3.883	0.000	16.977	51.653
원유	-244.0250	13.845	-17.626	0.000	-271.185	-216.865
KOSPI	6.9065	0.430	16.064	0.000	6.063	7.750
USDKRW=X	-23.4672	4.782	-4.908	0.000	-32.848	-14.087
JPYKRW=X	3133.4387	487.215	6.431	0.000	2177.622	4089.255

Omnibus:	58.532	Durbin-Watson:	0.042
Prob(Omnibus):	0.000	Jarque-Bera (JB):	71.080
Skew:	0.470	Prob(JB):	3.67e-16
Kurtosis:	3.652	Cond. No.	1.26e+07

- 도
- 회
- 시

확인 미흡

>> 주식 가격을 이해하고 예측하기 위한 다양한 데이터를 활용하는 데 어려움이 있었다.  
주식 예측에 있어 도메인 지식이 충분히 뒷받침 되어야 함을 깨닫는 계기가 되었다.