



주린이 주목! 지금 팔아?! 말아?!

주식 초보자의 리스크 관리를 위한 매수/매도 추천 서비스



주식 시장의 상황

주식 시장이 급격히 안 좋아지며
많은 사람이 투자 손해를 보게 됨



주식 초보자들의 상황

투자 정보를 알지 못하는
주식 초보자들이 큰 손해를 봄



주식 초보자도 쉽게 참고할 수 있는
투자 인사이트를 제공하기 위한 서비스를 만들자!



시계열자료분석팀 주제분석

주식 초보자의 리스크 관리를 위한 매도 / 매수 추천 서비스

주식 등락률에 영향을 미치는 정보를 모델에 입력 → 모델이 매도/매수를 학습

투자자들에게 현재 상황에 따른 매도 또는 매수를 추천!

▶ 투자 시 참고할 수 있는 **인사이트**를 제공하는 서비스

2 데이터 소개

수집 데이터



SK 하이닉스



신한금융지주회사

신한지주



현대차

다양한 종목에서 응용 가능한 “강건한 모델” 구축을 위해
위 3가지 주식 종목을 분석 대상으로 선정 !

개별 지표

- ✓ 주식 시세 추이 데이터
- ✓ 투자자별 거래 실적 데이터
- ✓ 외국인 보유량 데이터
- ✓ 공매도 데이터
- ✓ 국내 기사 데이터
- ✓ 영문 기사 데이터
- ✓ 네이버 증권 토론방 데이터
- ✓ 네이버 검색량 데이터

(출처) 한국거래소, 네이버 증권,
네이버 데이터랩, 빅카인즈, CNN

공통 지표

- ✓ 코스피 데이터
- ✓ 비트코인 거래 데이터
- ✓ 경제 심리 지수
- ✓ 뉴스 심리 지수
- ✓ 산업 생산 지수
- ✓ 소비자 물가 지수
- ✓ 소비자 신뢰 지수
- ✓ 소비자 심리 지수
- ✓ 실업률
- ✓ 한국은행 기준금리
- ✓ 환율

(출처) 한국거래소, 한국은행 경제통계시스템, KOSIS, invest.com

개별 지표

- ✓ 주식 시세 추이 데이터
- ✓ 투자자별 거래 실적 데이터
- ✓ 외국인 보유량 데이터
- ✓ 공매도 데이터
- ✓ 국내 기사 데이터
- ✓ 영문 기사 데이터
- ✓ 네이버 증권 토론방 데이터
- ✓ 네이버 검색량 데이터

공통 지표

- ✓ 코스피 데이터
- ✓ 비트코인 거래 데이터
- ✓ 경제 심리 지수
- ✓ 농산물 가격 지수
- ✓ 산업 생산 지수
- ✓ 소비자 물가 지수
- ✓ 소비자 신뢰 지수
- ✓ 소비자 심리 지수
- ✓ 실업률



여론 및 투자자 심리에 따른 등락을 파악하기 위해
여론을 반영할 수 있는 데이터 활용 !

2 데이터 소개

수집 데이터

개별 지표

- ✓ 주식 시세 추이 데이터
- ✓ 투자자별 거래 실적 데이터
- ✓ 외국인 보유량 데이터
- ✓ 공매도 데이터
- ✓ 국내 기사 데이터
- ✓ 영문 기사 데이터
- ✓ 네이버 증권 토론방 데이터
- ✓ 네이버 검색량 데이터

공통 지표



데이터 수집 기간 : 2017/06/08 ~ 2023/03/31

네이버 토론방 서비스가 2017/06/08 부터 시작된 관계로

- ✓ 스피 데이터
- ✓ 비트코인 거래 데이터
- ✓ 경제 심리 지수
- ✓ 뉴스 심리 지수
- ✓ 산업 생산 지수
- ✓ 소비자 물가 지수
- ✓ 소비자 신뢰 지수
- ✓ 소비자 심리 지수
- ✓ 실업률
- ✓ 한국은행 기준금리
- ✓ 환율

국내 기사 데이터 감성 분석

분석 데이터

2017-06-08 ~ 2023-03-31까지의 종목별 기사 제목 데이터

| 날짜 | 제목 | label |
|------------|-------------------------------|-------|
| 2017-06-08 | [fnRASSI] 장마감, 거래소 하락 종목 | -1 |
| 2017-06-08 | '오늘의 증시 메모 [6월 8일] | 0 |
| ⋮ | ⋮ | |
| 2023-03-31 | 신한금융 "데이터센터 전력 재생에너지로 조달 " | 1 |
| 2023-03-31 | 신한금융, 데이터센터 전력 100% 재생 에너지 추진 | 1 |



| 날짜 | 기사 감성 점수 |
|------------|----------|
| 2017-06-08 | 0.148148 |
| 2017-06-09 | 0.142857 |
| ⋮ | ⋮ |
| 2023-03-30 | 0.166667 |
| 2023-03-31 | 0.384615 |

경제 뉴스 기사로 학습한
BERT 모델을 통해 라벨링

외국 기사 데이터 감성 분석

기사 데이터

CNN에서 크롤링 한 2017년 7월부터 2023년 3월까지 총 9995행의 데이터

| 날짜 | 제목 | score |
|------------|--|---------|
| 2017-07-19 | US general warns of ... control killer ... | -0.6908 |
| 2017-07-21 | Pompeo signals want for N. Korea ... | 0.0772 |
| ⋮ | ⋮ | |
| 2023-03-25 | How AI turned the ancient sport of Go ... | 0 |
| 2023-03-29 | US & South Korea stage joint military ... | 0 |

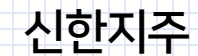
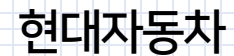
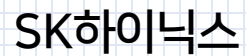


| 날짜 | 기사 감성 점수 |
|------------|----------|
| 2017-07-14 | 0.365775 |
| 2017-07-15 | 0.261127 |
| ⋮ | ⋮ |
| 2023-03-30 | 0.120400 |
| 2023-03-31 | 0.079550 |



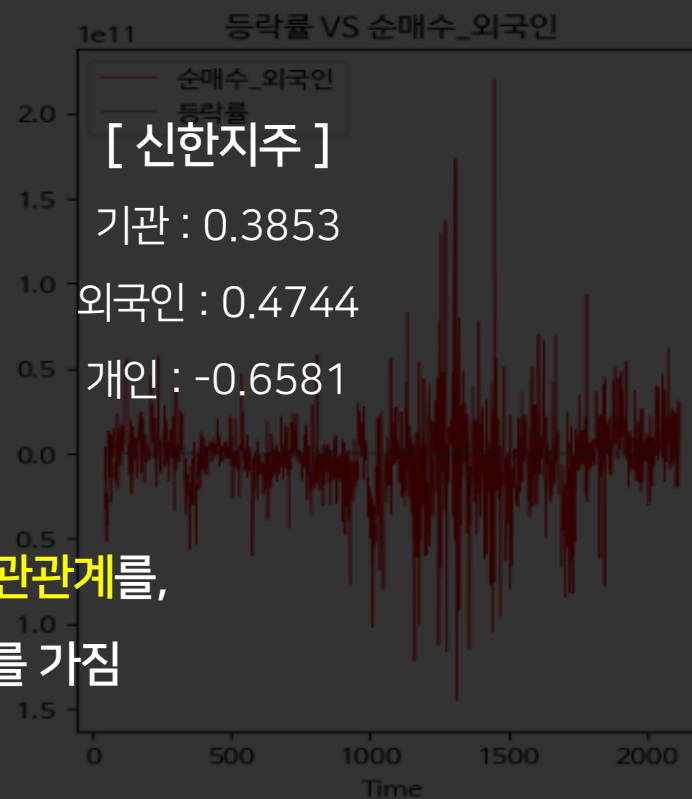
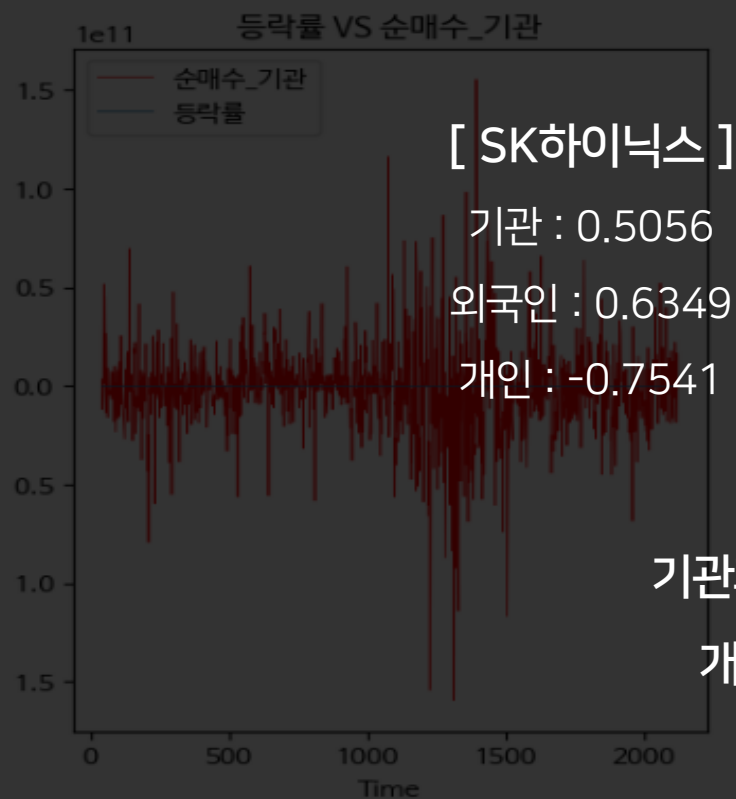
`nltk.sentiment.vader`을 통해 분석

등락률과 상관관계



4 EDA

순매수_기관,개인,외국인

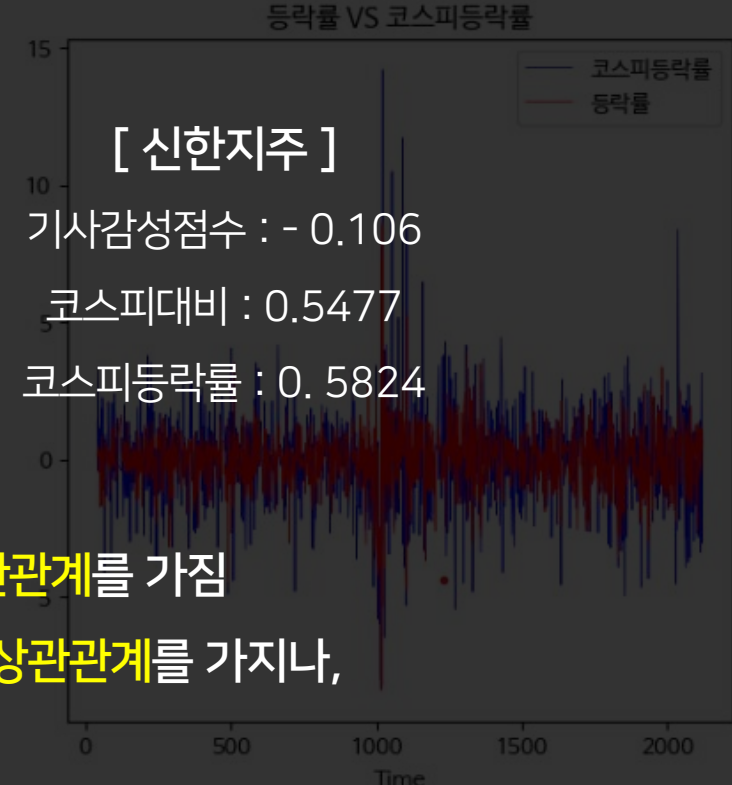
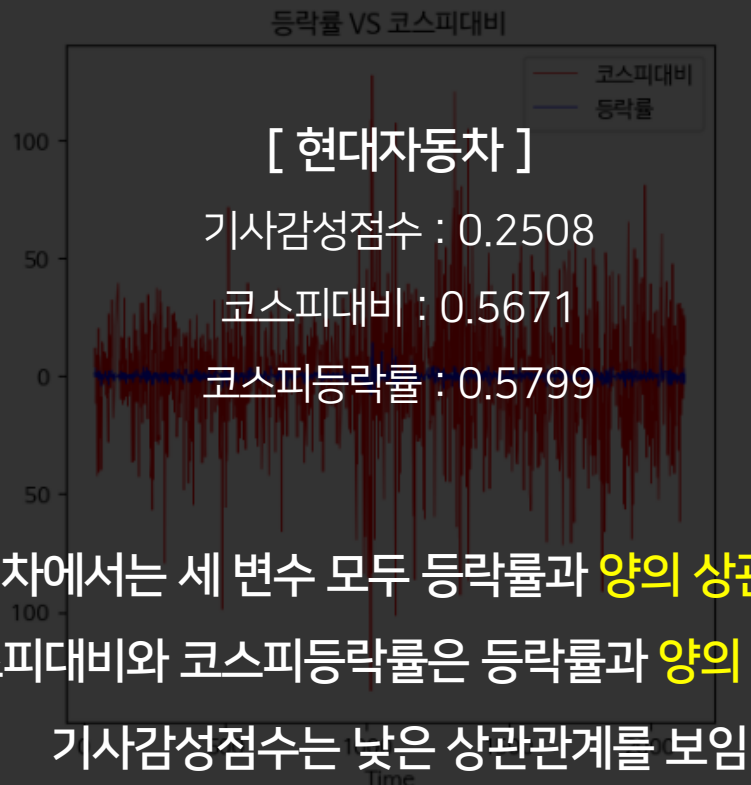
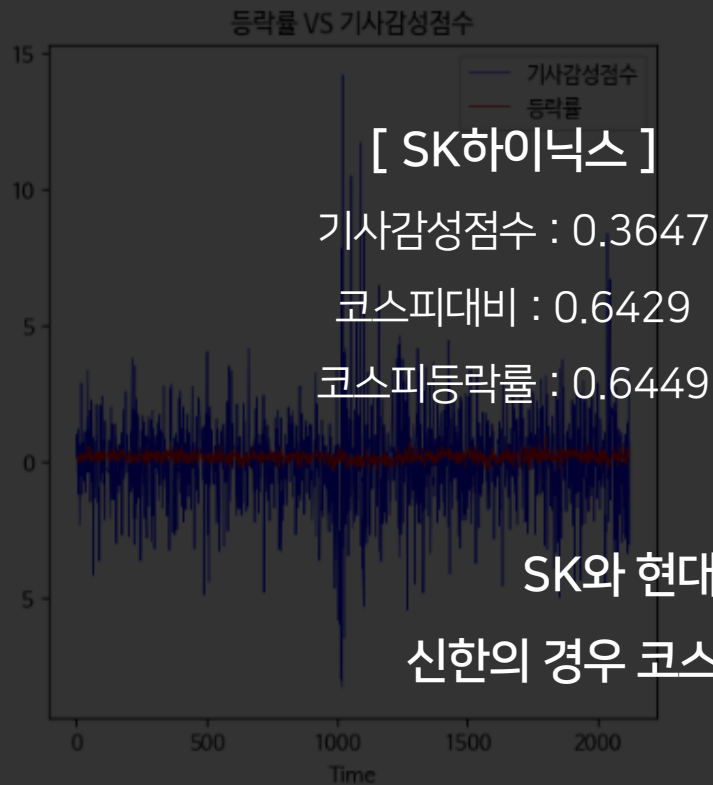


기관과 외국인의 순 매수량은 등락률과 양의 상관관계를,
개인의 순 매수량은 등락률과 음의 상관관계를 가짐

신한지주

4 EDA

코스피대비, 코스피등락률

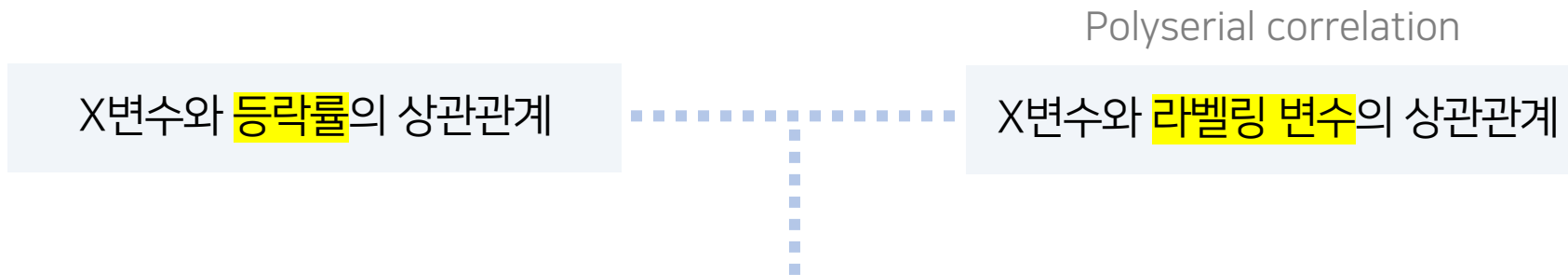


SK와 현대차에서는 세 변수 모두 등락률과 양의 상관관계를 가짐
신한의 경우 코스피대비와 코스피등락률은 등락률과 양의 상관관계를 가지나,
기사감성점수는 낮은 상관관계를 보임

신한지주

4 EDA

Y변수 라벨링



두 상관관계가 비슷해지는 지점을 기준으로 라벨링 !



3일 등락률: 5% 기준 라벨링

5일 등락률: 5% 기준 라벨링

3일 후 등락률 $\leq -5\%$: 매도 (sell)

3일 후 등락률 $\geq 5\%$: 매수 (buy)

Otherwise : 유지 (maintain)

Y변수 라벨링 ; 라벨 Y변수와의 상관계수 비교

| 신한지주 | | | |
|---------|---------|--------|--------|
| 순매수_기관 | 0.3853 | 기사감성점수 | - |
| 순매수_외국인 | 0.4744 | 코스피대비 | 0.5477 |
| 순매수_개인 | -0.6581 | 코스피등락률 | 0.5824 |

| SK 하이닉스 | | | |
|---------|---------|--------|--------|
| 순매수_기관 | 0.5056 | 기사감성점수 | 0.3647 |
| 순매수_외국인 | 0.6349 | 코스피대비 | 0.6429 |
| 순매수_개인 | -0.7541 | 코스피등락률 | 0.6449 |

| 현대차 | | | |
|---------|---------|--------|--------|
| 순매수_기관 | 0.4678 | 기사감성점수 | 0.2508 |
| 순매수_외국인 | 0.4444 | 코스피대비 | 0.5671 |
| 순매수_개인 | -0.5798 | 코스피등락률 | 0.5799 |

해석

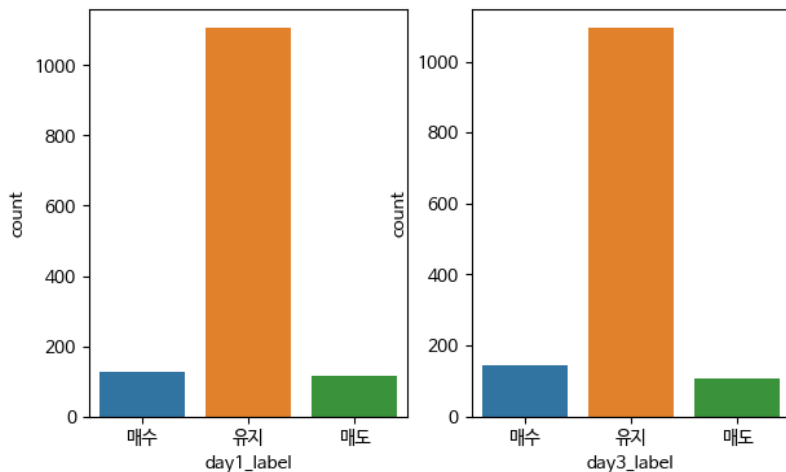
양의 상관계수를 갖는 변수들은
해당 변수가 증가할 때,
매수일 확률이 높아지고
매도일 확률은 낮아짐

해석

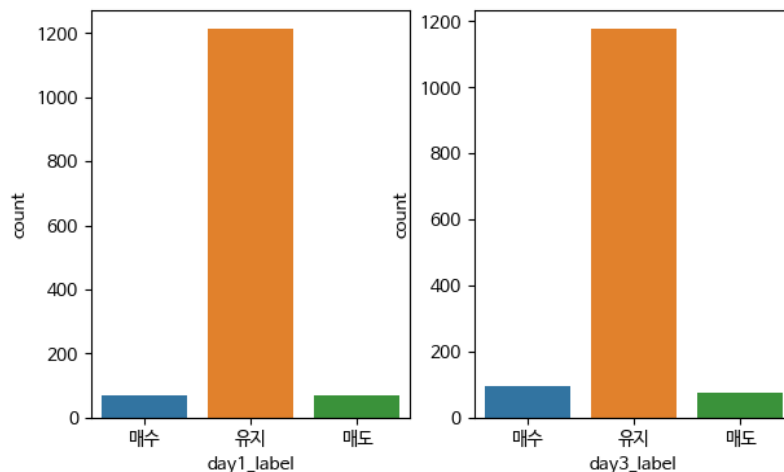
음의 상관계수를 갖는 변수들은
해당 변수가 증가할 때,
매수일 확률이 낮아지고
매도일 확률은 높아짐

종목별 매도 / 유지 / 매수 라벨링 결과

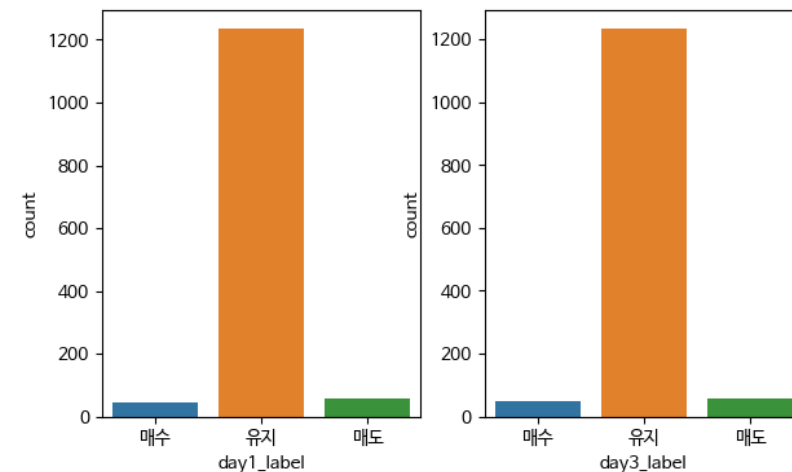
SK 하이닉스 라벨 분포



현대차 라벨 분포



신한지주 라벨 분포



모든 종목에서 '유지' 대비
'매수' / '매도'의 클래스 불균형 심각



전체 정확도는 높지만, 모델 예측이 전체적으로 '유지'에 편향되어
'매수'와 '매도'를 제대로 예측하지 못하는 문제 발생

5 모델링 과정

변수선택

1

[인과관계 검정]

2

[VIF]

3

[Feature Importance]

4

[KS 검정]

5

[Full Model]

But...

대부분의 경우에선

Full model의
성능이 가장 좋았다...



Input

현재 시점의 X변수 (minmax scaled / full model)

Output

앞으로 5일 이내의 주가 등락 예측에 따른 현재 시점에서의 매수/유지/매도 추천

- t+1 시점부터 t+6 시점까지, 5일 동안 등락률이 5%이상 상승할 것으로 예상되면 매수 추천
- t+1 시점부터 t+6 시점까지, 5일 동안 등락률이 5%이상 하락할 것으로 예상되면 매도 추천
- t+1 시점부터 t+6 시점까지, 5일 동안 등락률의 절댓값이 5% 이내로 예상되면 유지 추천



클래스 불균형이 가장 덜한 SK하이닉스 데이터로 모델 적합 후, 동일한 모델을 3종목 모두에 적용



모델 선정 기준: 매수/매도를 얼마나 잘 맞추는가? (매수/매도 정확도 높은 것 중 f1-score 높은 모델)

01 매수, 매도, 유지 정확도의 평균

```
cm=confusion_matrix(y_test, y_pred)
bacc=cm[0,0]/sum(cm[0]) # 매수 정확도
sacc=cm[2,2]/sum(cm[2]) # 매도 정확도
macc=cm[1,1]/sum(cm[1]) # 유지 정확도
rst=np.mean([bacc,sacc,macc])
```

02 매수, 매도 정확도의 평균

```
cm=confusion_matrix(y_test, y_pred)
bacc=cm[0,0]/sum(cm[0]) # 매수 정확도
sacc=cm[2,2]/sum(cm[2]) # 매도 정확도
rst=np.mean([bacc,sacc])
```

03 매수, 매도 정확도, F1 Score의 평균

```
cm=confusion_matrix(y_test, y_pred)
bacc=cm[0,0]/sum(cm[0]) # 매수 정확도
sacc=cm[2,2]/sum(cm[2]) # 매도 정확도
f1=sum(scores)/len(scores)
rst=np.mean([bacc,sacc,f1])
```

04 매수, 매도의 정확도, 정밀도의 평균

```
cm=confusion_matrix(y_test, y_pred)
bacc=cm[0,0]/sum(cm[0]) # 매수 정확도
sacc=cm[2,2]/sum(cm[2]) # 매도 정확도
bpre=cm[0,0]/np.sum(cm, axis=0)[0] #매수 정밀도
spre= cm[2,2]/np.sum(cm, axis=0)[2] #매도 정밀도
rst=np.mean([bacc,sacc,bpre,spre])
```

5 모델링 과정

시도한 모델

모델 종류

- LSTM
- CNN
- SVM
- Logistic regression
- 나이브 베이즈
- XGB
- LGBM
- LSTM 회귀
- LSTM-CNN

할일을 눈앞에 산더미처럼 쌓아 놓으니
더 하기 싫어진다



SK 하이닉스 기준 4가지 score로
optuna 진행해 최적의 모델 선택

6 최종 모델

최종 모델 소개

XGB classifier

데이터 : SK 하이닉스

변수 : Full Model

평가지표 : custom한 optuna score

분류 : 다중 분류(매수, 매도, 유지)

LSTM regressor

데이터 : SK 하이닉스

변수 : VIF로 선별된 변수

특징 : 라벨 인코딩 한 결과를

회귀로 예측하고

Threshold 결정 함수로 분류

6 최종 모델

XGB Classifier

1. 변수 선택

데이터 : SK 하이닉스



클래스 불균형이 가장 심하지 않은 SK하이닉스 데이터를 기준으로
하이퍼파라미터 튜닝 후 나머지 종목에 적용할 예정 !

변수 : Full Model

변수

X : '종가', '대비', '등락률', '거래량', '거래대금', '시가총액', '외국인 보유수량', '외국인 지분율', '토론방', '순매수_기관', '순매수_기타법인', '순매수_개인', '순매수_외국인', '검색어', '보도량', '기사감성점수', '뉴스심리지수', '비트코인종가', '비트코인거래량', '비트코인변동', '코스피종가', '코스피등락률', '코스피거래량', '코스피거래대금', '코스피시가총액', '한은금리', '원/미국달러', '원/위안', '원/일본엔(100엔)', '원/유로', '경제심리지수(원계열)', '경제심리지수(순환변동치)', '산업생산지수', '물가상승률', '소비자신뢰지수', '소비자심리지수', '경제활동참가율(%)', '실업률(%)', '고용률(%)', '코스피대비'

Y : 'day5_label'

6 최종 모델

XGB Classifier

2. 라벨인코딩

target label (day5_label)에 대한 라벨인코딩 진행

| | |
|--------------|---|
| buy(매수) | 0 |
| maintain(유지) | 1 |
| sell(매도) | 2 |

3. MinMax Scaling

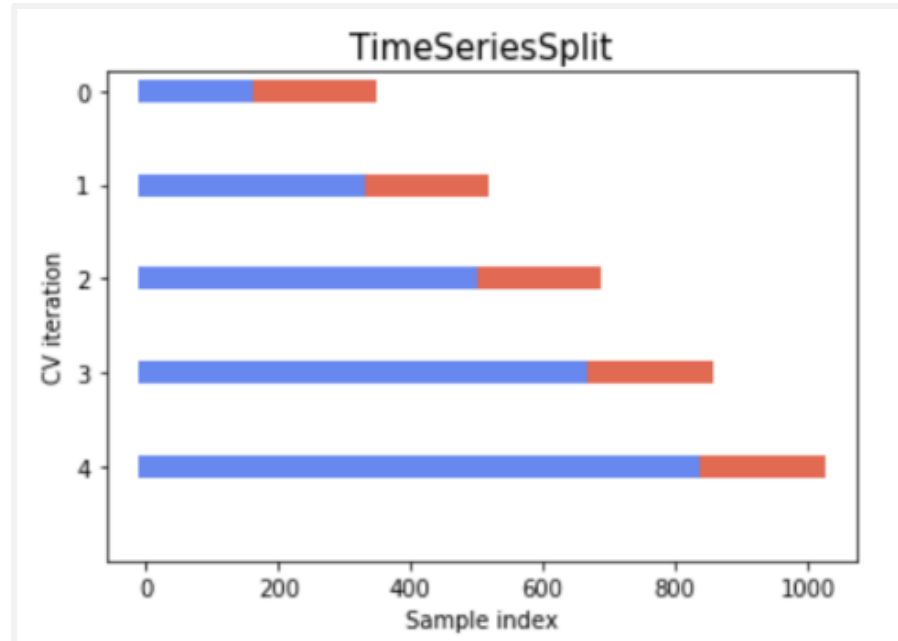
$$\frac{x - \text{Min}(X)}{\text{Max}(X) - \text{Min}(X)}$$

- 모든 연속형 X 변수에 대해서 MinMax scaling 진행
- 변수의 범위를 바꿔주는 정규화 스케일링 (범위 : [0, 1])
- 종목 간 스케일 차이를 줄여 동일한 하이퍼파라미터에 적합하기 위함

6 최종 모델

XGB Classifier

3. Expanding Window CV



n_splits = 4 인 Expanding Window CV 활용 !

Split 횟수가 많아지면 validation set의 크기가 작아지기 때문에,
하나의 validation set에서의 클래스 불균형이 너무 심각해지는 문제를 예방하기 위함

6 최종 모델

XGB Classifier

4. 클래스 가중치

각 클래스별 비율의 역수를 클래스별 샘플 가중치로 활용 !

```
class_weights = class_weight.compute_sample_weight( class_weight='balanced', y=y_train )
```

▶ 불균형한 학습 데이터에 대해 클래스별 샘플 가중치를 계산해주는 함수

```
xgb_model=xgb.XGBClassifier(**params, random_state = 42)  
xgb_model.fit(x_train, y_train, sample_weight=class_weights)
```

이런 식으로 fit 함수에서 활용 가능 !



6 최종 모델

XGB Classifier

5. Optuna 하이퍼파라미터 튜닝

XGBoost Classifier 하이퍼파라미터

- max_depth : 트리의 최대 깊이, 깊어질수록 복잡한 모델
- learning_rate : 학습율
- n_estimators : 트리의 개수
- min_child_weight : 분기에 필요한 최소 hessian 가중치, 작을수록 복잡한 모델
- gamma : 분기에 필요한 최소한의 loss 감소 비율, 작을수록 복잡한 모델
- subsample : 각 트리마다의 데이터 샘플링 비율
- colsample_bytree : 각 트리마다의 feature 샘플링 비율
- reg_alpha : L1 가중치, 작을수록 복잡한 모델
- reg_lambda : L2 가중치, 작을수록 복잡한 모델

6 최종 모델

XGB Classifier

5. Optuna 하이퍼파라미터 튜닝

거래 손익에 직접적 영향을 줄 수 있는 '매수', '매도'를 잘 맞추는 모델을 만들기 위해
클래스 개수가 많은 유지를 제외한 매수, 매도의 **정확도**를 평가 지표에 포함

$\frac{TP + TN}{TP + TN + FP + FN}$ 모델이 과도하게 매수, 매도로만 예측하는 것을 예방하고
'유지'에 대한 예측력도 유지하기 위해, 매수, 매도의 **정밀도**를 평가 지표에 포함

Optuna 평가 지표

매수 정확도, 매수 정밀도, 매도 정확도, 매도 정밀도의 평균

6 최종 모델

XGB Classifier

5. Optuna 하이퍼파라미터 튜닝



투자자별 순매수량, 비트코인변동, 토론방 게시글 수, 뉴스심리지수, 기사 보도량, 기사감성점수, 검색어량

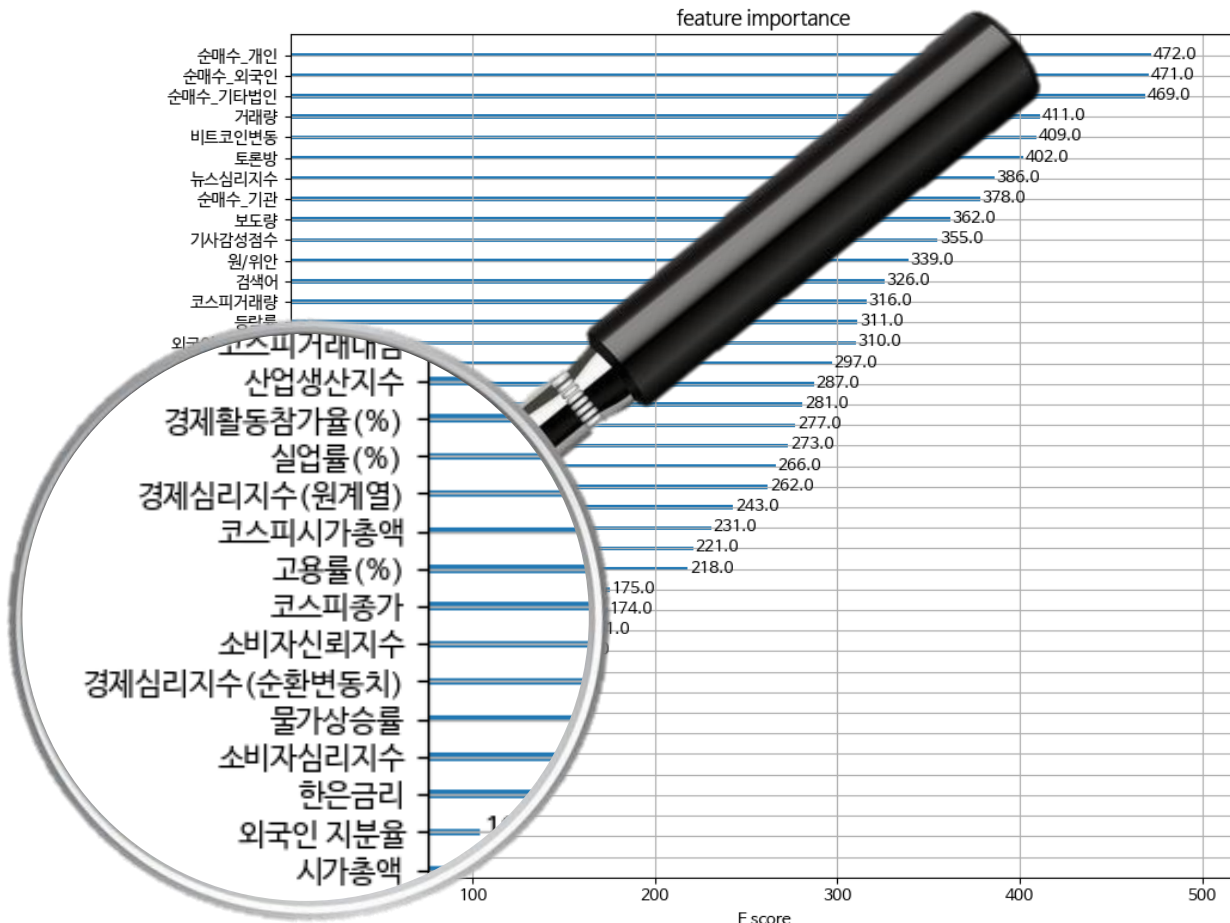


투자자별 순매수량 데이터와
여론 및 투자자 심리 관련 데이터가
중요한 변수로 작용

6 최종 모델

XGB Classifier

5. Optuna 하이퍼파라미터 튜닝



산업생산지수, 경제활동참가율, 실업률, 코스피, 경제심리지수, 고용률, 물가상승률, 한은금리, ...



반면, 거시경제 관련 데이터는 상대적으로 덜 중요한 변수로 나타남

6 최종 모델

XGB Classifier

6. 예측

- with test set

[SK하이닉스]

```
===== SK 하이닉스 =====  
[[27  7  3]  
 [45 90 63]  
 [ 6  7 33]]
```

전체 정확도 : 0.5338078291814946
전체 f1-score : 0.5563170430999059

매수 정확도 : 0.7297297297297297
매도 정확도 : 0.717391304347826
유지 정확도 : 0.45454545454545453

F1 score : 0.56

매수 정확도 : 0.73

매도 정확도 : 0.72

[현대차]

```
===== 현대차 =====  
[[ 2  1  0]  
 [24 117 11]  
 [ 0  3  3]]
```

전체 정확도 : 0.7577639751552795
전체 f1-score : 0.8229783067649851

매수 정확도 : 0.6666666666666666
매도 정확도 : 0.5
유지 정확도 : 0.7697368421052632

F1 score : 0.82

매수 정확도 : 0.66

매도 정확도 : 0.5

[신한지주]

```
===== 신한지주 =====  
[[ 10  8  1]  
 [ 40 163 37]  
 [ 0  13 13]]
```

전체 정확도 : 0.6526315789473685
전체 f1-score : 0.6975956808520171

매수 정확도 : 0.5263157894736842
매도 정확도 : 0.5
유지 정확도 : 0.6791666666666667

F1 score : 0.69

매수 정확도 : 0.52

매도 정확도 : 0.5

6 최종 모델

LSTM Regressor

1. 변수 선택

데이터 : SK 하이닉스

변수 : VIF를 통해서 선별한 변수

변수

X : '경제심리지수(순환변동치)', '시가총액', '비트코인종가', '원/미국달러', '소비자심리지수', '코스피거래대금', '순매수_개인', '순매수_외국인', '산업생산지수', '코스피거래량', '뉴스심리지수', '원/유로', '토론방', '실업률(%)', '원/일본엔(100엔)', '거래량', '기사감성점수', '외국인 보유수량', '코스피등락률', '경제활동참가율(%)', '검색어', '보도량', '순매수_기관', '비트코인거래량', '비트코인변동', '5일 등락률'

Y : 'day5_label'

6 최종 모델

LSTM Regressor

2. 라벨인코딩

target label (day5_label)에 대한 라벨인코딩 진행

| | |
|--------------|---|
| buy(매수) | 0 |
| maintain(유지) | 1 |
| sell(매도) | 2 |

3. MinMax Scaling

$$\frac{x - \text{Min}(X)}{\text{Max}(X) - \text{Min}(X)}$$

- 모든 연속형 X 변수에 대해서 MinMax scaling 진행
- 변수의 범위를 바꿔주는 정규화 스케일링 (범위 : [0, 1])
- 분류 모델보다 **회귀 모델**에 적합

6 최종 모델

LSTM Regressor

3. Window dataset 생성

EXAMPLE) window size = 3

sliding window

| 날짜 | 비트코인종가 | 실업률 | 거래량 | 검색어 | 보도량 | ... | day5_label |
|------------|--------|-----|---------|----------|-----|-----|------------|
| 2017-07-11 | 2324.3 | 3.4 | 3187332 | 8.10396 | 58 | ... | 1 |
| 2017-07-12 | 2403.1 | 3.4 | 3462150 | 8.16834 | 65 | ... | 1 |
| 2017-07-13 | 2362.4 | 3.4 | 5432312 | 11.22361 | 90 | ... | 1 |
| 2017-07-14 | 2234.2 | 3.4 | 2931832 | 9.64898 | 72 | ... | 0 |
| 2017-07-17 | 2233.4 | 3.4 | 2804598 | 9.12856 | 50 | ... | 0 |
| 2017-07-18 | 2320.2 | 3.4 | 2066194 | 7.92513 | 76 | ... | 1 |
| 2017-07-19 | 2282.6 | 3.4 | 2009799 | 7.69511 | 42 | ... | 1 |
| 2017-07-20 | 2866.0 | 3.4 | 1647153 | 7.71154 | 31 | ... | 1 |

→ X_train[0]

→ y_train[0]

6 최종 모델

LSTM Regressor

3. Window dataset 생성

Window size = 10으로 window dataset 생성

기존 data size : (1409, 27)



window sliding 적용

data size : (1399, 10, 27)



window size

4. train, validation, test split

data size : (1399, 10, 27)



Train set : (1120, 10, 27)

Validation set : (140, 10, 27)

Test set : (139, 10, 27)

6 최종 모델

LSTM Regressor

5. LSTM Regressor

- with train set

| | |
|---------------|----------|
| hidden_size | 2 |
| num_layers | 1 |
| learning_rate | 0.0001 |
| loss function | MSE loss |
| optimizer | Adam |
| epoch | 8000 |



데이터의 y label은 0, 1, 2로 구성된 범주형 변수지만 회귀를 통해서 예측

→ 최적의 모델 체크포인트로 저장

6 최종 모델

LSTM Regressor

5. LSTM Regressor

- with validation set

각 예측값을 범주형 데이터로 분리하기 위한 threshold(경계값) 결정

| | 매도 정확도 | 매수 정확도 | 유지 정확도 | 매도 정밀도 | 매수 정밀도 | f1 score | 평균 |
|------|--------|--------|--------|--------|--------|----------|--------|
| 조합 1 | 0 | 0 | X | 0 | 0 | X | |
| 조합 2 | 0 | 0 | 0 | X | X | X | ★BEST★ |
| 조합 3 | 0 | 0 | X | X | X | 0 | |
| 조합 4 | 0 | 0 | 0 | X | X | 0 | |



Threshold 함수의 역할

- Validation set 기준, 매수 정확도, 매도 정확도, 유지 정확도의 평균이 최대가 되는 조합으로 threshold 결정
- 결정된 threshold 바탕으로 매수/매도/유지 예측 자동화

6 최종 모델

LSTM Regressor

6. 예측

- with test set

[신한지주]

```
===== 신한지주 =====  
[[ 23 15 0]  
 [ 24 529 62]  
 [ 1 15 39]]
```

전체 정확도 : 0.8347457627118644
전체 f1-score : 0.8503657789754228

매수 정확도 : 0.6052631578947368
매도 정확도 : 0.7090909090909091
유지 정확도 : 0.8601626016260162

F1 score : 0.85

매수 정확도 : 0.61

매도 정확도 : 0.71

[SK하이닉스]

```
===== SK하이닉스 =====  
[[104 18 0]  
 [113 313 45]  
 [ 3 27 76]]
```

전체 정확도 : 0.7052932761087267
전체 f1-score : 0.7165113879684445

매수 정확도 : 0.8524590163934426
매도 정확도 : 0.7169811320754716
유지 정확도 : 0.6645435244161358

F1 score : 0.72

매수 정확도 : 0.85

매도 정확도 : 0.67

[현대차]

```
===== 현대차 =====  
[[ 9 9 3]  
 [ 47 248 57]  
 [ 4 7 14]]
```

전체 정확도 : 0.6809045226130653
전체 f1-score : 0.7416229778038823

매수 정확도 : 0.42857142857142855
매도 정확도 : 0.56
유지 정확도 : 0.7045454545454546

F1 score : 0.74

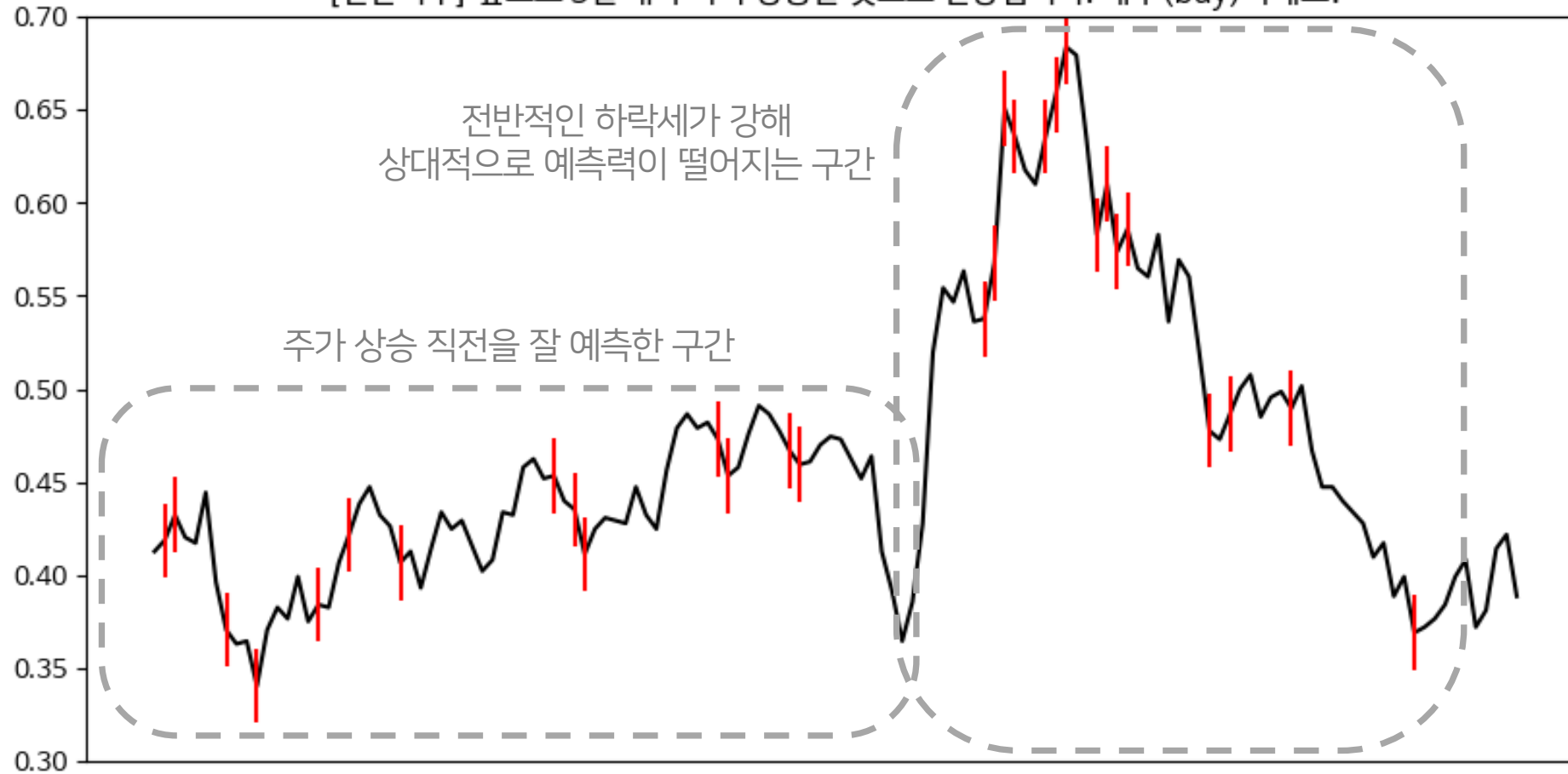
매수 정확도 : 0.43

매도 정확도 : 0.56

7 결과분석

예측 결과 시각화

[신한지주] 앞으로 5일 내 주가가 상승할 것으로 전망됩니다. 매수(buy)하세요!

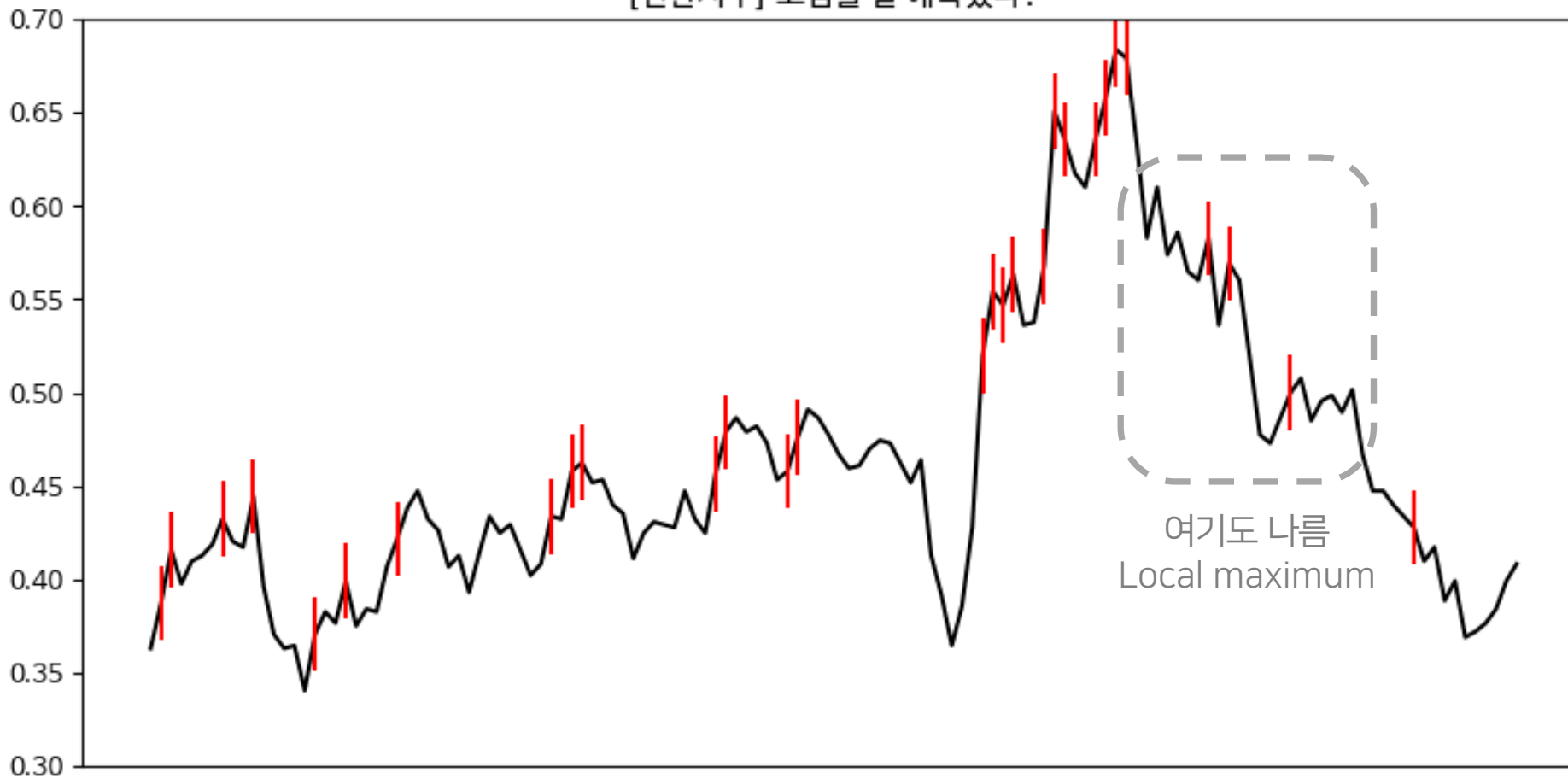


빨간색 지점을 기준으로
5일 이내 주가가 5% 이상
상승될 전망
▶ 매수 추천

7 결과분석

예측 결과 시각화

[신한지주] 고점을 잘 예측했나?

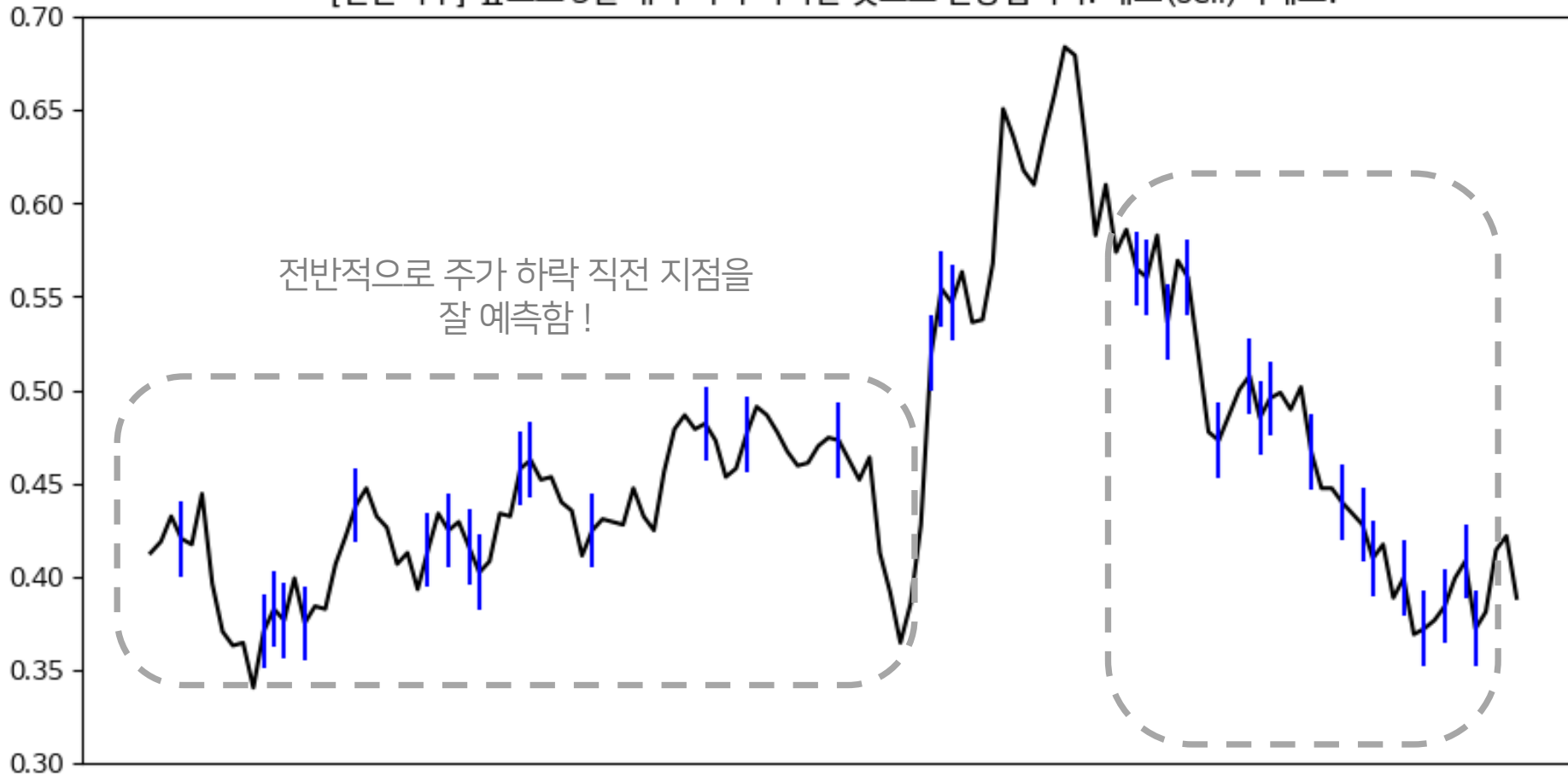


빨간색 지점을 5일 이후
시점으로 이동한 결과,
대체로 주가가 국소적 극대를
기록하는 지점에 잘 매칭됨!

7 결과분석

예측 결과 시각화

[신한지주] 앞으로 5일 내 주가가 하락할 것으로 전망됩니다. 매도(sell)하세요!

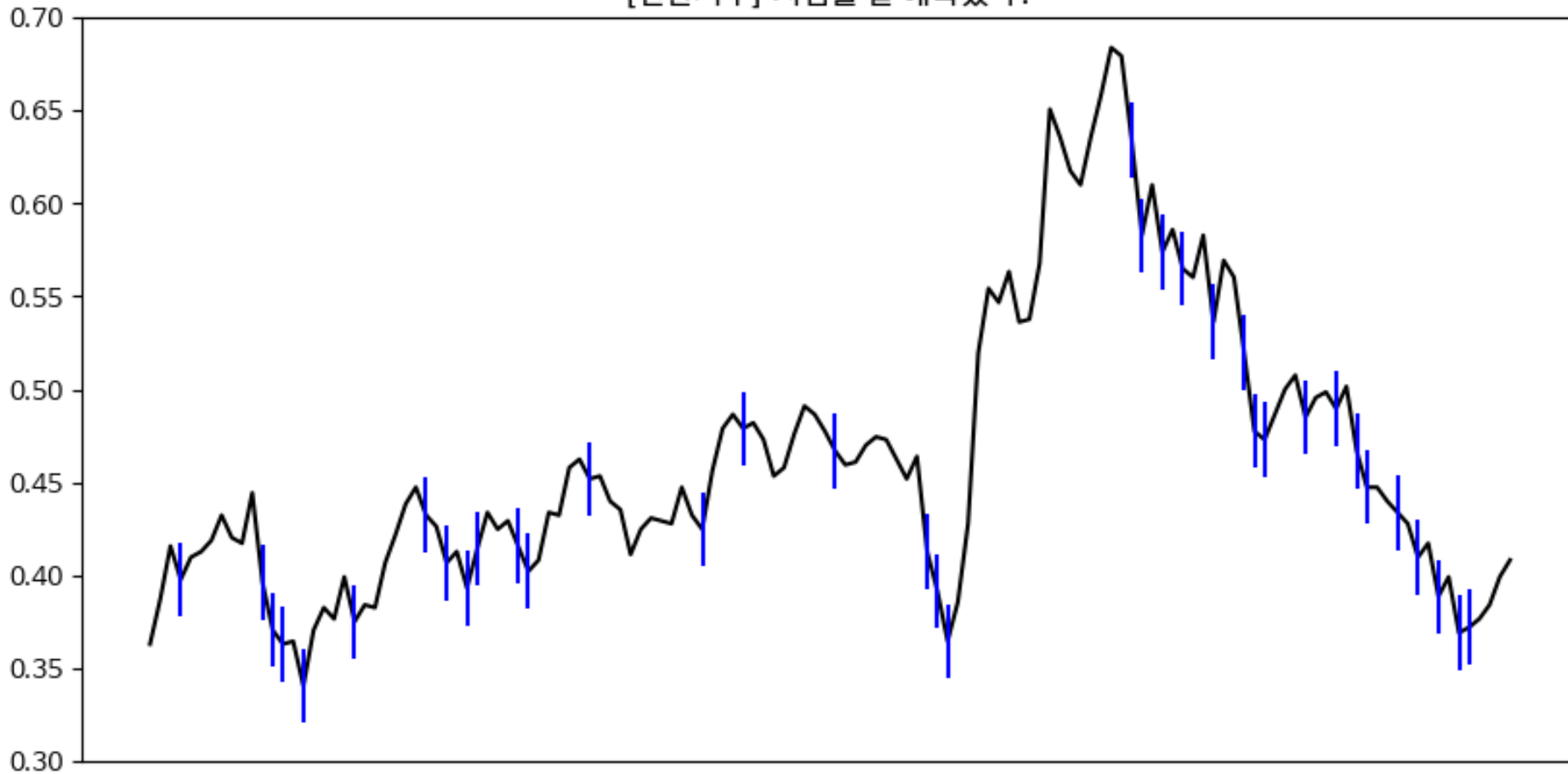


파란색 지점을 기준으로
5일 이내 주가가 5% 이상
하락할 전망
▶ 매도 추천

7 결과분석

예측 결과 시각화

[신한지주] 저점을 잘 예측했나?



파란색 지점을 5일 이후
시점으로 이동한 결과,
대체로 주가가 국소적 극소를
기록하는 지점에 잘 매칭됨!

7 결과분석

의의 및 한계

의의

- 등락률 예측을 위한 데이터로 정형 및 비정형 데이터 및 다양한 데이터를 사용하여 유의미한 결과 도출
- 주식 데이터의 특성(imbalanced data, time series data)을 반영하여 분석 진행
- 도메인에 크게 영향을 받지 않고 일정한 정확도를 보이는 강건한 모델

한계

- 실생활에 적용하기 위해서는 데이터 수집 자동화 필요
- 데이터셋으로 사용했던 3가지 종목 외에 더 다양한 종목에 적용 가능한지 시험해보지 않은 것

감사합니다!!!



지금까지
짱시계열이었습니다~~