

텍스트 마이닝 프로젝트

메타버스 감성분석 및 공포-탐욕 지표 제안

손우진, 김홍민, 장주현

주식 시장의 비이성적 행동과 공개정보의 역할

손판도 and 이형기. (2020). 주식시장의 비이성적 행동과 공개정보의 역할-한국 매스미디어로부터 증거 -. 경영과 정보연구, 39(3), 83-98.

- 비정보 투자자의 비이성적 투자행위는 매스 미디어의 정보를 통해 결정되는 동향이 있음
- Antweir와 Frank(2004)는 정보를 주고받는 커뮤니케이션행위와 주식거래량 간에 밀접한 관계가 존재한다는 것을 밝힘

주식 시장에서 생존하려면 감정을 경계하라

DBR 박승욱 저널 | 322호 2021년 6월 Issue 1호

- 신경과학의 발전에 힘입어 감정과 인지적 사고는 항상 뒤엉켜 상호작용을 함
- 주식시장의 이례적 행동의 대표적 예시인 주가 모멘텀과 역전 현상은 투자자의 심리적 편향에 기인하여 발생한 결과로 해석함

expectancy effects

Fear And Greed를 통해

주식시장 거품을 유추하여

매매 의사결정에 도움이 되길 기대함

Word2Vec (종토팅)

```
model.wv.most_similar('네이버')
```

```
(('엔씨', 0.9307807683944702),  
(('전문기대비', 0.9132906794548035),  
(('반토막에', 0.8859274387359619),  
(('났다', 0.8731440901756287),  
(('급하긴', 0.8701403737068176),  
(('유무중', 0.8695124387741089),  
(('한컴', 0.867789089679718),  
(('알체라', 0.865554690361023),  
(('개쪽락해네', 0.8645421862602234),  
(('이랜텍', 0.8616209030151367))]
```

```
model.wv.most_similar('엔씨')
```

```
(('네이버', 0.9307807683944702),  
(('전문기대비', 0.9166899919509888),  
(('반토막에', 0.9113982915878296),  
(('유무중', 0.9064014554023743),  
(('월구는', 0.9051628112792969),  
(('없앤', 0.9025605916976929),  
(('떨어지는건가요', 0.9005126953125),  
(('부양', 0.8990253806114197),  
(('대박난', 0.8988355398178101),  
(('떨어지니깐', 0.8980925679206848))]
```

```
model.wv.most_similar('한빛소프트')
```

```
(('칩스앤미디어', 0.9937265515327454),  
(('갈채', 0.9936501979827881),  
(('쿠팡', 0.9935990571975708),  
(('증가', 0.9928390979766846),  
(('전망에', 0.9926670789718628),  
(('삼반기', 0.9925553798675537),  
(('테마', 0.9925552010536194),  
(('지난해', 0.9924244284629822),  
(('소식에', 0.991966724395752),  
(('코로나', 0.9916753768920898))]
```

```
model.wv.most_similar('알체라')
```

```
(('한컴', 0.9874464273452759),  
(('제페토', 0.9817030429840088),  
(('etf', 0.9784061908721924),  
(('싸이월드', 0.9758068919181824),  
(('북허', 0.9723270535469055),  
(('윤석열', 0.9716570377349854),  
(('아이브', 0.9710423946380615),  
(('홀딩스', 0.9707745313644409),  
(('PE', 0.9703624248504639),  
(('개발중인', 0.9702233076095581))]
```

```
model.wv.most_similar('나무가')
```

```
(('이벤', 0.9915313720703125),  
(('어닉서프라이즈', 0.991205096244812),  
(('자이언트스텝', 0.989699125289917),  
(('스타트업', 0.9890900254249573),  
(('발표는', 0.9887019395828247),  
(('발표하고', 0.9886780977249146),  
(('억에', 0.988544762134552),  
(('궁금합니다', 0.9881877899169922),  
(('Q', 0.9876508712768555),  
(('엿다', 0.9873960018157959))]
```

Result

메타버스 테마주는
서로에 대해 자주 언급함

테마주는 커플링 종목임을 확인
(커플링 : 추세의 흐름이 비슷한 현상)

테마주 그룹은 서로 소통을 하며
감성적 교류를 통해
정보를 주고 받는다 해석함

주식 시장의 특성 상 공휴일엔 시장이 열리지 않음

주말에 뉴스는 계속해서 생성되지만 시장이 개장하지 않아 시계열 상 공백이 생김

-> 해당 공백을 직전 시점의 거래정보를 그대로 인용하는 것으로 대체

주식 거래량이 (매수 / 매도) 구분이 없이 제공됨

종목 별 거래량은 매수와 매도량이 더해진 값으로 정보가 제공됨

그리하여 해당 거래량 중 매수와 매도 중 어느 부분이 더 많이 이루어졌는지 해석이 모호함

-> 전일 대비 하락했다면 거래량을 음수화 / 상승했다면 양수화

거래량에 대한 우리만의 해석 Exam

1. 특정 종목 거래량 **폭발**
2. 해당 종목이 전일 대비 **하락**
3. 거래량은 **과매도**에 의한 것이라 해석
4. 거래량에 대한 **음수** 처리

Preprocessing Modification

감성 사전 구축 시 사용한 형태소 분류기 변경

기존 OKT 분류기에서 더욱 빠른 연산속도를 보여주는 Mecab으로 변경하여 사전 구축

전일 대비 상승 데이터: 0.5013026648801548
전일 대비 하락 데이터: 0.49869733511984515

<메타 버스>



워드 클라우드

공휴일에 대한 전처리 후 토큰 수 변화

Title

13,924 -> 14,564
약 500개 **증가**

Content

25,275 -> 27,256
약 2,000개 **증가**

Token Score

전일 대비 상승했다면
하락 비율을 토큰 Score에 **추가**

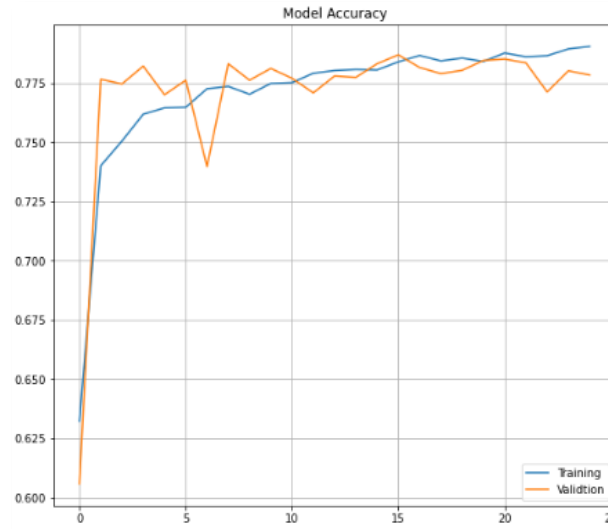
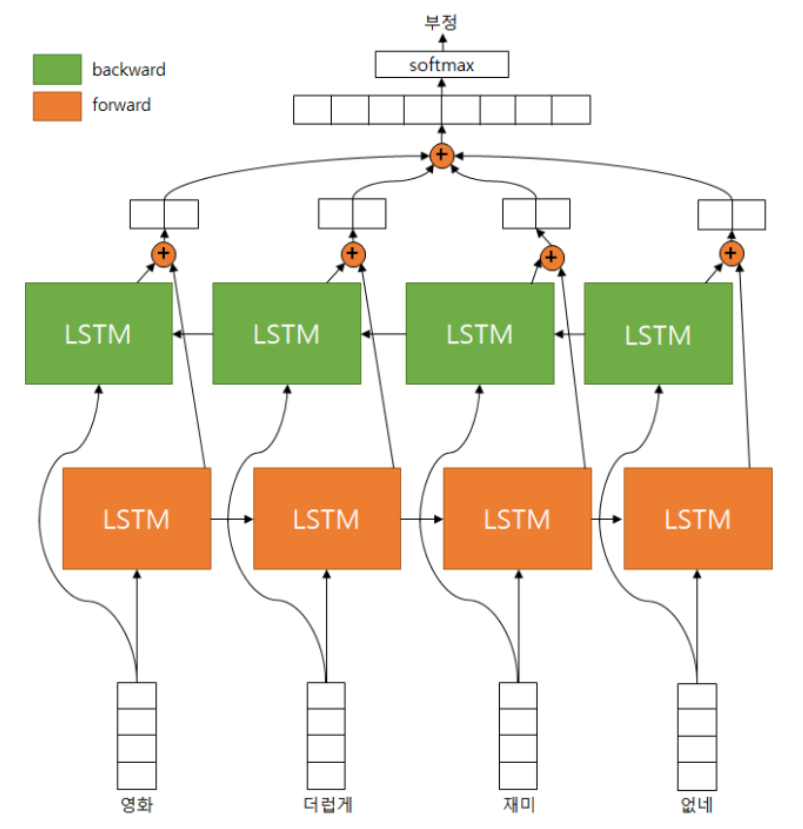
전일 대비 하락했다면
상승 비율을 토근 Score에서 차감

최근 시장이 하락세였기에
상승/하락 비율이 맞지 않음을 고려하여
나름의 정규화를 진행해준 셈

Modeling

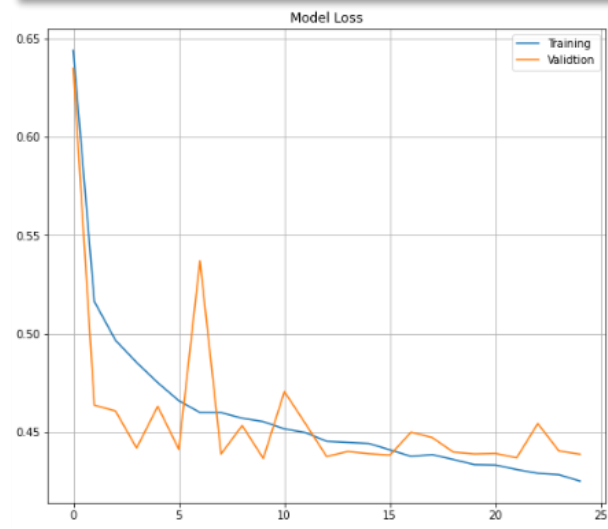
Bi-LSTM

사전 구축 모델로 Bi-LSTM을 선정하여 진행



개별 사전에 대한 정확도
(Exam : Naver)

Train 79.05%
Val 77.84%



Fear And Greed

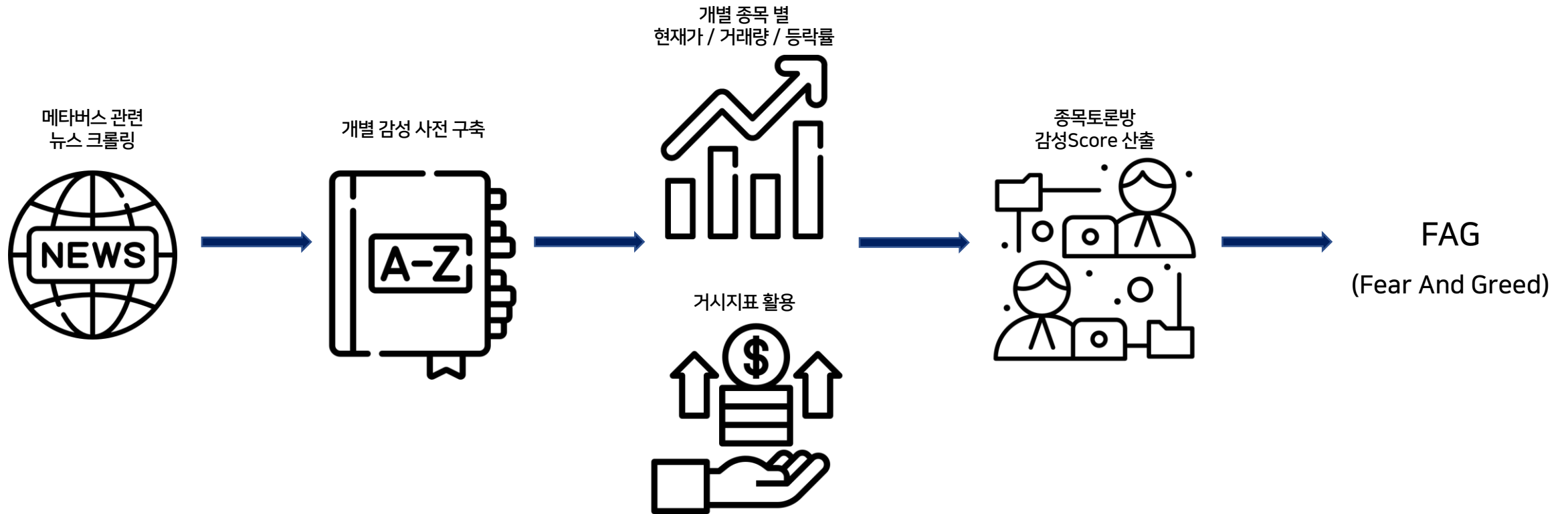
FAG (Fear And Greed)

- 감성 사전 기반으로 산출된 뉴스 감성 Score
- 종목 별 현재 가격 및 일일 거래 회전율, 등락률
- Kospi 200, S&P500 VIX, 환율(KRW / USD)

위 지표들을 결합하여 나온 Score에 종목 별 종목토론방 사람들의 감성 Score를 더해 최종 **FAG** 산출 예정



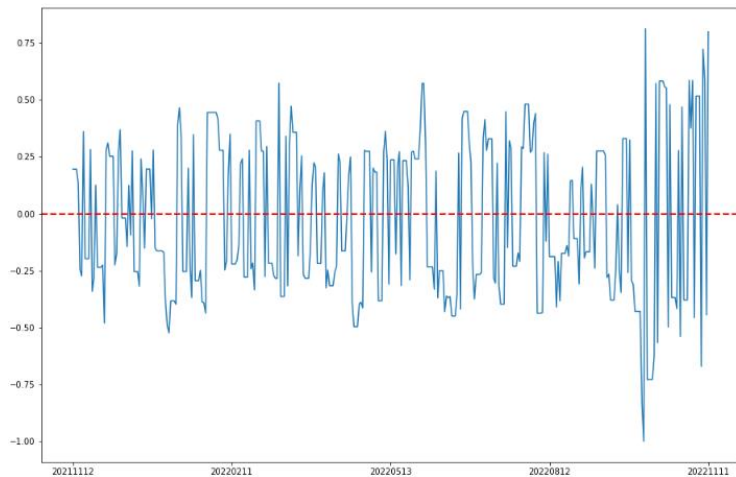
Fear And Greed Framework



Scaling Indicators

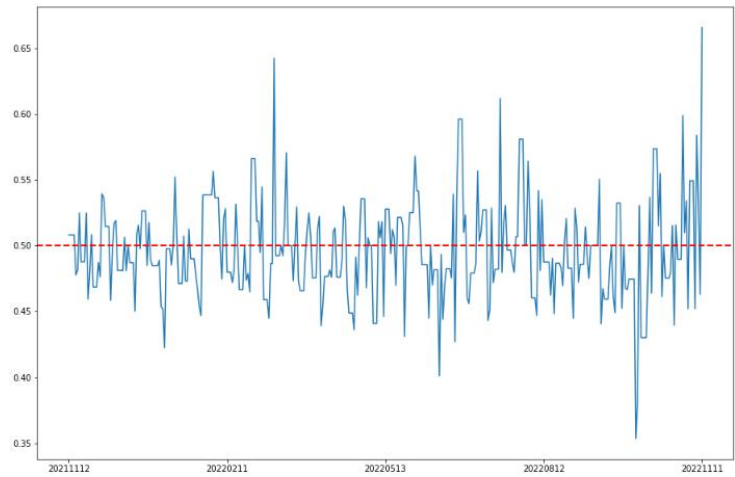
종목 별 일일거래량

거래량에 자연 로그 연산을 해준 뒤 min-max 스케일링 진행
등락여부에 따라 하락 : 음수 / 상승 : 양수 처리



종목 별 등락률

국내 시장은 상한가와 하안가에 $\pm 30\%$ 의 제한이 있음
(현재 등락률 - 하안 비율) / (상한 비율 - 하안 비율)로 스케일링



종목 별 현재가

스케일링의 기준을 52주 최저/최고가에서
최근 5년으로 범위 확대
(현재 가격 - 5년 최저가) / (5년 최고가 - 5년 최저가)로 스케일링



Scaling Indicators

Kospi

5년 최저-최고 스케일링의 값으로
스케일링 해준 뒤 지수에 반영



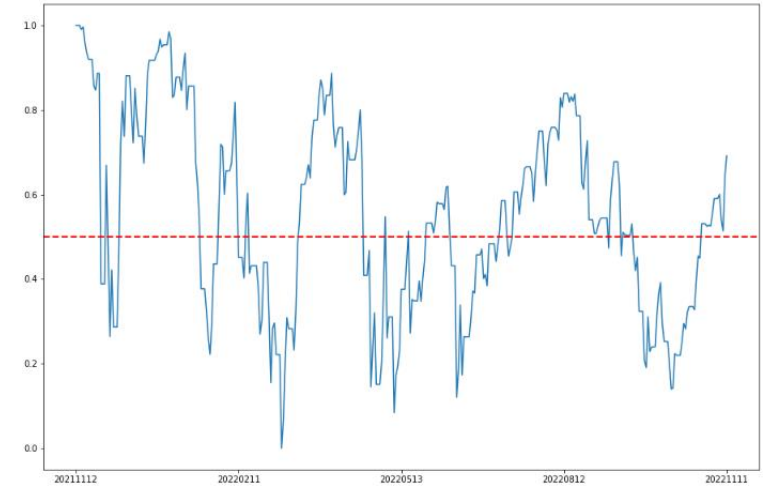
환율 KRW / USD

환율이 오르는 것은 시장의 공포감을 형성한다고 판단하여
1-(5년 최저-최고 스케일링의 값)으로 값을 역전하여 지수에 반영



S&P 500 VIX 변동률

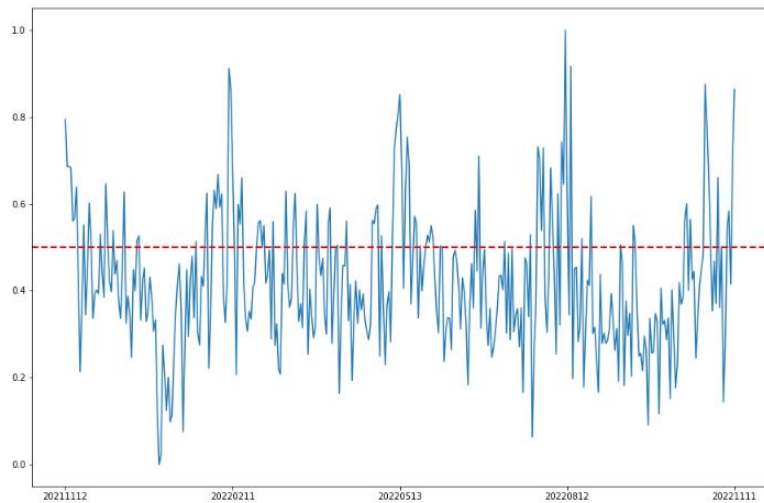
VIX지수는 상승할수록 공포감을 형성하므로
1-(min-max 스케일링의 값)으로 값을 역전하여 지수에 반영



Scaling Indicators

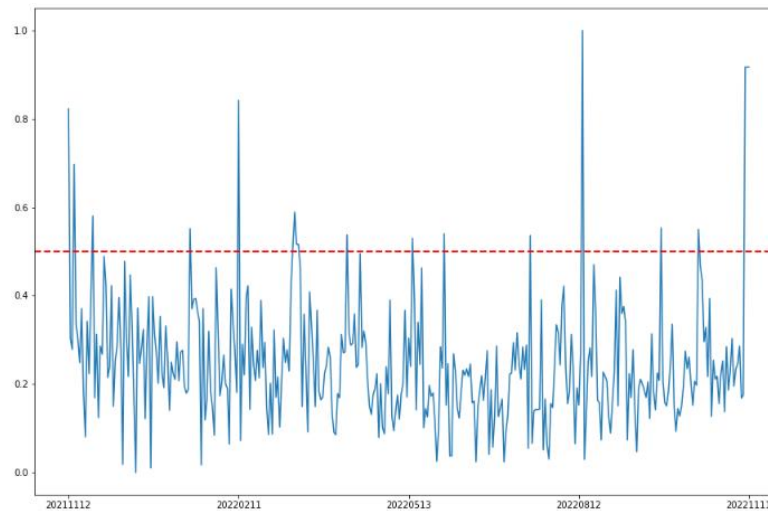
뉴스 감성 Score

min-max 스케일링을 해준 뒤 지수에 반영



종목토론방 감성 Score

min-max 스케일링을 해준 뒤 지수에 반영



Fear And Greed

Scaling Indicators의 Weight를 정한 뒤
최종 지표 산출

Indicators Weight

- 종토팡 Score 5%
 - 종목 별 현재 가격 / 일일거래량 / 등락률 10%
 - Kospi-200 / 환율 / VIX 15%
 - 뉴스 감성 Score 20%
- 종토팡의 경우 비속어와 어그로, 광고 등의 게시글이 존재하나 제일 솔직한 표현이 이루어지는 곳
 - Kospi와 환율은 금융권의 대표적 거시지표
 - VIX는 S&P500 시장의 시장변동성을 보여주는 지표
VIX 지수 상승 : 공포 상승
VIX 지수 하락 : 투심 완화
 - 현재 가격과 일일 거래량, 등락률은 종목에 대한 움직임을 보여주는 지표
 - 뉴스의 경우 정보에 대해 확실한 전달이 이뤄지는 전달자이자 정보에 대한 감성(호재 / 악재)을 확인하기 좋은 자료

Fear And Greed

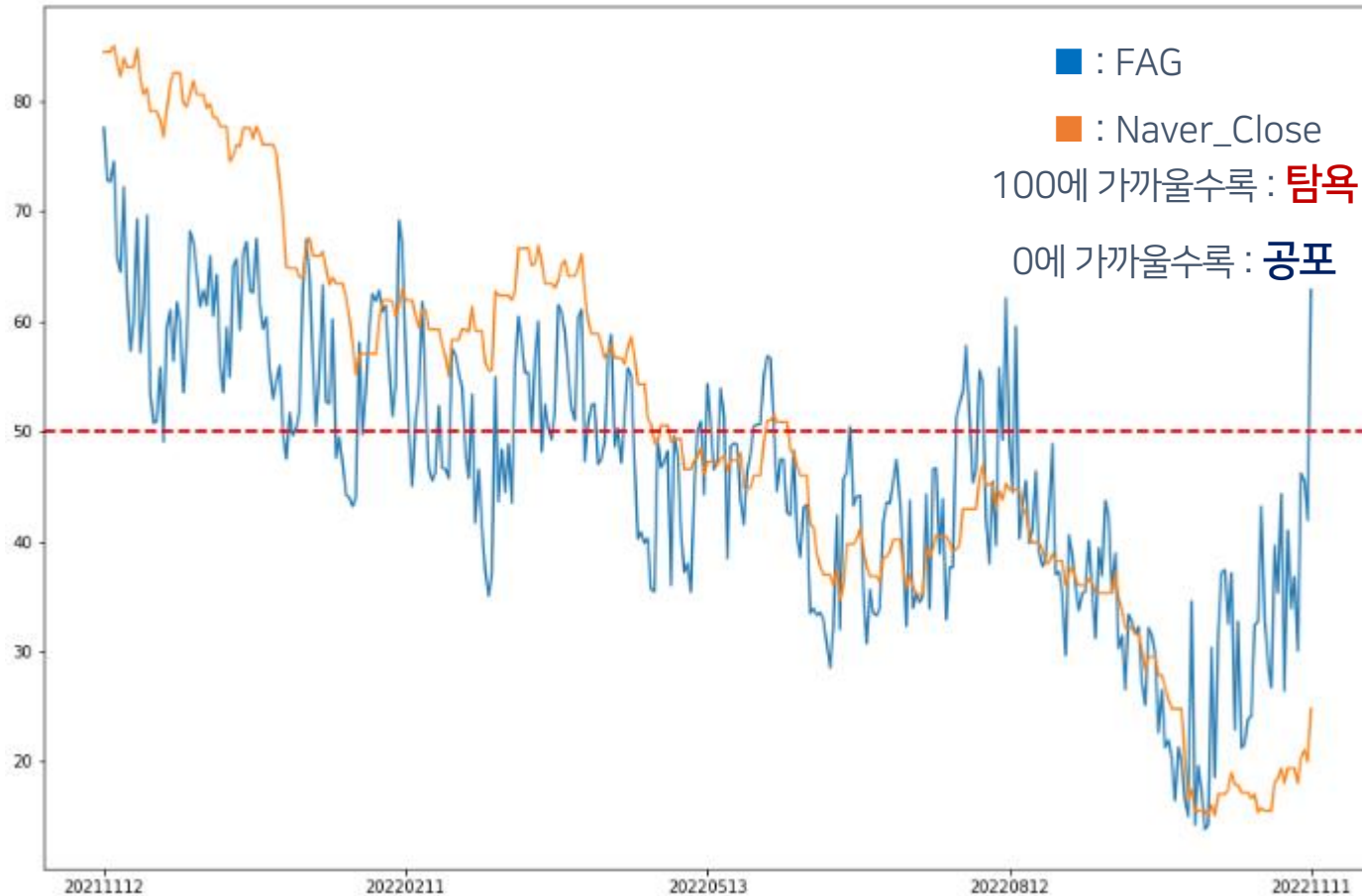
FAG(Fear And Greed) Plot



해석

코로나 거품을 반납한
최근 1년 시장이 하락세였던 것처럼
시장 분위기도 공포스러워졌다
10월 10일 이후 투자 심리가 완화됨

FAG(Fear And Greed) Plot With Naver_Close



Result

FAG 지표와 실제 네이버 종가의
그래프 동향이 비슷함을 확인

expectancy effects

종목 별 감성 Score가 투영된 FAG가
개별 종목 매매의 선행 지표로 고려됨을 기대함