

감정분석을 이용한 뉴스와 주가의 관계 분석

How to use

주제 선정

평소에 주식에 대한 관심이 많았다. 그러던 도중 사람들이 뉴스 기사가 주가에 영향을 미치기 때문에 기업이나 외인이 시세 조정에 이용한다는 주장을 보게 되었다. 그래서 실제로 뉴스가 주가에 미치는 영향이 있는지 궁금해졌다.

순서

분석은 다음과 같은 순서로 진행했다.

- 1. 데이터 수집
- 2. 뉴스 기사 제목 별 감정 분석
- 3. 뉴스와 주가간의 상관성 분석

1. 데이터 수집

1.1 뉴스 데이터 수집

뉴스 데이터 수집은 네이버페이 증권에 종목별 뉴스를 이용했다. 네이버 공식 API를 이용할 수도 있었지만 필요없는 내용을 제외하고 증권 뉴스만 가져오기 위해서 네이버페이 증권 뉴스를 이용했다.

LG화학 051910코스피 2023.12.01 기준(장마감) 실시간 기업개요

486,500
전일대비 ▼14,500 | -2.89%

전일 501,000 | 고가 501,000 (상한가 651,000) | 거래량 298,312
시가 499,500 | 저가 485,500 (하한가 351,000) | 거래대금 146,028 백만

투자정보

호가10단계

시가총액 34조 3,432억원
시가총액순위 9위
상장주식수 70,592,343
액면가 | 매매단위 5,000원 | 1주

외국인한도주식수(A) 70,592,343
외국인보유주식수(B) 30,578,293
외국인소진율(B/A) 43.32%

투자이전 | 목표주가 3.94배수 | 711,111
52주최고 | 최저 857,000 | 424,500

PER | EPS(2023.09) 29.23배 | 16,642원
추정PER | EPS 21.24배 | 22,901원
PBR | BPS (2023.09) 1.17배 | 416,754원
배당수익률 | 2022.12 2.06%

동일업종 PER 33.01배
동일업종 등락률 -2.19%

종합정보 | 시세 | 차트 | 투자자별 매매동향 | **뉴스·공시** | 종목분석 | 종목토론실 | 전자공시 | 공매도현황

종목뉴스 제목 내용

종목뉴스 안내

제목	정보제공	날짜
LG화학, 6일 ESG 컨퍼런스 개최	뉴스시스	2023.12.01 10:20
LG화학 '그린 커넥터' 활동 강화	헤럴드경제	2023.11.29 11:36
[충북소식] LG화학 청주공장, 해능보육원에 김장 김치 300kg 전달	연합뉴스	2023.11.28 15:42
LG화학, 대규모 재생에너지 확보 ... 넷제로 성큼	매일경제	2023.11.26 16:01
코스피, 개인 '사자'에 강보합세 ... LG화학 3%대 상승	한국경제	2023.11.24 09:18
LG화학 이종구 CTO 부사장으로 승진	헤럴드경제	2023.11.23 11:30
공공연구기관 기술이전, 5.9%↑ ...한양대, LG화학에 250억 특허 매...	아시아경제	2023.11.23 09:30
코스피, 美 증시 상승 속 강보합 출발 ... LG화학 2%대↑	이데일리	2023.11.23 09:14
LG화학, 이종구 CTO 부사장 승진 ...선학철 체제는 유지(종합)	연합뉴스	2023.11.22 19:02
선학철 LG화학 부회장 유임 ...LG엔솔 새 대표에 50대 김동명(종합)	뉴스1	2023.11.22 17:42

최근조회

MY STOCK

LG화학 486,500 ▼14,500

여기서 크롤링을 진행하면서 문제가 하나 발생했는데 일정 날짜가 넘어가게 되면 페이지 번호가 달라져도 보여지는 뉴스가 똑같다는 것이었다.

제목	정보제공	날짜	제목	정보제공	날짜
삼성, 삼성화재 안내건축교 30주년 기념식 개최	연합뉴스	2023.09.19 11:40	삼성, 삼성화재 안내건축교 30주년 기념식 개최	연합뉴스	2023.09.19 11:40
시각장애인-안내건의 행복한 동행...삼성 안내건사업 30주년 맞아	연합뉴스	2023.09.19 11:40	시각장애인-안내건의 행복한 동행...삼성 안내건사업 30주년 맞아	연합뉴스	2023.09.19 11:40
삼성자산운용, KODEX 반도체 ETF에 '삼성전자' 편입	조선비즈	2023.09.19 11:23	삼성자산운용, KODEX 반도체 ETF에 '삼성전자' 편입	조선비즈	2023.09.19 11:23
↳ 삼성자산운용, KODEX 반도체 ETF에 '삼성전자' 편입	뉴스1	2023.09.19 11:15	↳ 삼성자산운용, KODEX 반도체 ETF에 '삼성전자' 편입	뉴스1	2023.09.19 11:15
관련뉴스 5건 더보기 >			관련뉴스 5건 더보기 >		
삼성중, 디지털형해 나서...퀀오션 선택에 AI 플랫폼 탑재	매일경제	2023.09.19 11:09	삼성중, 디지털형해 나서...퀀오션 선택에 AI 플랫폼 탑재	매일경제	2023.09.19 11:09
삼성전자, KB국민카드-아멕스와 사물인터넷 협력	이코노미스트	2023.09.19 11:07	삼성전자, KB국민카드-아멕스와 사물인터넷 협력	이코노미스트	2023.09.19 11:07
↳ 갤럭시로 분실 카드 찾는다...삼성전자, KB국민카드-아멕스와 맞손	파이낸셜뉴스	2023.09.19 11:01	↳ 갤럭시로 분실 카드 찾는다...삼성전자, KB국민카드-아멕스와 맞손	파이낸셜뉴스	2023.09.19 11:01
관련뉴스 3건 더보기 >			관련뉴스 3건 더보기 >		
삼성전자, KB국민카드-아멕스와 '위치 확인 가능한 IoT카드' 협력	해럴드경제	2023.09.19 11:02	삼성전자, KB국민카드-아멕스와 '위치 확인 가능한 IoT카드' 협력	해럴드경제	2023.09.19 11:02
↳ 삼성전자, 국민카드-아멕스와 '위치확인 IoT카드' 개발	뉴스시	2023.09.19 11:01	↳ 삼성전자, 국민카드-아멕스와 '위치확인 IoT카드' 개발	뉴스시	2023.09.19 11:01
'위치 확인 가능' 사물인터넷 카드...삼성전자, MOU 체결	이데일리	2023.09.19 11:01	'위치 확인 가능' 사물인터넷 카드...삼성전자, MOU 체결	이데일리	2023.09.19 11:01
↳ 삼성전자, KB카드-아멕스와 '위치 확인 가능한 IoT카드' 협력	아시아경제	2023.09.19 11:00	↳ 삼성전자, KB카드-아멕스와 '위치 확인 가능한 IoT카드' 협력	아시아경제	2023.09.19 11:00
삼성전자, KB국민카드-아멕스와 '위치 확인 가능한 IoT카드' 협력	연합뉴스	2023.09.19 11:00	삼성전자, KB국민카드-아멕스와 '위치 확인 가능한 IoT카드' 협력	연합뉴스	2023.09.19 11:00
"삼성전자, 메모리 반도체 업황 회복 신호 나타나"...이전 9만전자 오	해럴드경제	2023.09.19 10:30	"삼성전자, 메모리 반도체 업황 회복 신호 나타나"...이전 9만전자 오	해럴드경제	2023.09.19 10:30
나... ↳ 유진금융 "삼성전자 3분기 영업이익, 시장 전망치 절반...1.6조 예상"	서울경제	2023.09.19 09:20	나... ↳ 유진금융 "삼성전자 3분기 영업이익, 시장 전망치 절반...1.6조 예상"	서울경제	2023.09.19 09:20
관련뉴스 3건 더보기 >			관련뉴스 3건 더보기 >		
키움증권 "삼성전자, 펀더멘털이 주가 뒷받침...9만 전자 된다"	조선비즈	2023.09.19 08:26	키움증권 "삼성전자, 펀더멘털이 주가 뒷받침...9만 전자 된다"	조선비즈	2023.09.19 08:26
↳ 삼성전자, 메모리 반도체 업황 회복 시그널 감지...키움	한국경제	2023.09.19 08:21	↳ 삼성전자, 메모리 반도체 업황 회복 시그널 감지...키움	한국경제	2023.09.19 08:21
관련뉴스 1건 더보기 >			관련뉴스 1건 더보기 >		
<< 맨앞 이전 491 492 493 494 495 496 497 498 499 500 다음 >> 맨뒤 >>			<< 맨앞 이전 491 492 493 494 495 496 497 498 499 500 다음 >> 맨뒤 >>		

이 부분은 네이버 증권 뉴스 자체의 문제였기 때문에 해결하지 못했고 결국에 가져올 수 있는 데이터 범위가 제한되게 되었다. 그래서 이번 분석에서는 데이터 수집 범위를 2023년 09월 21일 이후로 제한할 수 밖에 없었다.

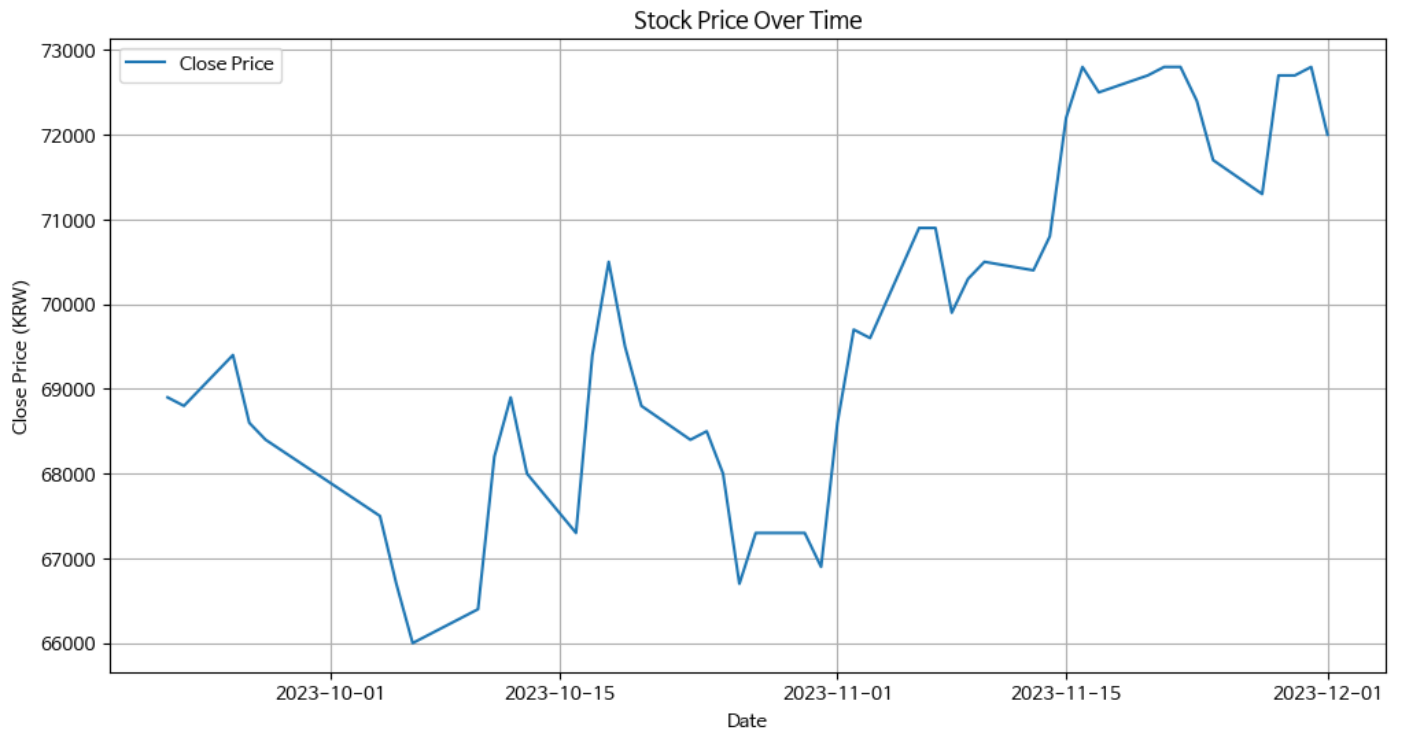
그리고 네이버 증권 뉴스에는 관련 뉴스도 제공해서 겹치는 뉴스를 제외하기 위해서 태그를 가져올 때 필터링 하기 위해서 아래와 같은 필터링을 수행했다.

```
for idx, row in enumerate(news_rows):
    # 관련 뉴스는 제외
    if idx == 0:
        continue
    if "class" in row.attrs and "relation_lst" in row["class"]:
        continue
    if row.find_parent("tr", class_="relation_lst"):
        continue
```

2. 뉴스 기사 감정 분석

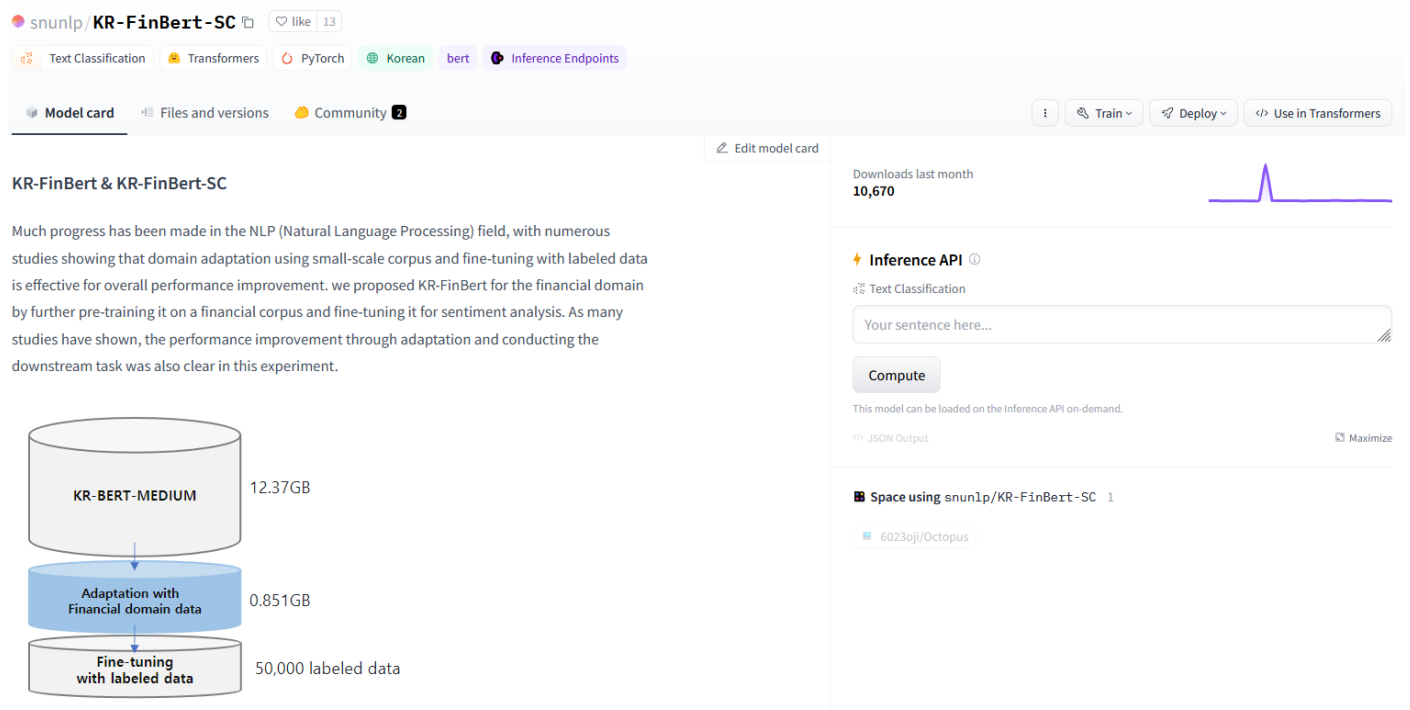
2.2 증권 주식 데이터 수집

증권 데이터 수집은 Yahoo Finance에서 제공하는 yfinace 파이썬 라이브러리를 사용하였다. 증권 데이터도 뉴스 데이터와 동일하게 2023년 09월 21일 이후로 데이터 수집 범위를 정했다.

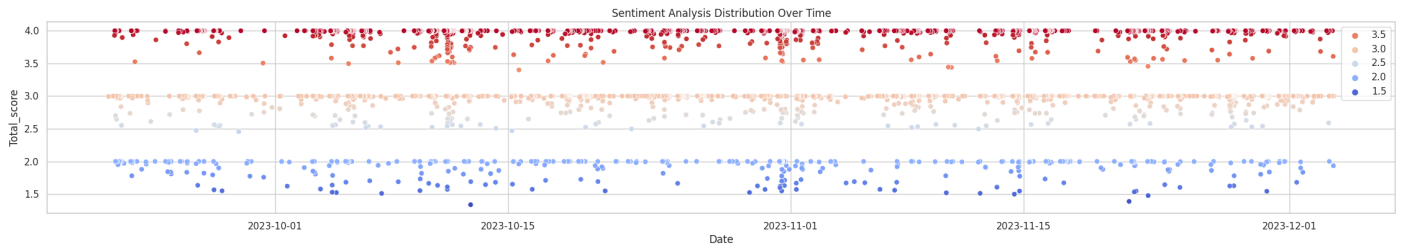


2.2 뉴스 기사 제목 별 감정 분석

금융 뉴스 기사의 감정을 분석하기 위해서 서울대학교 서울대학교 컴퓨터언어학_자연어처리 연구실에서 개발한 KR-FinBert-SC 금융 뉴스 분석 모델(<https://huggingface.co/snunlp/KR-FinBert-SC>) 을 사용하였다. 이 모델을 사용하게 된 이유는 먼저 뉴스 기사를 분석하는데 금융 분야에 특화된 모델이 필요했고, 두번째로 기존의 금융 뉴스 분석 모델을 있었지만 한국어 지원은 되지 않았기에 앞서 선택한 모델을 사용하게 되었다.



이후에 positive, natural, negative로 구분된 감점을 분석하기 위해 계량화 작업을 진행하였다. 계량화 한 날짜별 감정 분포도는 아래와 같이 나타났다.



3. 뉴스와 주가간의 상관성 분석

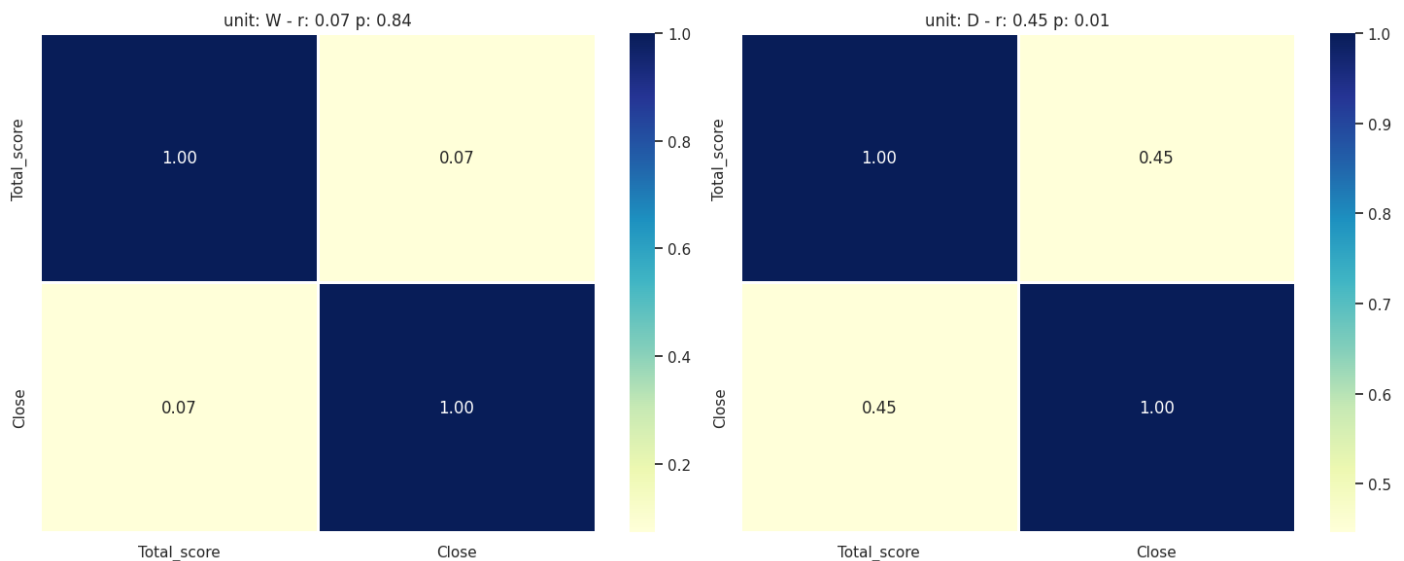
뉴스 데이터와 주식 데이터간의 상관관계 분석을 위해 다음과 같은 단계를 거쳤다.

1. 두 데이터들의 Date열을 datetime 형식으로 변환하고 두 datetime 형식을 같은 형식으로 맞추다음 합친다.
2. 주가 데이터를 log변환한후 차분하고 감정 점수와 같이 min-max-scaling 방식으로 정규화 한뒤 0.5를 빼서 음과 양을 표현할 수 있도록 했다.
3. 상관분표 표를 통해 일별, 주별 감성 점수와 주식 데이터간의 상관관계를 나타냈다.

아래의 표는 뉴스 기사에 대한 감성 점수와 각 기간(주, 일) 별 주가의 피어슨 상관 검정 결과값을 나타낸 것이다. r은 상관계수, p는 p-value를 의미한다.

분석한 결과를 봤을 때, 주별로는 0.7의 상관관계를 일별로는 0.45의 상관관계를 나타냈다. 상관계수는 보통 -1 ~ 1 사이의 범위를 가진다.

이 결과에서는 주별로는 봤을때는 상관관계가 높지는 않았지만, 일별로는 어느정도의 상관관계가 있는 것으로 보인다.



4. 결론

결과적으로 이번 분석에서는 뉴스와 주가간의 상관관계는 어느정도 있는 것으로 보인다.

그러나 일단 수집할 수 있었던 데이터의 범위가 제한적이었기 때문에 좀 더 과거에도 상관관계가 있었을지 아니면 미래에도 상관관계가 있을지는 미지수이다.

또 아쉬운 점으로, 장 마감 후에 나온 뉴스 기사는 당일 주가에는 영향을 미치지 못하지만 그 다음날의 주가에는 영향을 미칠 수 있기 때문에 이런 것들을 고려해서 시간별로 세분화 해서 분석할 수 있었다면 좀 더 좋은 결과가 나오지 않았을까 생각한다.

Reference

김유신, 김남규, 정승렬. (2012). 뉴스와 주가 : 빅데이터 감성분석을 통한 지능형 투자의사결정모형. 지능정보연구, 18(2), 143-156.

이예지. "뉴스 빅데이터 분석을 통한 종목별 주가예측." 국내석사학위논문 충북대학교, 2014. 충청북도

<https://songseungwon.tistory.com/124>

https://github.com/CSJ516/Stock_Prediction_project

<https://github.com/seyoongit/news>

https://github.com/park-gb/financial-news-sentiment-classifier/blob/main/sentiment_classifier.ipynb