

2023 Big Contest

[



]

“ 어린이를 위한 주식교육 프로그램? ”

뉴스 데이터와 KOSPI 정보를
분석 및 이용하여
구현된 프로그램

흥미유발 및 지속적인
금융교육이 가능하도록 함

게임 형식으로
어린이들에게 접근하는
교육 프로그램

자연스럽게
주식 및 경제지식을
습득하여 금융능력을 함양함



주식탐험대 뉴스와 주가등락 예측



뉴스와 주가의 상관관계

- 금융 시장은 정보의 흐름에 의존하므로 투자자들은 최신의 경제, 금융, 정치 뉴스를 파악하여 투자를 결정
- 뉴스키워드와 내용은 투자 심리에 영향을 미치고 주가에 반영되므로 주가 예측의 변수로 사용되기에 적합



뉴스를 이용한 주가예측

- 주식 시장은 정보의 흐름과 심리적 요소에 크게 영향을 받기 때문에 뉴스와 주가는 밀접한 관련성을 가짐
- 각각의 뉴스마다 긍정 또는 부정적인 감성이 담겨있으므로 텍스트를 분석하면 예측 정보로 활용가능함



뉴스와 KOSPI 정보를 이용

과거의 기업 KOSPI 등락 흐름과 뉴스를 감성분석한 결과를 활용하여 주식 예측 모델에 반영하면 주식등락 예측 정확성을 높일 수 있음

주식탐험대 분석수행절차

“ 데이터 수집 ”

주가데이터(KOSPI)

- yahoo finance 이용하여 특정기간 동안의 기업별 KOSPI 정보를 수집.
- 수집된 정보 중 종가(Close) 데이터만 사용함

회사(KS)	수집기간
LG Chem, Ltd. (051910.KS)	2020/1/1 ~ 2022/12/31
Hyundai Engineering & Construction Co., Ltd. (000720.KS)	2018/1/1 ~ 2021/12/31
Celltrion, Inc. (068270.KS)	2018/1/1 ~ 2021/12/31
Samsung Biologics Co.,Ltd. (207940.KS)	2020/1/1 ~ 2022/12/31
Ncsoft Corporation (036570.KS)	2018/1/1 ~ 2022/12/31
SK hynix Inc. (000660.KS)	2020/1/1 ~ 2022/12/31

뉴스데이터

- 통합데이터 빅카인즈 이용하여 기업별, 일자별로 뉴스를 수집
- 일자, 제목, 본문 내용 데이터를 사용함

일자	제목	키워드	특성추출 (가중치순 상위 50개)	본문
20221230	"돈이 얼마나 많길래" 직원에게 1000만 원 현금카드 뿌린 택진형	직원,1000만,현금카드,택진형,연봉,1000만,체크,카드,선물,온라인,커뮤니티,1..	엔씨소프트,1000만,김택진,그동안,상반기,리버티,택진형,100만,쓰론,현금카드,직...	[헤럴드경제=김현일 기자] "억대 연봉에 1000만원 체크카드 선물까지" \n \n...
20221230	직원 연봉 2억에 재택근무, 부러움 샀는데 "파티 끝났다"	2억,직원,연봉,재택근무,파티,회사,출근,회사,카카오,직원,2억,평균,연봉,육박,재...	재택근무,사무실,넥슨,직장인,근무제,놀이,코로나19,임직원,2억,1인,직원들...	[헤럴드경제=박세정 기자] "이제 익숙해졌는데 회사로 출근하러니, 회사 옮기고 싶다..."
20221230	테크기업들 재택 끝내고 몸집줄이기 판교에 불어닥친 '칼바람'	테크,기업,재택,몸집,판교,칼바람,네카오,채용,구조,보수,게임사,정리,재택,출근,내...	인건비,네이버,스타트업,구조조정,칼바람,내년도,판교,관계자,네트워킹,신사업,엔씨소프...	[아시아경제 최유리 기자] "내년도 사업계획서를 세 번째 수정 중인데 예산이 계속 ...
20221230	[기자수첩] '김택진 대표'가 이상하다는 창업자들	김택진,대표,창업자들,택진이형,사람,대표,얼마전,창업자,유명,기업,창업자,대표이사,...	창업자,대표이사,책임자,센터장,스타트업,창업자들,김택진,법법자,김범수,스마일게이트,...	[아시아경제 강나홀 기자] "택진이형"은 이상한 사람이예요 아직도 대표잖아요. ...
20221230	[폼아는기자들] 넥슨 오웬 대표 인터뷰 "노 템플릿(No Template)"	넥슨,오웬,대표,인터뷰,템플릿,Template,발행,스타트업,유료,뉴스레터,무료,가...	넥슨,일본,온라인,한국,닌텐도,미국,오웬,창업자,김정주,스타트업,ea,최전선,기자들...	@주3회 발행하는 유료 뉴스레터 [스타트업]입니다. 무료 가입은 https://pa...
...
20180101	5대 증권사가 선정한 황금개미에 투자 유망주는?	증권사,선정,황금개미,투자,유망주,시대,산업,혁명,로봇,인공지능,AI,자율주행,융성...	센터장,네이버,증권사,현대모비스,기대감,인공지능,클로바,포스코,미래에셋대우,유망주,...	4차 산업혁명 시대를 맞아 로봇 인공지능(AI) 자율주행 \n융성 인식 등의 신규 ...

“ 데이터 분석 ”

Okt 분석

- 문장을 형태소 단위로 분해하고 각각의 형태소에 품사를 태깅
- 중요한 정보를 담는 명사들만을 추출하여 사용한다.

TF-IDF

- 특정 단어가 얼마나 자주 나타나는지, 그렇지 않은지를 측정하여 각각의 단어가 가진 중요도를 계산
- 단어의 중요도가 계산되었으므로 이를 각각의 단어에 반영해 준다.

감성분석(Sentiment Analysis)

- 분해된 단어의 정보의 특징(feature)을 추출하여 텍스트가 긍정인지 부정인지 중립인지를 분류
- 분석을 효과적으로 할 수 있도록 텍스트를 수치화하기 위함이 목적이다.

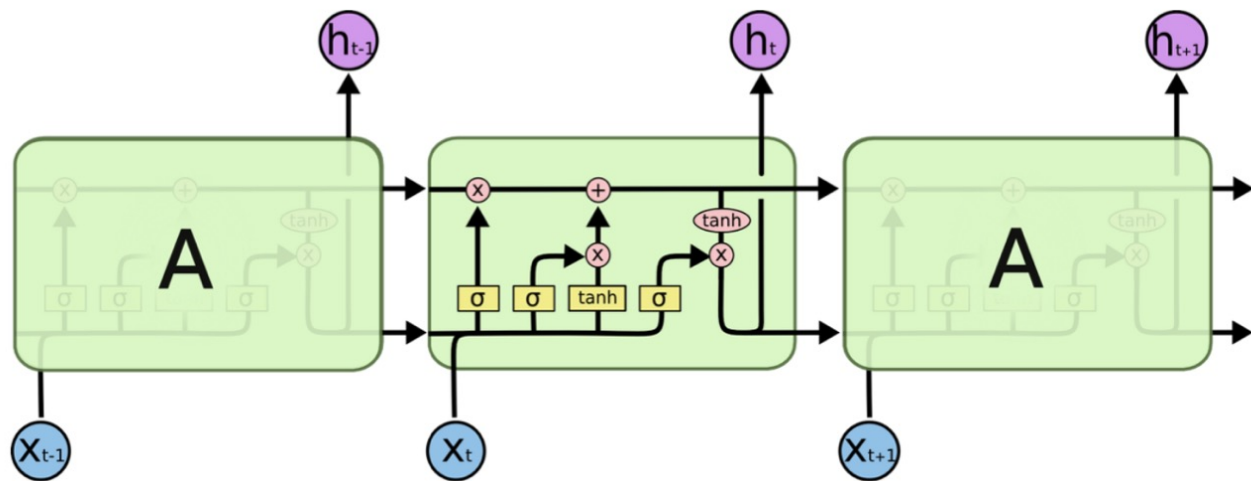
텍스트 요약(Summarization)

- 문제에 사용될 뉴스 기사를 만들기 위해 진행하며 추출적(extractive)요약을 사용하여 정보는 압축되지만 원본 손실이 발생하지 않도록 함

주식탐험대 분석수행절차

“ 모델링 (Modeling) ”

- 시계열 예측 (Timeseries forecasting)에 특화되어 있는 LSTM(장단기 메모리, Long Short-Term Memory) 모델을 미래 주가 등락 예측에 사용
- 아키텍처는 RNN(Recurrent Neural Network)의 한계인 기울기 소실 (Gradient Vanishing) 문제를 극복하고자 선택
- 시간적, 순차적 특성을 가지고 있는 주식 데이터의 특성을 고려하여 본 아키텍처를 모델링에 활용.



주식탐험대 분석수행절차

“ 모델학습(Learning) ”

- 'keras'의 LSTM 모델 패키지를 불러와 학습에 활용 ('Sequential' 모델 활용)
- 학습이 잘 되었는지 확인하기 위해 train data 80%, test data 20%로 나누어서 학습을 진행
- 회사별로 KOSPI와 관련된 뉴스 기사가 다르므로, 회사별로 나누어서 개별적으로 모델을 학습
- 각 회사마다 batch size와 epoch은 조절하면서 모델을 학습했고 출력된 정확도(Accuracy)

회사(KS)	accuracy	val_accuracy	Epoch
LG Chem, Ltd. (051910.KS)	0.6708	0.7042	20
Hyundai Engineering & Construction Co., Ltd. (000720.KS)	0.648	0.5677	30
Celltrion, Inc. (068270.KS)	0.6561	0.6474	30
Samsung Biologics Co.,Ltd. (207940.KS)	0.7074	0.6879	30
Ncsoft Corporation (036570.KS)	0.6284	0.6376	30
SK hynix Inc. (000660.KS)	0.712	0.7194	30

주식탐험대 분석수행절차

“ 주가등락 결과예측 ”

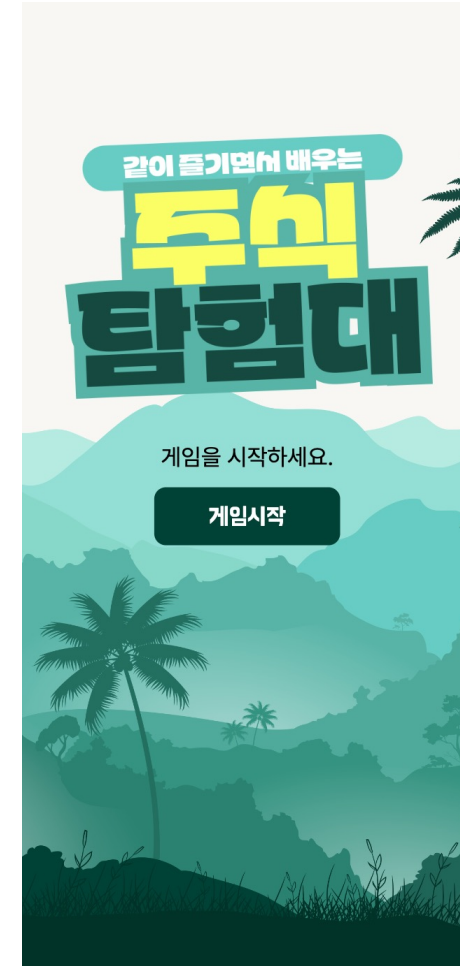
- 요약한 뉴스 본문에서 okt를 활용해 명사만 뽑아서 학습시 생성된 단어사전(vocab)에서 해당 단어를 찾아 tf-idf값을 곱해서 단어의 score를 구해 수치화
- 이 값을 학습된 모델의 input으로 넣고, 주가의 상승(1)/하락(0) 여부 예측을 테스트하며, 예측 된 값은 array([[1.28374006e-17]], dtype=float32) 형식으로 출력된다. 이 경우 0에 가까운 값이 나왔기 때문에 모델 예측 결과값은 '하락'이다.
- 이렇게 나온 결과를 가지고 실제로 문제에 활용하기 위해서 json 형식으로 변환하여 사용하였다.

```
{
  "id": 1,
  "headline": "대한민국, 아시아 백신 허브기지로 정착 예정",
  "lead": "앞서 글로벌 제약사들은 급증하는 코로나19 백신 수요를 감당하기 위해 안정적으로 백신을 생산할 수 있는 CMO 찾기에 집중해 왔다.삼성바이오로직스의 화이자 백신 연간 생산량은 최소 10억회분(5억명분) 이상으로 예상된다. 앞서 계약을 체결한 한국코러스는 연간 6억5000만도스로 알려져 있다.휴온스글로벌 관계자는 '러시아 백신이 생산되면 전세계 60여개국에 공급이 가능하고 대한민국이 글로벌 백신 허브기지로 정착하는 데 기여하게 될 것'이라고 밝혔다.",
  "invest": "바이오주",
  "answer": 0
},
```

```
{
  "id": 3,
  "headline": "곽노정 SK하이닉스 사장, 반도체산업협회장 선출",
  "lead": "한국반도체산업협회는 최근 서면으로 정기총회를 개최해 곽노정 SK하이닉스 안전개발제조총괄 사장(사진)을 제13대 협회장으로 선출했다고 7일 밝혔다. 협회장은 "회원사와의 교류 확대를 통해 업계의 다양한 요구를 수렴하고, 반도체협회가 이를 대변함으로써 '세계 1위 반도체 강국' 실현을 위한 전략 파트너로 거듭나겠다"고 밝혔다.",
  "invest": "IT주",
  "answer": 0
},
```

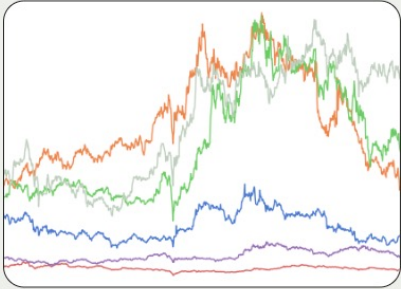

주식탐험대 뉴스기반 주식 공부 프로그램

“ 교육 프로그램 구현 ”



주식탐험대 뉴스기반 주식 공부 프로그램

주식탐험대



NEWS FEED

이슈되고 있는 뉴스기사를 안내합니다.

LG화학, 배터리 사업부 분사 수면 아래로... '실익 없다' 판단

옥죄는 中에 출렁이는 韓 증시...게임·엔터주는 관심을까

셀트리온, 코로나19 진단키트 7월 출시 예정

REMEMBER TIP

모르는 단어에는
힌트를
사용해보세요.

명예의 전당에
함께하세요!

퀴즈풀기

Q1.

화학주

Hint

LG화학, 첫 매출 30조 돌파... LGD, 14분기만 최대 영업이익

LG화학이 창사 이후 처음으로 매출 30조원을 돌파했다. LG화학은 지난해 연간 매출이 30조575억원, 영업이익이 2조3532억원이라고 공시했다. 4분기 매출은 전년 동기 대비 19.9% 증가했다. 영업이익은 전년 동기 대비 흑자 전환에 성공했다. LG화학 관계자는 '전기차 판매 확대 영향으로 매출 성장 기조가 유지될 것으로 전망한다'며 '연간 조 단위 영업이익의 창출을 목표로 하고 있다'고 밝혔다.

화학주가 상승한다.

화학주가 하락한다.

다음 문제 >

Q3.

화학주

Hint

LG화학이 전기차용 배터리 사업을 분사하는 논의를 사실상 중단했다. 당분간 분사에 우호적인 환경이 조성되기 어렵다는 게 대체적인 관측이다. LG화학은 지난해 하반기부터 사내 일각에서 검토하던 전지사업본부 분사 검토를 중단하고 당분간 현 체제를 유지하는 쪽으로 선회했다. 업계에서는 영업손실을 기록하고 있는 LG화학이 배터리 사업 분사를 중단한 결정을 '예견된 수순'으로 받아들이고 있다.

분사하다 : 해당하는 사업 분야를 독립시켜 여러 개로 쪼갬다

사내 : 회사 내

화학주가 상승한다.

화학주가 하락한다.

다음 문제 >

Q5.

화학주

Hint

우크라 전쟁 공포 증시 뒤흔들었다...무려 145개 종목이 신저가 경신

우크라이나를 둘러싼 전운이 최고조에 달하며 증시가 패닉에 빠졌다. 코스피는 34개 종목이 신저가를 경신했다. 푸틴 러시아 대통령이 특별 군사작전을 선포했다는 소식이 들려오면서다. 푸틴 대통령은 이러한 러시아의 움직임에 외국이 간섭할 경우 즉각 보복할 것이라고 강조했다. 바이든 미국 대통령은 우크라이나 침공에 대해 '정당한 이유가 없는 공격'이라고 비판의 목소리를 높였다.

화학주가 상승한다.

화학주가 하락한다.

다음 문제 >

주식탐험대 뉴스기반 주식 공부 프로그램

퀴즈 점수

SCORE: 3000

명예의 전당

1위	엘	8000점
2위	테스터2	7000점
3위	주식왕	7000점
4위	대원왕랄프	7000점

틀린 문제 확인하기

새 게임 시작

다시 체크해봅시다.

2번 문제 옥죄는 중에 출렁이는 韓 증시...게임
·엔터주는 관심을까

중국의 규제 강화에 국내 증시가 출렁이고 있다. 특히 중국 매출 비중이 컸던 게임·엔터테인먼트·미디어 업종을 향한 우려가 크다. 중국은 청소년이 평일에 야에 게임을 할 수 없도록 막고 금요일 또는 공휴일 하루 1시간만 게임이 가능하도록 했다. 전문가들은 중국 사업부 매출 비중에 따라 국내 게임 업체의 충격 정도가 다를 것이라고 진단했다.

정답: IT주 상승

3번 문제 LG화학, 배터리 사업부 분사 수면 아래로... '실익 없다' 판단

LG화학이 전기차용 배터리 사업을 분사하는 논의를 사실상 중단했다. 당분간 분사에 우호적인 환경이 조성되기 어렵다는 게 대체적인 관측이다. LG화학은 지난해 하반기부터 사내 일각에서 검토하던 전지사업본부 분사 검토를 중단하고 당분간 현 체제를 유지하는 쪽으로 선회했다. 업계에서는 영업손실을 기록하고 있는 LG화학이 배터리 사업 분사를 중단한 결정을 '예견된 수순'으로 받아들이고 있다.

분석 결과를 이용해서 'django'로 프로그램 제작

- 메인: 랜덤으로 문제 중 10개를 선택
- 뉴스피드: 랜덤문제 10개 중 임의로 3개만 선택해서 출력
- 퀴즈: 뉴스 기사를 읽고 주가 하락 or 상승을 예측
- 힌트: 힌트버튼을 누르면 어려운 경제용어의 뜻풀이가 출력
- 퀴즈결과: 사용자가 맞춘 답을 점수로 변환하여 출력
- 오답노트: 틀린 문제는 한 번 더 확인할 수 있도록 제공