

03 PJT

관심 종목 토론 데이터 수집

| 공통 요구사항

- 크롤링을 통해 주식 종목의 토론 데이터를 수집합니다.
 - 토스증권에서 종목 토론 데이터를 크롤링합니다.
 - 크롤링 대상 사이트: <https://tossinvest.com/>
 - 크롤링 시 필요한 selenium 패키지의 메서드는 생성형 AI 등으로 검색해서 활용합니다.
- Django 프로젝트의 이름과 앱 이름은 자유롭게 구성합니다.
- .gitignore 파일을 추가하여 불필요한 파일 및 폴더는 제출하지 않도록 합니다.
- 명시된 요구사항 이외에는 자유롭게 작성해도 무관합니다.

| URL

- `crawlings` 앱은 다음 URL 요청에 맞는 역할을 가집니다.

URL 패턴	역할
<code>/crawlings/index/</code>	사용자의 입력을 받아 종목 토론 데이터를 검색 검색한 결과(토론 목록)를 출력
<code>/crawlings/delete_comment/</code>	댓글 목록에서 하나를 삭제

| View

- `crawlings` 앱은 다음 역할을 가지는 `view` 함수를 가집니다.

View Method	역할
<code>index</code>	사용자 입력을 받을 수 있도록 <code>stock_finder.html</code> 렌더링 사용자가 입력한 종목을 검색하여 데이터를 크롤링 크롤링된 데이터를 <code>stock_finder.html</code> 을 이용하여 화면에 출력
<code>delete_comment</code>	특정 토론 데이터 삭제

| Templates

- 사용 템플릿 목록

A. stock_finder.html

- 파일명은 자유롭게 구성 하셔도 좋습니다

| 요구사항

- A. 종목 입력 화면 구성하기
- B. 종목 데이터 크롤링하기
- C. 크롤링 데이터 출력하기
- D. 댓글 삭제하기
- E. [도전] OpenAI API를 활용한 종목 분석

A. 종목 입력 화면 구성하기

- /crawlings/index/ 로 접속 시 사용자로부터 회사 이름을 입력 받을 수 있도록 화면을 구성합니다.
- 요구사항 A에서는 화면만 구성합니다.

출력 결과 예시

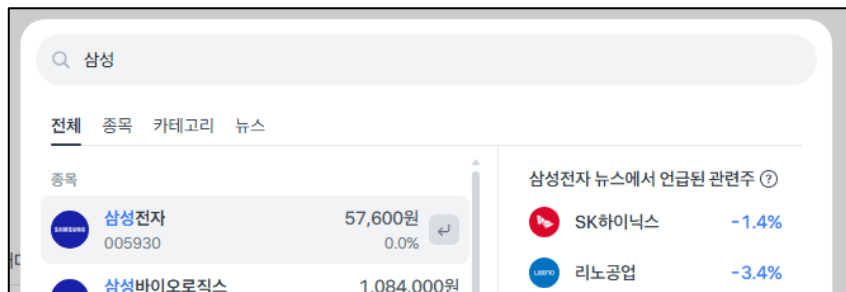
주식에 대한, 투자자들의 목소리

주식 댓글 분석기

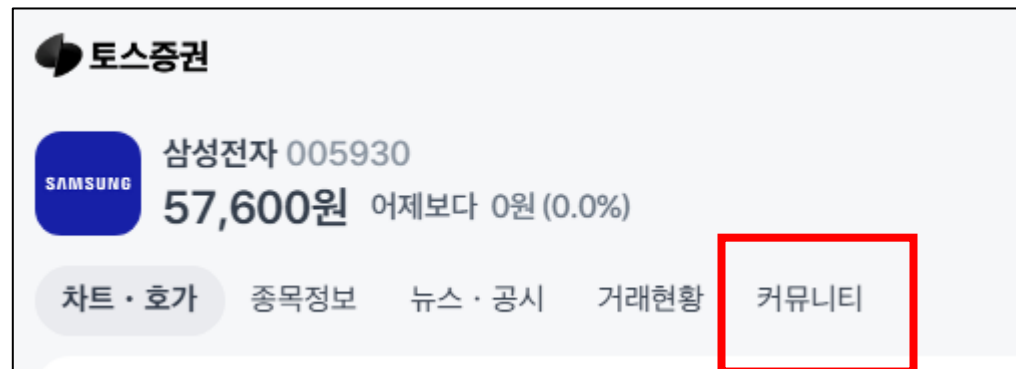
회사 이름을 입력하세요:

B. 종목 데이터 크롤링하기 (1/3)

- 요구사항 A에서 검색버튼 클릭 시 회사 종목명을 전달받아 크롤링을 수행합니다.
 - 같은 이름으로 시작하는 회사가 여러 개라면, 가장 상단의 회사를 검색합니다.
- selenium 을 활용하여 토스 증권에서 사용자가 입력한 종목의 토론 데이터를 수집해 옵니다.
 - 토스증권 사이트에서 검색을 직접 해보고 구조를 파악한 후 크롤링을 진행합니다.



삼성전자 검색



커뮤니티 탭 클릭

B. 종목 데이터 크롤링하기 (2/3)

- 커뮤니티 탭의 댓글 내용들을 수집합니다.
- 수집 후 데이터베이스에 저장할 수 있도록 구성합니다.
 - 수집 과정은 다음 페이지 참고
- 테이블 구조
 - 회사 이름 (text)
 - 종목 코드 (text)
 - 댓글 내용 (text)
 - 저장 일자 (datetime)



커뮤니티 탭 화면

| B. 종목 데이터 크롤링하기 (3/3)

- 크롤링 및 저장은 아래와 같은 순서로 진행합니다.
 - selenium 을 통해 토스 증권 메인 페이지를 열어줍니다.
 - selenium 을 통해 사용자가 입력한 종목을 검색합니다.
 - selenium 을 통해 커뮤니티 탭의 내용에서 댓글들을 수집합니다.
 - 수집한 댓글들을 저장합니다.
 - 댓글은 여러 개 입니다.
 - 하나의 종목에 대해 여러 개의 댓글을 어떻게 데이터베이스에 저장할 지 고민해야 합니다.

C. 크롤링 데이터 크롤링하기

- 저장한 데이터를 모두 출력합니다.

주식에 대한, 투자자들의 목소리

주식 댓글 분석기

회사 이름을 입력하세요:

삼성전자 (005930)

수집된 댓글:

- 삼성아 . 나 . 또 들어가게하지말고 올라가 !!!
- 내일 주주총회니까 버텨볼게?
- 젠슨황 믿는다?
- 여러분 다같이 영차영차 해보죠
- 영차영차
- 개미 99퍼가 털릴때까지 안올리는게
- 외인들 기를쓰고 매집하고 있는데....조금 남기겠다고 낮은 금액에
- 버거들 입속에 넣어주는 개미들.....;-;
- 외계인들 매수 멈췄는데
- 팔지도 않고 내일 어떻게 되려나

결과 화면 예시

D. 댓글 삭제하기

- 각 댓글들을 삭제할 수 있도록 버튼을 구성합니다.
- 삭제 버튼 클릭 시 DB에서 해당 댓글 내용을 삭제합니다.

주식에 대한, 투자자들의 목소리

주식 댓글 분석기

회사 이름을 입력하세요:

삼성전자 (005930)

수집된 댓글:

- 삼성아 . 나 . 또 들어가게하지말고 올라가 !!!
- 내일 주주총회니까 버텨볼게?
- 젠슨황 믿는다?
- 여러분 다같이 영차영차 해보죠
- 영차영차
- 개미 99퍼가 털릴때까지 안올리는게
- 외국인들 기를쓰고 매집하고 있는데....조금 남기겠다고 낮은 금액에
- 버거들 입속에 넣어주는 개미들.....;-;
- 외계인들 매수 멈췄는데
- 팔지도 않고 내일 어떻게 되려나

결과 화면 예시

도전 과제

[도전] OpenAI API 를 활용한 종목 분석

- OpenAI API 를 활용하여 댓글들의 전체 여론을 분석합니다.
- 여론에 대해 긍정, 부정, 중립 등으로 판단합니다.
 - 댓글들을 API 로 전달합니다.
 - 댓글 내용들을 분석할 수 있도록 프롬프트를 구성합니다.
 - 모델은 “gpt-4o-mini” 를 사용합니다.
- 페이지 하단에 분석 결과를 출력합니다.

주식에 대한, 투자자들의 목소리

주식 댓글 분석기

회사 이름을 입력하세요 · 삼성전자

검색

삼성전자 (005930)

수집된 댓글:

• 삼성아, 나, 또 들어가게하지말고 올라가 !!! 삭제

• 내일 주주총회니까 버티볼게? 삭제

• 켄슨형 믿는다? 삭제

• 여러분 다같이 영차영차 해보죠 삭제

• 영차영차 삭제

• 개미 99퍼가 탈릴때까지 안올리는게 삭제

• 외인들 기를쓰고 매집하고 있는데...조금 남기겠다고 낮은 금액에 삭제

• 버거들 입속에 넣어주는 개미들.....; 삭제

• 외계인들 매수 멈췄는데 삭제

• 팔지도 않고 내일 어떻게 되려나 삭제

'gpt-4o-mini' 분석:

이 댓글들은 삼성전자 주식에 대한 투자자들의 감정과 기대를 반영하고 있습니다. 다음과 같은 몇 가지 주요 포인트로 분석할 수 있습니다: 1. **투자의 불안정성**: "삼성아, 나, 또 들어가게하지말고 올라가 !!!"와 같은 요청은 투자자들이 주가 상승을 간절히 바라는 심리를 보여줍니다. 이는 주식 시장의 불안정함에 대한 우려를 나타내기도 합니다. 2. **존재하는 기대**: "내일 주주총회니까 버티볼게?"라는 댓글은 주주총회가 주가에 긍정적인 영향을 미칠 것이라는 기대를 내포하고 있습니다. 이는 특정 이벤트에 대한 기대감이 여전히 존재하고 있음을 시사합니다. 3. **시장 경쟁에 대한 의견**: "개미 99퍼가 탈릴때까지 안올리는게"와 "외인들 기를쓰고 매집하고 있는데..."와 같은 댓글은 소규모 투자자(개미)와 외국인 투자자 간의 힘의 불균형을 우려하는 목소리입니다. 이는 시장의 불공정성을 지적하며, 소규모 투자자들이 손해를 볼 가능성에 대한 우려를 나타냅니다. 4. **강정의 존재**: "영차영차" 같은 은유의 표현이 보이지만, 뒤따르는 부정적인 표현은 투자자들의 혼란과 고민을 반영합니다. 특히 외국인 투자자들의 매수 중단에 대한 언급은 시장에 대한 신뢰와 불안정성이 복합적으로 나타나는 지점입니다. 종합적으로, 삼성전자에 대한 댓글들은 투자자들의 간절한 희망과 함께 불안감이 혼재된 모습을 보여줍니다. 주주총회에 대한 기대감과 더불어, 대규모 투자자와 소규모 투자자 간의 불균형에 대한 우려가 지배적입니다. 이러한 요소들을 종합적으로 고려할 때, 현재 여론을 부정적으로 판단할 수 있습니다.

결과 화면 예시

제출

| 제출 시 주의사항

- 제출기한은 금일 18시까지입니다. 제출기한을 지켜 주시기 바랍니다.
- 반드시 README.md 파일에 단계별로 구현 과정 중 학습한 내용, 어려웠던 부분, 새로 배운 것들 및 느낀 점 등을 상세히 기록하여 제출합니다.
 - 단순히 완성된 코드만을 나열하지 않습니다.
- 위에 명시된 요구사항은 최소 조건이며, 추가 개발을 자유롭게 진행할 수 있습니다.
- <https://lab.ssafy.com/> 에 프로젝트를 생성하고 제출합니다.
 - 프로젝트 이름은 '프로젝트 번호 + pjt' 로 지정합니다. (ex. 03-pjt)
- 반드시 각 반 담당 교수님을 Maintainer 로 설정해야 합니다.