

노트북: 첫 번째 노트북

만든 날짜: 2020-10-27 오전 8:32

URL: <https://continuous-development.tistory.com/115?category=736681>

Python

[Python] python 에서 Seleium을 통한 동적 크롤링 - 4

2020. 10. 8. 17:11 수정 삭제 공개

일동제약 상세 페이지 크롤링 하기

```
In [1]: from selenium import webdriver
import time
path = './chromedriver.exe'
driver = webdriver.Chrome(path)
```

기본적인 함수를 넣고 확인해보고 페이지에 따라 이동하는 것을 구하기 위해서 아래와 같이 페이지가 바뀌게끔 url을 만든다.

ildong.com/kor/product/list.id?page=8&halt=&prdDisease=&prdCategory=5&searchVal=

Chrome이 자동화된 테스트 소프트웨어에 의해 제어되고 있습니다.

ILDONG 일동제약 기업 제품 투자 R&D 홍보 고객

제품검색

일동제약의 다양한 제품을 확인하실 수 있습니다

홈 > 제품 > 제품검색

제품검색

검색어 검색	제품명, 성분명, 효능/효과로 검색 가능 <input checked="" type="checkbox"/> 제품내 검색 <input type="checkbox"/> 성분 검색 <input type="checkbox"/> 효능/효과 검색			외형별 검색	약확정보원 바로가기		
일반의약품	전문의약품	의약외품	의료기기	건강기능식품	식품	기타	생신중단품목

```
In [4]: for page in range(1,10):
        url = 'https://www.ildong.com/kor/product/list.id?page='+str(page)+'&halt=&prdDisease=&prdCategory=5&searchVal=&searchOption=0'
        driver.get(url)
```

해당 페이지에 상세페이지로 이동하여 클릭하고 안에 있는 데이터를 긁어 오고 나오려 한다.


```
In [57]: details_list = []
for page in range(1,9):
    url = 'https://www.ildong.com/kor/product/list.id?page='+str(page)+'&halt=&prdDisease=&prdCategory=5&search='
    driver.get(url)
    prList = driver.find_element_by_class_name('prList')
    a_tag = prList.find_elements_by_tag_name('a') # 해당 페이지로 이동하기 위한 a tag를 찾는다.
    for idx in range(0,9):
        a_tag[idx].click() # 0이 a_tag를 index 순서대로 클릭하고
        detail = driver.find_element_by_class_name('detailCnt2') # 내가 필요로 하는 값을 가져온다.
        details_list.append(detail.text) # 그 다음 미리 만들어 놓은 리스트에 append로 추가해준다.
        driver.back() # 기존웹 클리어(초기화)한다.
        prList = driver.find_element_by_class_name('prList') # 그다음 위와 같이 다시 값을 가져오는 작업을 한다.
        a_tag = prList.find_elements_by_tag_name('a')
        time.sleep(2) # 2초 간격으로 진행한다.
```

이걸 진행하면 자동화 프로그램에서 깜빡거리면서 페이지가 바뀌고 크롤링하는 모습을 볼 수 있다.

Chrome이 자동화된 테스트 소프트웨어에 의해 제어되고 있습니다.

일동윌당스

일동제약

일동바이오사이언스

일동아미네크

일동에스텍스

일동아카데미스

윤리경영

채용

KOR

ILDONG 일동제약

기업

제품

투자

R&D

홍보

고객

제품검색

일동제약의 다양한 제품을 확인하실 수 있습니다

홈 > 제품 > 제품검색

제품검색

검색어 검색

제품명, 성분명, 효능/효과로 검색 가능

☒ 제품내 검색
☐ 성분 검색
☐ 효능/효과 검색

외형별 검색

약특정보원 바로가기

일반의약품

전문의약품

의약품

의료기기

건강기능식품

식품

기타

생산출단품목

건강기능식품

키드큐 면역젤리

키드큐 면역젤리

건강기능식품

키드큐 비타젤리

키드큐 비타젤리

건강기능식품

키드큐 아이언&C 유어블

키드큐 아이언&C 유어블

건강기능식품

키드큐 초어를 합시다

키드큐 초어를 합시다

'Python' 카테고리의 다른 글

[Python] Numpy를 통한 배열 indexing(Boolean indexing, fancy indexing)

[Python] Numpy에 대한 기초 정리와 사용법 정리

[Python] python 에서 Selenium을 통한 동적 크롤링 - 4

[Python] python 에서 Selenium을 통한 동적 크롤링 - 3

[Python] python 에서 Selenium을 통한 동적 크롤링 - 2

[Python] python 에서 Selenium을 통한 동적 크롤링 - 1

python Selenium

python Selenium 크롤링

python 동적 크롤링



나무늘보스

혼자 끄적끄적하는 블로그 입니다.