

# 파이썬(2) - 11주차 / 1905096(진태양)

## BeautifulSoup의 search API

- BeautifulSoup 라이브러리는 `.find_all` 과 `.select` 두 가지의 Element Search API를 제공한다.
- BeautifulSoup는 `.select`를 `.find_all`의 subset으로 간주하고 있다. 즉, 메인은 `.find_all`이고, 간혹 필요할 때 `.select`를 사용한다. 그렇기 때문에 `.find_all`은 생략하고도 사용할 수 있다.

```
# 아래 두 코드는 완전히 같은 동작을 보여준다.  
soup.find_all('div')  
soup('div')
```

- `.select`는 CSS Selector와 같이 `>`를 이용해 하위 태그로 접근할 수 있어, `.find_all`과 `.find`을 연속해서 사용하는 것 보다 `.select`로 한번에 검색하는 것이 간편하다.
- tag 안의 문자열 추출은 `.get_text()`, `.string`, `.contents` method를 이용한다.

## 관련 실습 코드

```
import urllib.request  
  
from bs4 import BeautifulSoup  
  
# 학과 홈페이지 공지사항 페이지 request 수행  
url = 'http://computer.ync.ac.kr/computer/CMS/Board/Board.do?mCode=MN045'  
html = urllib.request.urlopen(url).read()  
soup = BeautifulSoup(html, 'html.parser')  
  
# 결과 출력을 위한 url_prefix 사전 정의  
url_prefix = 'http://computer.ync.ac.kr/computer/CMS/Board/Board.do'  
  
# Class 명으로 Element를 가져온다. select method는 css selector 혹은 XPath로 Element  
# 에 접근 할 수 있게 해준다.  
titles = soup.select('.board-list-table > tbody > tr > td.subject > p.stitle >  
a')  
  
print('=====  
print('[영남이공대학교 컴퓨터정보과·소프트웨어콘텐츠계열 최신 공지사항 10건]')  
print('=====  
  
# titles 배열을 enumerate 객체로 변환하고 이를 순환하며 내용 출력 수행  
for idx, title in enumerate(titles):  
    subject = title.get_text().strip()  
    link = url_prefix + title.get('href', None)  
    res = '[%s] %s => %s' % ((idx + 1), subject, link)  
    print(res)
```

## 실행 결과

=====

[영남이공대학교 컴퓨터정보과·소프트웨어콘텐츠계열 최신 공지사항 10건]

=====

[1] 2020학년도 동계 계절학기 수강신청 안내 =>

[http://computer.ync.ac.kr/computer/CMS/Board/Board.do?  
mCode=MN045&mode=view&mgr\\_seq=34&board\\_seq=85576](http://computer.ync.ac.kr/computer/CMS/Board/Board.do?mCode=MN045&mode=view&mgr_seq=34&board_seq=85576)

[2] 아마존웹서비스 인턴쉽 모집 =>

[http://computer.ync.ac.kr/computer/CMS/Board/Board.do?  
mCode=MN045&mode=view&mgr\\_seq=34&board\\_seq=85470](http://computer.ync.ac.kr/computer/CMS/Board/Board.do?mCode=MN045&mode=view&mgr_seq=34&board_seq=85470)

[3] (주)제이에스시스템 신규 채용 =>

[http://computer.ync.ac.kr/computer/CMS/Board/Board.do?  
mCode=MN045&mode=view&mgr\\_seq=34&board\\_seq=85450](http://computer.ync.ac.kr/computer/CMS/Board/Board.do?mCode=MN045&mode=view&mgr_seq=34&board_seq=85450)

[4] (주)우리이노베이션 신규 채용 =>

[http://computer.ync.ac.kr/computer/CMS/Board/Board.do?  
mCode=MN045&mode=view&mgr\\_seq=34&board\\_seq=85424](http://computer.ync.ac.kr/computer/CMS/Board/Board.do?mCode=MN045&mode=view&mgr_seq=34&board_seq=85424)

[5] (주)한송네오텍 신규 채용 =>

[http://computer.ync.ac.kr/computer/CMS/Board/Board.do?  
mCode=MN045&mode=view&mgr\\_seq=34&board\\_seq=85416](http://computer.ync.ac.kr/computer/CMS/Board/Board.do?mCode=MN045&mode=view&mgr_seq=34&board_seq=85416)

[6] 대구시 시민해킹단 모집 =>

[http://computer.ync.ac.kr/computer/CMS/Board/Board.do?  
mCode=MN045&mode=view&mgr\\_seq=34&board\\_seq=85348](http://computer.ync.ac.kr/computer/CMS/Board/Board.do?mCode=MN045&mode=view&mgr_seq=34&board_seq=85348)

[7] (주)유니에버 신규 채용 =>

[http://computer.ync.ac.kr/computer/CMS/Board/Board.do?  
mCode=MN045&mode=view&mgr\\_seq=34&board\\_seq=85116](http://computer.ync.ac.kr/computer/CMS/Board/Board.do?mCode=MN045&mode=view&mgr_seq=34&board_seq=85116)

[8] (주)위니텍 신규 채용 => [http://computer.ync.ac.kr/computer/CMS/Board/Board.do?  
mCode=MN045&mode=view&mgr\\_seq=34&board\\_seq=85115](http://computer.ync.ac.kr/computer/CMS/Board/Board.do?mCode=MN045&mode=view&mgr_seq=34&board_seq=85115)

[9] (주)트론트 신규 채용 => [http://computer.ync.ac.kr/computer/CMS/Board/Board.do?  
mCode=MN045&mode=view&mgr\\_seq=34&board\\_seq=85102](http://computer.ync.ac.kr/computer/CMS/Board/Board.do?mCode=MN045&mode=view&mgr_seq=34&board_seq=85102)

[10] 제2회 2020 YNC 메이커톤 참가 안내 =>

[http://computer.ync.ac.kr/computer/CMS/Board/Board.do?  
mCode=MN045&mode=view&mgr\\_seq=34&board\\_seq=85019](http://computer.ync.ac.kr/computer/CMS/Board/Board.do?mCode=MN045&mode=view&mgr_seq=34&board_seq=85019)