# -\*- coding: utf-8 -\*-

import requests

from bs4 import BeautifulSoup

import urllib

import time

opener = urllib.request.build\_opener()

opener.addheaders = [('User-agent', 'Opera/9.80 (Android 2.3.4; Linux; Opera Mobi/build-1107180945; U; en-GB) Presto/2.8.149 Version/11.10')]

urllib.request.install\_opener(opener)

url='https://www.go100.com.tw/exam\_download\_3.php'

response = requests.get(url)

response.encoding = 'utf-8'

soup = BeautifulSoup(response.text, 'html.parser')

#print(soup)

all\_as = soup.find\_all('a')

for index, a\_tag in enumerate(all\_as):

if a\_tag.text == '試題內容':

print(a\_tag['href'])

if 'pdf' in a\_tag['href'] or 'docx' in a\_tag['href'] or 'doc' in a\_tag['href']:

urllib.request.urlretrieve(a\_tag['href'], './past\_paper/'+a\_tag['href'].split('/')[-1])

else:

urllib.request.urlretrieve(a\_tag['href'], './past\_paper/'+a\_tag['href'].split('/')[-1]+'.pdf')

time.sleep(1)