

KEMENTERIAN RISET TEKNOLOGI DAN PENDIDIKAN TINGGI
PROGRAM STUDI ILMU KOMPUTER DEPARTEMEN ILMU KOMPUTER DAN ELEKTRONIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS GADJAH MADA

TEMU KEMBALI INFORMASI

Tugas 3 Arsitektur dan Komponen Sistem Temu Kembali Informasi



DISUSUN OLEH:

ADAM YOGISYAH PUTRA	20/455439/PA/19654
MUHAMMAD ARSYA PUTRA	20/462186/PA/20158
HIZKYA FIRSTADIPA HARTOKO	20/455447/PA/19662

DOSEN:

Dr. Lukman Heryawan, S.T., M.T.

Metode Akuisisi Data Sistem Rekomendasi Obat Berdasarkan Gejala

Metode yang digunakan dalam akuisisi data pada sistem rekomendasi berdasarkan gejala adalah menggunakan metode web scraping. Web Scraping adalah proses ekstraksi data dan informasi dari suatu website untuk digunakan untuk keperluan tertentu. Web scraping dilakukan dengan bahasa pemrograman python dengan bantuan library beautiful soup. Data yang diambil merupakan data dari website Halodoc sebagai acuan rekomendasi data obat berdasarkan gejala. Langkah-langkah akuisisi data dengan web crawling adalah sebagai berikut.

1. Menginstall dan mengimport library yang diperlukan

```
import requests
from bs4 import BeautifulSoup
import csv
```

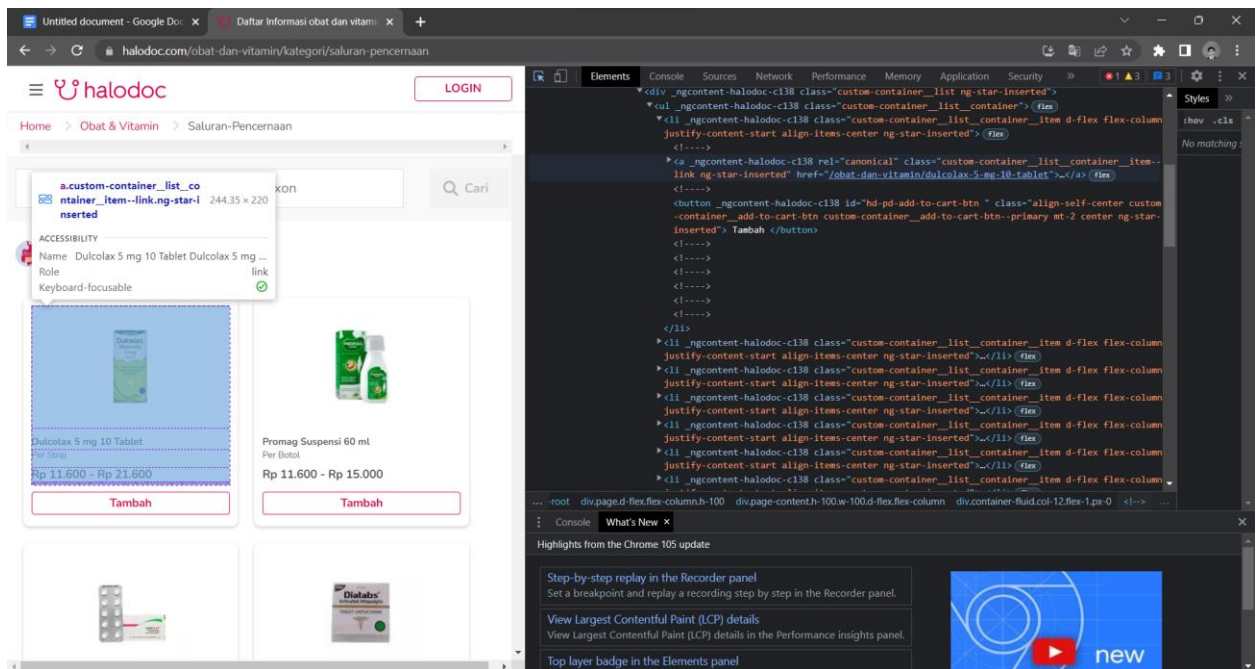
Kita perlu menginstall beautiful soup untuk parsing html dan digunakan sebagai web scraper, selain itu kita juga membutuhkan requests untuk melakukan HTTP request ke website yang diinginkan. Terakhir kita perlu import CSV agar dapat mengekspor data output dalam bentuk CSV

2. Menginisialisasi objek dari library yang diperlukan

```
req = requests.get(url, headers=headers)
soup = BeautifulSoup(req.text, "html.parser")
```

Instansiasi objek kelas beautiful soup agar kita dapat melakukan scraping dan parsing pada website yang dituju

3. Menelaah struktur website yang dituju



```
clas = "custom-container_list_container_item d-flex flex-column justify-content-start align-items-center ng-star-inserted"  
items = soup.findAll("li", clas)  
clas_a = "custom-container_list_container_item--link ng-star-inserted"
```

Menginspeksi halaman web yang ingin di scrape, pada proses ini kita perlu memperhatikan tag HTML yang berisi data yang kita inginkan. Kemudian kita juga perlu mengetahui request header yang digunakan pada website sehingga kita dapat mengambil data dari website tersebut

4. Melakukan request ke website yang dituju.

```
detail_url = "https://www.halodoc.com" + link.get("href")  
req_detail = requests.get(detail_url, headers=headers)  
soup_detail = BeautifulSoup(req_detail.text, "html.parser")
```

Selanjutnya kita melakukan request dengan menggunakan library request yang sudah kita import sebelumnya

5. Memberikan struktur pada data

```

desc = ind = dos = atur = '-'
for prod in product_detail:

    checkDesc = prod.find("div", "drug-list col-md-12").find('div','ttl-list').text
    if checkDesc.lower() == 'deskripsi':
        desc = prod.find('div', 'drug-detail col-md-12 margin-b-20 ng-star-inserted').find('div').text
    elif checkDesc.lower() == 'indikasi umum':
        ind = prod.find('div', 'drug-detail col-md-12 margin-b-20 ng-star-inserted').find('div').text
    elif checkDesc.lower() == 'dosis':
        dos = prod.find('div', 'drug-detail col-md-12 margin-b-20 ng-star-inserted').find('div').text
    elif checkDesc.lower() == 'aturan pakai':
        atur = prod.find('div', 'drug-detail col-md-12 margin-b-20 ng-star-inserted').find('div').text

    dict_data = {
        'product_name':isDefined(product_name),
        'category':isDefined( category),
        'deskripsi':desc,
        'indikasi_umum':ind,
        'dosis':dos,
        'aturan_pakai':atur,
    }

    array_data.append(dict_data)

```

Mengambil data yang diperlukan dan menyusunnya sedemikian sehingga kita dapat mudah melihatnya di kemudian waktu

6. Mengekspor data ke format yang diinginkan

```

try:

    filename = nama_kategori+'.csv'
    keys = array_data[0].keys()
    a_file = open(filename, "w")
    dict_writer = csv.DictWriter(a_file, keys)
    dict_writer.writeheader()
    dict_writer.writerows(array_data)
    a_file.close()
except IOError:
    print("I/O error")

```

```
crawl.py antiseptik.csv X
tki-scraping-obat-main > antiseptik.csv
1 product_name,category,deskripsi,indikasi_umum,dosis,aturan_pakai
2
3 Albothyl Konsentrate 5 ml,Antiseptik," ALBOTHYL KONSENTRATE mengandung Policresulen. Obat ini dapat digunakan sebagai hemostatik, antiseptik
4 Hemostatik lokal.
5 Pembersihan dan regenerasi jaringan pada luka bakar, luka, proses inflamasi kronik, lesi dekubitus, ulkus kruris, kondiloma akuminata, st
6 Vaginitis bakterial, kandidiasis vaginal trikomoniasis.
7 Iritasi atau inflamasi lokal dan kerusakan jaringan vagina atau serviks. "," PENGGUNAAN OBAT INI HARUS SESUAI DENGAN PETUNJUK DOKTER.
8 - Hemostasis lokal : Gunakan secara langsung.
9 - Luka bakar: Larutan dengan konsentrasi 1 : 3 sampai dengan 1 : 8.
10 - Luka bakar yg luas: Gunakan hanya larutan konsentrat saja atau lapisi daerah yg sakit dengan Albothyl gel atau lapisan kasa. ", Untuk p
11
12 Betadine Vaginal Douche 10% 100 ml Plus Dengan Alat,Antiseptik," BETADINE VAGINAL DOUCHE merupakan obat pembersih vaginal yang mengandung
13 Sebagai disinfektan Vagina. Digunakan untuk mengobati gejala iritasi ringan, gatal-gatal & keputihan yang disebabkan oleh Trichomonas vag
14 Terapi : 30-960 mL betadine dalam air hangat 1 kali sehari selama 6 hari.
15
16 Upaya pembersihan rutin : 15-960 mL betadine dalam air hangat 1-2 kali/minggu "," Isi botol plastik yang tersedia dengan larutan sampai b
17
18 Albothyl Konsentrate 10 ml,Antiseptik," Albothyl Konsentrate mengandung Policresulen. Obat ini dapat digunakan sebagai hemostatik, antiseptik
19 Hemostatik lokal.
20 Pembersihan dan regenerasi jaringan pada luka bakar, luka, proses inflamasi kronik, lesi dekubitus, ulkus kruris, kondiloma akuminata, st
21 Vaginitis bakterial, kandidiasis vaginal trikomoniasis.
22 Iritasi atau inflamasi lokal dan kerusakan jaringan vagina atau serviks. "," PENGGUNAAN OBAT INI HARUS SESUAI DENGAN PETUNJUK DOKTER.
23 Hemostasis lokal : Gunakan secara langsung.
24 Luka bakar: Larutan dengan konsentrasi 1 : 3 sampai dengan 1 : 8.
25 Luka bakar yg luas: Gunakan hanya larutan konsentrat saja atau lapisi daerah yang sakit dengan Albothyl gel atau lapisan kasa.
26 ", Untuk pemakaian luar.
27
28 Alkohol Swab Onemed 2 Ply 100 Lembar,Antiseptik," Alkohol Swab Onemed merupakan non woven tissue 2 ply steril yang mengandung Ethyl Alco
29
30 Betadine Skin Cleanser Antiseptic 60 ml,Antiseptik," BETADINE SKIN CLEANSER ANTISEPTIC merupakan antiseptik untuk kulit yang digunakan se
31 Mampu mengatasi ketombe dan jerawat
32 Disinfeksi tangan sebelum operasi. ", Gunakan secukupnya. , Tuangkan pada badan yang telah dibasahi dengan air.
33
34 Alkohol 70% 200 mL,Antiseptik, ALKOHOL 70% merupakan cairan antiseptik yang digunakan untuk
35
36 Alkohol 70% Ika Cool 100 ml,Antiseptik, ALKOHOL 70% IKA COOL merupakan cairan antiseptik ya
37
```

Terakhir kita dapat melakukan ekspor pada data yang telah kita scrape dan kemudian mengolahnya sesuai kebutuhan

Lampiran

Repository : <https://github.com/HizkiFirst/tki-scraping-obat>