

LAPORAN DATA SCRAPING

Disusun untuk Memenuhi Matakuliah Data Scraping
Dibimbing oleh Pelsri Ramadar N.S., M.Kom



Oleh:

FAVIAN NURRUDDIN

1123102138

**PROGRAM STUDI S1 TEKNIK INFORMATIKA
SEKOLAH TINGGI ILMU KOMPUTER PGRI BANYUWANGI
2024**

LAPORAN MATERI 10

1. Code app.py

Kode ini memiliki dua rute utama: satu untuk menampilkan halaman utama dan satu lagi untuk mengambil data populer dari situs *Bola.com*. Kode ini memanfaatkan library *requests* untuk melakukan pengambilan data dari URL eksternal, dan *BeautifulSoup* untuk memarsing data HTML.

a. Script / Setting Program

```
from flask import Flask, render_template
import requests
from bs4 import BeautifulSoup

app = Flask(__name__)

@app.route("/")
def Home():
    return render_template("index.html")

@app.route("/detik-popular")
def DetikPopular():
    htmlDoc = requests.get("https://www.bola.com/tag/bolacom")
    soup = BeautifulSoup(htmlDoc.text, "html.parser")
    popularArea = soup.find(attrs={'class': 'articles--iridescent-list'})

    images = popularArea.find_all(attrs={'class': 'articles--iridescent-list--item articles--iridescent-list-text-item'})

    return render_template("populer.html", gambar=images)

if __name__ == "__main__":
    app.run(debug=True)
```

b. Langkah Uji Coba

- **from flask import Flask, render_template:** Mengimpor modul Flask untuk membangun aplikasi berbasis web dan fungsi `render_template` untuk menampilkan file HTML.
- **import requests:** Mengimpor library *requests* untuk melakukan permintaan HTTP ke sebuah URL.
- **from bs4 import BeautifulSoup:** Mengimpor library *BeautifulSoup* untuk memanipulasi dan mengekstrak data dari dokumen HTML.

- **app = Flask(__name__):** Membuat objek Flask sebagai instansi utama untuk aplikasi web. Parameter `__name__` digunakan untuk konfigurasi.
- **@app.route("/"):** Menentukan rute utama (/) untuk aplikasi, yaitu sebagai halaman awal.
- **def Home(): return render_template("index.html"):** Fungsi Home menampilkan file HTML `index.html` sebagai respons untuk rute utama.
- **@app.route("/detik-popular"):** Menambahkan rute baru (/detik-popular) untuk menyajikan konten populer.
- **Def DetikPopular():**
htmlDoc = requests.get("https://www.bola.com/tag/bolacom"): Fungsi DetikPopular mengambil dokumen HTML dari URL yang ditentukan dengan menggunakan pustaka `requests`.
- **soup = BeautifulSoup(htmlDoc.text, "html.parser"):** Membuat objek BeautifulSoup untuk menganalisis (parsing) konten HTML yang diperoleh.
- **popularArea = soup.find(attrs={'class': 'articles--iridescent-list'}):** Mencari elemen HTML dengan atribut class tertentu (`articles--iridescent-list`) yang mengindikasikan area artikel populer.
- **images = popularArea.find_all(attrs={'class': 'articles--iridescent-list--item articles--iridescent-list--text-item'}):** Menemukan semua elemen dalam area populer dengan atribut class tertentu yang merepresentasikan daftar artikel berbentuk teks.
- **return render_template("populer.html", gambar=images) :** Mengembalikan respons berupa file `populer.html` yang dirender dengan data `images` sebagai parameter untuk ditampilkan.
- **if __name__ == "__main__": app.run(debug=True):** Menjalankan aplikasi Flask jika file ini dijalankan secara langsung, dengan mengaktifkan mode debug untuk mempermudah proses pengembangan.

2. Code index.html

a. Script / Setting Program

```
<div class="collapse navbar-collapse" id="navbarSupportedContent">
  <ul class="navbar-nav me-auto mb-2 mb-lg-0">
    <li class="nav-item">
      <a class="nav-link active" aria-current="page" href="/">Home</a>
    </li>
    <li class="nav-item">
      <a class="nav-link" href="/detik-popular">Berita Populer</a>
    </li>
  </ul>
</div>
```

b. Penjelasan Code

Pada index.html terdapat navbar yang berisi tombol link “Home” dan “Berita Populer” dengan penjelasan berikut:

- **Home**: Tautan "Home" yang aktif saat berada di halaman utama.
- **Berita Populer**: Tautan "Berita Populer" yang mengarah ke halaman daftar berita populer. Kode href yang mengarah ke rute “detik-popular” yang sudah didefinisikan di “app.py” untuk menjalankan Flask dan mengirim request ke situs www.bola.com, kemudian parsing HTML, mencari artikel populer yang sesuai class terkait, dan merender template populer.html.

3. Code populer.html

a. Script / Setting Program

```
<div class="container">
  <div class="row">
    <h2 class="mb-3">Link Berita</h2>
    {% for image in gambar %}
      <div class="col-md-4 mb-4">
        <div class="card">
          
          <div class="card-body">
            <h5 class="card-title">{{ image.find('a')['title']|safe }}</h5>
            <p class="card-text">{{ image.find('time')|safe }}</p>
            <a href="{{ image.find('a')['href']|safe }}" class="btn btn-primary">Baca
Selengkapnya</a>
          </div>
        </div>
      </div>
    </div>
  </div>
```

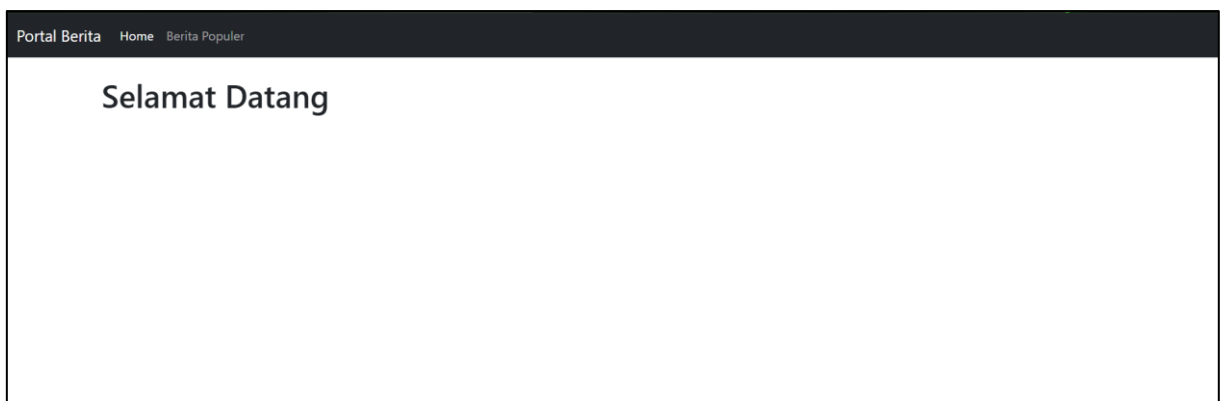
```
{% endfor %}  
</div>  
</div>
```

b. Penjelasan Code

- **{% for image in gambar %}** dan **{% endfor %}**: pernyataan loop yang digunakan untuk iterasi setiap elemen melalui list gambar yang sudah diambil dari Flask “app.py”.
- **{{ image.find('img')['src']|safe }}**: Mengambil sumber gambar dan menampilkan gambar berita dengan mencari tag dengan atribut “src”. “|safe” berfungsi untuk memastikan data yang diambil aman.
- **{{ image.find('a')['title']|safe }}**: Menampilkan judul berita yang diambil dari atribut title dari tag <a> dalam data image dan data yang diambil tetap aman.
- **{{ image.find('time')|safe }}**: Menampilkan tanggal atau waktu berita dengan mencari tag <time> di dalam image dan data yang diambil aman.
- **{{ image.find('a')['href']|safe }}**: Mengarah ke URL berita yang diambil dari atribut href tag <a> dalam data image.

Hasil Keseluruhan Code :

a. Index.html




b. Populer.html

Portal Berita

Home

Detik Populer


Link Berita



Timnas Indonesia Bermain Imbang di Kandang Sendiri saat Menghadapi Laos. Pengamat Vietnam: Kita Terlalu Tinggi Menilai Tim Garuda

13 Des 2024 18:15


Baca Selengkapnya



Pratama Arhan Pede Timnas Indonesia Mampu Bangkit di Piala AFF 2024: Singgung Mental Juara Skwad Garuda

13 Des 2024 18:00


Baca Selengkapnya



Kisah Shin Tae-yong yang Belum Pernah Menang atas Vietnam di Piala AFF

13 Des 2024 15:30


Baca Selengkapnya



50 Pemain Terhebat Real Madrid Sepanjang Masa: Cristiano Ronaldo Luar Biasa

13 Des 2024 14:15


Baca Selengkapnya



5 Partner Lionel Messi yang Paling Ganas: Alasan Kuat Tak Bisa Move on dari Barcelona

13 Des 2024 13:45


Baca Selengkapnya



6 Bintang Top yang Telat Bersinar: Bapuk Dulu, Lalu Gemilang hingga Pengujung Karier

13 Des 2024 13:00


Baca Selengkapnya



Mengulas Plus Minus Timnas Indonesia saat Ditahan Imbang Laos: Produktivitas Bek, Kesalahan Fatal, hingga Macetnya Skema Open-Play

13 Des 2024 12:30


Baca Selengkapnya



Piala AFF 2024: Pakai Pesawat Komersil, Timnas Indonesia Langsung Terbang ke Hanoi

13 Des 2024 12:00


Baca Selengkapnya



Mengulas Performa Pemain Senior Timnas Indonesia saat Ditahan Laos pada Piala AFF 2024: Pratama Arhan dan Ferri Bersinar

13 Des 2024 11:30


Baca Selengkapnya



10 Bapak-bapak Teman Cristiano Ronaldo pada Laga Debut di Manchester United, Ngapain Aja Sekarang?

13 Des 2024 11:00


Baca Selengkapnya



7 Pemain Enggak Jelas yang Diboyong Sir Alex Ferguson ke Manchester United, Bahkan Fotonya Pun Langka

13 Des 2024 10:45

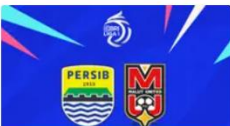
Baca Selengkapnya



Timnas Indonesia Memang Gagal Kalahkan Laos di Piala AFF 2024, tapi 3 Pemain Ini Performanya Tetap Keren

13 Des 2024 08:30


Baca Selengkapnya



Prediksi Persib Vs Malut United di BRI Liga 1: Lanjutkan Tren Positif

13 Des 2024 08:00


Baca Selengkapnya



Rapor Timnas Indonesia Vs Laos di Piala AFF 2024: Marselino Kali Ini Dapat Nilai Merah

13 Des 2024 07:15


Baca Selengkapnya



6 Pelatih Hebat yang Pernah Tolak Melatih MU: Tak Berminat atau Takut Tekanan Besar?

13 Des 2024 06:30


Baca Selengkapnya



Stadion Manahan Tetap Bergemuruh meski Timnas Indonesia Tak Mampu Kalahkan Laos pada Piala AFF 2024

13 Des 2024 06:00


Baca Selengkapnya



Timnas Indonesia Harus Move On setelah Dilmbangi Laos dan Fokus biar Dapat Poin dari Vietnam

13 Des 2024 05:15


Baca Selengkapnya



Mempertanyakan Lini Serang Timnas Indonesia saat Ditahan Laos: Malah Ferrari yang Dikawal Musuh

13 Des 2024 05:00


Baca Selengkapnya



Kandang-Tandang Jeda Cuma 3 Hari. Waktu Habis di Jalan, Shin Tae-yong Sebut Piala AFF 2024 Membunuh Pemain Timnas Indonesia

12 Des 2024 23:55


Baca Selengkapnya



STY Minta Bantuan Wartawan Agar Ivar Jenner Gabung Timnas Indonesia di Piala AFF 2024: Tolong Buatkan Tulisan..

12 Des 2024 23:48

Baca Selengkapnya



Piala AFF 2024: Timnas Indonesia Ditahan Laos, Shin Tae-yong Minta Maaf ke Seluruh Supporter

12 Des 2024 23:15

Baca Selengkapnya

PENUTUP

Kesimpulan

Keseluruhan kode ini membangun sebuah aplikasi web menggunakan Flask dengan Bootstrap untuk tampilan responsif. Server akan menyajikan halaman utama yang berisi tautan ke berbagai artikel populer yang diambil dari situs eksternal menggunakan BeautifulSoup. Artikel-artikel ini ditampilkan dalam bentuk kartu dengan gambar, judul, waktu, dan tautan untuk membaca lebih lanjut. Navbar memungkinkan navigasi antara halaman utama dan halaman artikel populer. Kode ini menggabungkan fitur web dinamis dengan tampilan yang responsif agar pengguna dapat dengan mudah mengakses konten berita melalui berbagai perangkat.

Mengetahui:

Dosen Pengampu Mata Kuliah

Pelsri Ramadar N.S., M.Kom

NIKP. 0704068408