



**ANALISIS HUBUNGAN POPULARITAS *SEIYUU* DENGAN
ANIME MENGGUNAKAN METODE PENGAMBILAN DATA
WEB SCRAPING PADA SITUS MYANIMELIST.NET
SKRIPSI**

**Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana
Komputer (S.Kom)**

NAMA: VENERDI RAFID MUHAJIR

NPM: 18360034

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI INFORMASI
INSTITUT SAINS DAN TEKNOLOGI NASIONAL
JAKARTA**

MARET 2022

HALAMAN PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya saya sendiri, dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar.

Nama : Venerdi Rafid Muhajir

NPM : 18360034

Tanggal : 02 Maret 2022

Jakarta, 02 Maret 2022

Venerdi Rafid Muhajir

HALAMAN PERNYATAAN NON PLAGIAT

Saya yang bertanda tangan di bawah ini :

Nama : Venerdi Rafid Muhajir

NPM : 18360034

Mahasiswa : Teknik Informatika

Tahun Akademik : 2018

Menyatakan bahwa saya tidak melakukan kegiatan plagiat dalam penulisan tugas akhir yang berjudul Analisis Hubungan Popularitas *Seiyuu* dengan Anime Menggunakan Metode Pengambilan Data *Web Scraping* pada Situs Myanimelist.net. Apabila suatu saat nanti terbukti saya melakukan plagiat, maka saya akan menerima sanksi yang telah ditetapkan. Demikian Surat Pernyataan ini saya buat dengan sebenar-benarnya.

Jakarta, 02 Maret 2022

Venerdi Rafid Muhajir

HALAMAN PENGESAHAN

Skripsi ini diajukan oleh :

Nama : Venerdi Rafid Muhajir

NPM : 18360034

Program Studi : Teknik Informatika

Judul Skripsi : Analisis Hubungan Popularitas *Seiyuu* dengan Anime Menggunakan Metode Pengambilan Data *Web Scraping* pada Situs Myanimelist.net.

Telah berhasil dipertahankan di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer pada Program Studi Teknik Informatika, Fakultas Sains dan Teknologi Informasi, Institut Sains Dan Teknologi Nasional

DEWAN PENGUJI

Pembimbing : Aryo Nur Utomo S.T, M,Kom ()

Penguji : Ir Andi Suprianto S.Kom, M.Kom ()

Penguji : B. Sumardiyono S.T, M.Kom ()

Penguji : Siti Nurmiati S.Kom, M.Kom ()

Ditetapkan di: Jakarta

Tanggal : 02 Maret 2022

KATA PENGANTAR

Alhamdulillah, puji syukur kehadiran Allah SWT yang telah memberikan rahmat dan inayah-Nya sehingga kami dapat menyelesaikan skripsi ini.

Skripsi ini saya buat sebagai tugas akhir saya di ISTN. Saya selaku penulis mengucapkan terima kasih kepada semua pihak yang telah membantu pembuatan skripsi ini.

Dalam penyusunan skripsi ini, tentu tak lepas dari pengarahan dan bimbingan dari berbagai pihak. Maka penulis ucapkan rasa hormat dan terima kasih kepada semua pihak yang telah membantu. Pihak-pihak yang terkait itu diantaranya sebagai berikut :

1. Ibu Neneng Rosmawarni, S.Si, M.Kom selaku Ketua Program Studi Teknik Informatika.
2. Ibu Aryo Nur Utomo S.Si, M.Kom selaku Dosen Pembimbing Skripsi;
3. Bapak Dadang Permana dan Ibu Balkis Erma selaku orang tua serta saudara – saudara penyusun yang mendukung dari awal, pertengahan, sampai akhir rangkaian penulisan skripsi ini terselesaikan;
4. Seluruh teman – teman HIMPUNAN MAHASISWA TEKNIK INFORMATIKA (HIMTIF) ISTN yang tidak dapat saya sebutkan satu persatu yang telah memberikan bantuan serta dukungan kepada penyusun selama ini.

Kami menyadari, bahwa skripsi yang kami buat ini masih jauh dari kata sempurna baik segi penyusunan, bahasa, maupun penulisannya. Oleh karena itu, kami sangat mengharapkan kritik dan saran yang membangun dari semua pembaca guna menjadi acuan agar penulis bisa menjadi lebih baik lagi di masa mendatang.

Depok, 02 Maret 2022

Penulis

Venerdi Rafid Muhajir

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika Institut Sains Dan Teknologi Nasional, saya yang bertanda tangan di bawah ini:

Nama : Venerdi Rafid Muhajir
NPM : 18360034
Program Studi : Teknik Informatika
Fakultas : Fakultas Sains dan Teknologi Informasi
Jenis Karya : Skripsi

demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Institut Sains dan Teknologi Nasional **Hak Bebas Royalti Noneksklusif (Non- exclusive Royalty- Free Right)** atas karya ilmiah saya yang berjudul :

“Analisis Hubungan Popularitas *Seiyuu* dengan Anime Menggunakan Metode Pengambilan Data *Web Scraping* pada Situs *Myanimelist.net*” beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Noneksklusif ini Institut Sains dan Teknologi Nasional berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (*database*) *soft copy* dan *hard copy*, merawat, dan mempublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta
Pada tanggal : 02 Maret 2022

Yang menyatakan

(Venerdi Rafid Muhajir)

ABSTRAK

Nama : Venerdi Rafid Muhajir

Program Studi : Teknik Informatika

Judul : Analisis Hubungan Popularitas *Seiyuu* dengan Anime Menggunakan Metode Pengambilan Data *Web Scraping* pada Situs Myanimelist.net.

Saat ini, volume data yang sangat besar dihasilkan setiap hari dengan kecepatan yang belum pernah terjadi sebelumnya dari berbagai sumber, termasuk data di internet. Keberadaan internet membuat jumlah data yang terdapat di website – website semakin bertambah, sehingga dibutuhkan metode yang efisien untuk mengumpulkan banyaknya data yang ada di suatu website, dan web scraping termasuk salah satunya. Pada penelitian ini, peneliti mengimplementasikan metode *web scraping* pada situs myanimelist.net yang merupakan komunitas dan database seputar anime dan manga terbesar di dunia. Tujuan dari penelitian ini adalah untuk menganalisis hubungan antara *seiyuu* (pengisi suara) dan anime yang diperankannya melalui analisis korelasi. Metode penelitian ini secara garis besar terdiri dari lima tahapan, meliputi *data extracting*, *data cleansing*, *data processing*, *statistic test*, dan *data visualization*. Alat penelitian yang digunakan untuk *scraping* adalah *beautifulsoup* dan *pandas* yang merupakan *library* python, sedangkan alat untuk proses dan analisis data hasil *scraping* adalah *google colab* dan *pandas* yang merupakan aplikasi pengolahan data. Hasil dari penelitian ini menunjukkan bahwa *web scraping* dapat diimplementasikan pada situs myanimelist.net dengan efektif. Peneliti dapat mengumpulkan 54298 baris data dari 275 orang dalam waktu 8 hari. Hasil analisis korelasi menunjukkan popularitas *seiyuu* memiliki hubungan sedang dengan popularitas karakter yang mereka perankan dan jumlah anime yang mereka mainkan.

Kata Kunci: *big data*, *web scraping*, *anime*

ABSTRACT

Name : Venerdi Rafid Muhajir
Study Program : Technical Information
Title : Analysis of the Relationship between Seiyuu Popularity and Anime Using the Web Scraping Data Retrieval Method on the Myanimelist.net Site

Today, enormous volumes of data are generated every day at an unprecedented speed from various sources, including data on the internet. The existence of the internet makes the amount of data contained on the website increase, so we need an efficient method to collect a lot of data on a website, and web scraping is one of them. In this study, the researchers implemented the web scraping method on the myanimelist.net site which is the largest community and database of anime and manga in the world. The purpose of this study is to analyze the relationship between the seiyuu (voice actors) and the anime they play through correlation. This research method broadly consists of five stages, including data extracting, data cleansing, data processing, statistical testing, and data visualization. The research tools used for scraping are beautifulsoup and pandas which are python libraries, while the tools for processing and analyzing scraping data are collaborative google and pandas which are data processing applications. The results of this study indicate that web scraping can be implemented on the myanimelist.net site effectively. Researchers were able to collect 54298 rows of data from 275 people within 8 days. The results of the analysis show that the popularity of seiyuu has a moderate relationship with the popularity of the characters they play and the number of anime they play.

Keyword: big data, web scraping, anime

DAFTAR ISI

HALAMAN PERNYATAAN ORISINALITAS.....	iii
HALAMAN PERNYATAAN NON PLAGIAT	iv
HALAMAN PENGESAHAN.....	v
KATA PENGANTAR	vi
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPERLUAN AKADEMIS.....	vii
ABSTRAK	vii
ABSTRACT	ix
DAFTAR ISI.....	x
DAFTAR GAMBAR	xii
DAFTAR TABEL	xiv
DAFTAR KODE.....	xv
DAFTAR LAMPIRAN.....	xvi
1. PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	2
1.3. Tujuan Penelitian	3
1.4. Manfaat penelitian.....	3
1.5. Batasan Masalah.....	3
2. TINJAUAN PUSTAKA	5
2.1. Web Scraping.....	5
2.2. Myanimelist.net.....	5
2.3. Anime	5
2.4. Seiyuu.....	6
2.5. Python	6
2.6. BeautifulSoup4	6
2.7. HTML	7
2.5. Pandas	7
2.6. Google Colaboratory	7
2.7. Analisi Korelasi.....	7
3. METODOLOGI PENELITIAN	9

3.1. Metodologi Penelitian	9
3.2. Alat Penelitian.....	14
3.3. Data Extracting.....	15
3.3.1. Scraping Halaman Daftar People	18
3.3.2. Scraping Halaman People	21
3.3.3. Scraping Halaman Anime	27
3.4. Data Cleansing (pembersihan data)	33
3.5. Data Processing (pemrosesan data).....	37
3.6. Statistic test (Uji Statistik)	40
3.6.1. Uji Korelasi	41
3.7. Data Visualization (Visualisasi Data)	42
3.7.1. Merancang aplikasi web Django	42
3.7.2. Membuat model database.....	48
3.3.3. Membuat konfigurasi URL dan tampilan	49
3.7.4. Membuat program visualisasi data.....	52
3.7.4.1. Halaman chart.html	53
3.3.4.2. Halaman tabel.html	56
4. HASIL DAN PEMBAHASAN	58
4.1. Hasil Uji Statistik	58
4.1.1. Hasil Uji Statistik Korelasi	58
4.2. Hasil Data Visualization (Visualisasi Data).....	61
5. PENUTUP.....	69
5.1. Kesimpulan	69
5.2. Saran.....	70
DAFTAR PUSTAKA	71
LAMPIRAN.....	73

DAFTAR GAMBAR

Gambar 3.1 Alur Penelitian.....	9
Gambar 3.2 alur proses penelitian.....	10
Gambar 3.3 Alur proses <i>web-scraping myanimelist.net</i>	17
Gambar 3.4 halaman <i>myanimelist.net/people.php</i>	18
Gambar 3.5 <i>Output scraping tag href</i>	20
Gambar 3.6 Tampilan halaman <i>people</i>	21
Gambar 3.7 Data yang diambil di halaman <i>seiyuu</i>	23
Gambar 3.8 <i>Output</i> nama karakter anime	26
Gambar 3.9 <i>Output</i> jumlah favorite karakter	27
Gambar 3.10 Data yang diambil di halaman anime	28
Gambar 3.11 <i>Output scraping</i> halaman anime	31
Gambar 3.12 Hasil <i>scraping</i> <i>myanimelist</i>	32
Gambar 3.13 <i>Source code</i> dan Tampilan <i>dataframe</i> di google collaboratory	34
Gambar 3.14 Mengurutkan nama sesuai banyak anime yang diperankan	36
Gambar 3.15 Menghapus 67 nama dari <i>dataframe</i>	37
Gambar 3.16 Tampilan <i>dataframe</i> setelah 67 nama dihapus	37
Gambar 3.17 Pengelompokan data berdasarkan nama <i>seiyuu</i> dan nama karakter	38
Gambar 3.18 hasil akhir pengelompokan data	39
Gambar 3.19 Tampilan akhir <i>dataframe</i> setelah diproses	40
Gambar 3.20 Rumus uji korelasi.....	41
Gambar 3.21 Pembuatan file <i>base.html</i>	44
Gambar 3.22 Proses migrasi model	49
Gambar 3.23 Proses migrasi model	49
Gambar 3.24 Tampilan folder <i>django</i>	50
Gambar 4.1 Hasil uji korelasi dengan <i>pandas</i>	58
Gambar 4.2 Tampilan data setelah penyesuaian	60
Gambar 4.3 Proses memasukkan data <i>csv</i> ke <i>database</i>	61
Gambar 4.4 Proses menjalankan server web.....	61

Gambar 4.5 <i>Bar Chart</i> dari data jumlah favorite <i>seiyuu</i>	62
Gambar 4.6 <i>Bar Chart</i> dari data jumlah favorite <i>seiyuu</i> dan jumlah favorite karakter	62
Gambar 4.7 <i>Scatter plot</i> dari data jumlah favorite <i>seiyuu</i> dan jumlah favorite karakter	63
Gambar 4.8 <i>Bar Chart</i> dari data jumlah favorite <i>seiyuu</i> dan jumlah member.....	64
Gambar 4.9 <i>Scatter plot</i> dari data jumlah favorite <i>seiyuu</i> dan jumlah anime yang diperankannya.....	64
Gambar 4.10 <i>Bar Chart</i> dari data jumlah favorite <i>seiyuu</i> dan rata- rata jumlah member.....	65
Gambar 4.11 <i>Scatter plot</i> dari data jumlah favorite <i>seiyuu</i> dan rata- rata jumlah member.....	65
Gambar 4.12 <i>Bar Chart</i> dari data jumlah favorite <i>seiyuu</i> dan rata - rata rating anime.....	66
Gambar 4.13 <i>Scatter Plot</i> dari data jumlah favorite <i>seiyuu</i> dan rata - rata rating anime	67
Gambar 4.14 Tabel keseluruhan data.....	67
Gambar 4.15 Halaman admin	68

DAFTAR TABEL

Tabel 3.1 Struktur proyek Django.....	42
Tabel 3.2 Struktur folder aplikasi django	43
Tabel 4.1 Hasil uji korelasi	59
Tabel 5.1 Rincian jumlah data setelah scraping dan pembersihan data.....	69

DAFTAR KODE

Kode 3.1 <i>Source code universal beautifulsoup</i>	15
Kode 3.2 <i>Rancangan dataframe pandas</i>	16
Kode 3.3 <i>Source code scraping tag href</i>	19
Kode 3.4 <i>Merubah source scraping</i>	20
Kode 3.5 <i>Source code pembeda seiyuu</i>	22
Kode 3.6 <i>Source code scraping halaman seiyuu</i>	23
Kode 3.7 <i>Source code scraping halaman anime</i>	29
Kode 3.8 <i>Source code pembuatan dataframe dan ekspor ke csv</i>	31
Kode 3.9 <i>Source code base.html</i>	44
Kode 3.10 <i>Source code model.py</i>	48
Kode 3.11 <i>Source code views.py</i>	50
Kode 3.12 <i>Source code urls.py</i>	52
Kode 3.13 <i>Source code bar chart</i>	53
Kode 3.14 <i>Source code scatter plot</i>	55
Kode 3.15 <i>Source code penampil bar chart dan scatter plot</i>	55
Kode 3.16 <i>Source code penampil bar chart dan scatter plot</i>	56

DAFTAR LAMPIRAN

Lampiran 1. Lembar Konsultasi Bimbingan Skripsi	73
Lampiran 2. Tabel daftar 275 orang yang diambil.....	74
Lampiran 3. <i>Source code myanimelistScrapper.py</i>	81
Lampiran 4. <i>Source code</i> lengkap aplikasi web visualisasi data	86