

DATA SCRAPING INFORMASI KULINER DARI MEDIA SOSIAL INSTAGRAM

Tugas Akhir

**Diajukan untuk Memenuhi Salah Satu Persyaratan Mencapai Derajat
Sarjana Teknik Informatika**



Dibuat Oleh :

MELITA DWI NUGROHO

15 07 08365

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS ATMA JAYA YOGYAKARTA
2019**

LEMBAR PENGESAHAN

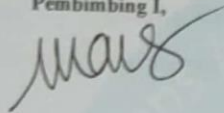
Data Scraping Informasi Kuliner dari Media Sosial Instagram

Yogyakarta, 08 Juli 2019

Melita Dwi Nugroho
15 07 08365

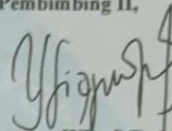
Menyetujui,

Pembimbing I,



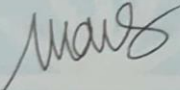
(Martinus Maslim, S.T., M.T.)
NPP: 01.13.847

Pembimbing II,



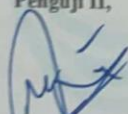
(Y. Sigit Purnomo WP., S.T., M.Kom.)
NPP: 04.01.710

Tim Penguji:
Penguji I



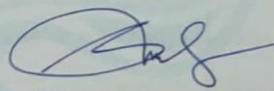
(Martinus Maslim, S.T., M.T.)
NPP: 01.13.847

Penguji II,



(Patricia Ardhanari, S.Si., MT.)
NPP: 03.94.512

Penguji III,



(Yulius Harjoseputro, ST., MT.)
NPP: 04.15.897

Mengetahui,
Dekan Fakultas Teknologi Industri



(Dr. A. Teguh Siswanto)
NPP: 09.93.464

Pernyataan Orisinalitas Karya

Saya yang bertanda tangan di bawah ini:

Nama Lengkap : Melita Dwi Nugroho

NPM : 150708365

Program Studi : Teknik Informatika

Fakultas : Teknologi Industri

Judul Penelitian : Data *Scraping* Informasi Kuliner dari Media Sosial Instagram

Menyatakan dengan ini:

1. Tugas Akhir ini adalah benar tidak merupakan salinan sebagian atau keseluruhan dari karya penelitian lain.
2. Memberikan kepada Universitas Atma Jaya Yogyakarta atas penelitian ini, berupa Hak untuk menyimpan, mengelola, mendistribusikan, dan menampilkan hasil penelitian selama tetap mencantumkan nama penulis.
3. Bersedia menanggung secara pribadi segala bentuk tuntutan hukum atas pelanggaran Hak Cipta dalam pembuatan Tugas Akhir ini.

Demikianlah pernyataan ini dibuat dan dapat dipergunakan sebagaimana mestinya.

Yogyakarta, 08 Juli 2019

Yang menyatakan,

Melita Dwi Nugroho

150708365

Halaman Persembahan

*“Through hard work, perseverance and a faith in God, you can live
your dreams.”*

-Ben Carson

Tugas Akhir ini dipersembahkan untuk:

Tuhan Yesus Kristus,

Papa, Mama, Keluarga dan

Semua Teman yang selalu mendampingi dikala duka dan suka.

KATA PENGANTAR

Puji syukur kepada Tuhan yang Maha Esa, atas semua berkat dan tuntunan yang diberikan-Nya kepada penulis sehingga, penulis dapat menyelesaikan tugas akhir yang berjudul “*Data Scraping* Informasi Kuliner dari Media Sosial Instagram“ ini dengan baik. Tugas akhir adalah tugas yang diwajibkan kepada mahasiswa Program Studi Teknik Informatika Fakultas Teknologi Industri Universitas Atma Jaya Yogyakarta yang telah lulus dari mata kuliah teori, praktikum dan kerja praktek. Adapun tujuan dari tugas akhir ini adalah sebagai salah satu syarat untuk memperoleh derajat Sarjana Komputer (S.Kom.) dari Program Studi Teknik Informatika Fakultas Teknologi Informasi Universitas Atma Jaya Yogyakarta.

Penulis menyadari bahwa dalam proses pembuatan tugas akhir ini, terdapat banyak pihak yang telah memberikan dukungan, bimbingan, pikiran, tenaga dan doa untuk penulis baik secara langsung maupun tidak langsung. Oleh karena itu, penulis mengucapkan terima kasih kepada :

1. Tuhan Yesus Kristus yang telah memberikan karunia, petunjuk, kesabaran, kekuatan dan perlindungan kepada penulis sehingga penulis mampu memulai masa kuliah hingga mengakhiri masa kuliah melalui selesainya tugas akhir ini dengan baik dan penuh dengan berkat.
2. Papa, Mama dan Kakak, tiga orang hebat dan sangat berarti untuk penulis, yang selama ini telah berjuang untuk anaknya. Orang tua yang telah sabar mendidik, memotivasi dan selalu mendoakan penulis sehingga penulis bisa sampai pada titik ini.
3. Bapak Martinus Maslim, S.T., M.T., selaku Dosen Pembimbing I yang telah sabar dan meluang waktu serta tenaga untuk memberikan bimbingan, petunjuk dan nasihat kepada penulis sehingga tugas akhir ini dapat diselesaikan dengan baik.
4. Y. Sigit Purnomo WP., S.T., M.Kom., selaku Dosen Pembimbing II yang telah sabar dan meluang waktu serta tenaga untuk memberikan bimbingan, petunjuk dan nasihat kepada penulis sehingga tugas akhir ini dapat diselesaikan dengan baik.
5. Seluruh Dosen dan Staf Pengajar Fakultas Teknologi Industri Universitas Atma Jaya Yogyakarta yang telah membantu penulis selama masa kuliah di Program

Studi Teknik Informatika Fakultas Teknologi Industri Universitas Atma Jaya Yogyakarta.

6. Rosiana Fransisca, Theresia Marsha Prawatya dua orang sahabat sekaligus teman seperjuangan yang telah setia menemani, memberi nasihat, mendukung dan menyemangati penulis selama masa perkuliahan.
7. Teman-teman seperjuangan, Nadya, Indah, Lia, Samuel Djoni, Ian, Verry, Apri, Dea, Fara, Grelly, Gege, Megan, Rani, Dio, Travis, Wahyu, Rizka, Reynaldo, Bobby, Komang yang telah memberikan semangat, bantuan, nasihat maupun doa untuk penulis.
8. Seluruh teman-teman FTI, atas kebersamaan selama ini di masa perkuliahan.
9. Semua teman-teman yang mendukung dan menyemangati penulis yang tidak dapat disebutkan satu-persatu.
10. Semua orang yang secara tidak langsung memberikan dukungan dan semangat.

Penulis menyadari bahwa tugas akhir ini masih jauh dari kata sempurna karena keterbatasan waktu dan pengetahuan yang dimiliki penulis. Oleh karena itu, segala kritik dan saran yang bersifat membangun sangat diharapkan. Akhir kata, semoga tugas akhir ini dapat berguna dan bermanfaat bagi semua pihak.

Yogyakarta, 08 Juli 2019

Penulis,

Melita Dwi Nugroho

Daftar Isi

DATA <i>SCRAPING</i> INFORMASI KULINER DARI MEDIA SOSIAL INSTAGRAM.....	i
LEMBAR PENGESAHAN	ii
Pernyataan Orisinalitas Karya.....	iii
Halaman Persembahan	iv
KATA PENGANTAR	v
Daftar Isi.....	vii
Daftar Gambar.....	ix
Daftar Tabel	xii
Intisari	xiii
BAB I.	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian	3
1.5 Metode Penelitian.....	4
1.6 Metodologi Penelitian	5
BAB II.....	6
2.1 Penelitian Terdahulu	6
BAB III.	10
3.1 Media Sosial (Instagram)	10
3.2 <i>Foodgram (Food Blogger</i> di Instagram) dan Kuliner.....	11
3.3 <i>Web Scraping</i> atau <i>Web Crawling</i>	12
3.4 <i>Web Server</i> dan <i>Web Service</i>	13
3.5 JSON.....	16
3.6 Python dan <i>Flask Microframework</i>	16
3.7 <i>Beautiful Soup</i> dan <i>Selenium</i>	18
3.8 <i>Elastic Search</i>	19
BAB IV.	20
4.1 Analisis Sistem.....	20
4.2 Lingkup Masalah.....	31
4.3 Fungsi Produk	31
4.4 Perancangan	37
BAB V.....	61

5.1.	Definisi Perangkat Lunak.....	61
5.2.	Implementasi <i>Data Scraping</i>	62
5.2.1.	<i>Scraping</i> bagian 1 - Add <i>Foodgram</i> User dan GetPostLink.....	62
5.2.2.	<i>Scraping</i> bagian 2 - GetDataPost.....	69
5.2.3.	<i>Scraping</i> bagian 3 - UpdateLikesCount.....	75
5.3.	Implementasi Sistem.....	76
5.3.1.	<i>Log in</i>	76
5.3.2.	<i>Edit Data User / Change Password</i>	78
5.3.3.	Registrasi.....	79
5.3.4.	<i>Forgot Password</i>	80
5.3.5.	Antarmuka <i>Master (User – Location – Foodgram)</i>	81
5.3.6.	Antarmuka <i>Data Scraping</i>	86
5.3.7.	Antarmuka Normalisasi.....	89
5.3.8.	Antarmuka Verifikasi.....	98
5.3.9.	Antarmuka <i>Rejected</i>	103
5.3.10.	Antarmuka <i>Top Food</i> dan Pencarian.....	105
5.3.11.	Antarmuka <i>Hashtag</i> Report.....	111
5.3.12.	Antarmuka Kategori Place Type.....	111
5.3.13.	Antarmuka <i>User/Hashtag Click</i>	112
5.3.14.	Antarmuka SIBM - Modal Detail Post.....	112
5.4.	Hasil Pengujian Sistem <i>Scraping</i>	113
5.5.	Hasil Pengujian terhadap Pengguna.....	114
5.6.	Kritik dan Saran dari Pengguna.....	122
5.7.	Analisis Kelebihan dan Kekurangan Aplikasi.....	123
5.7.1.	Kelebihan.....	123
5.7.2.	Kekurangan.....	123
BAB VI.	124
6.1	Kesimpulan.....	124
6.2	Saran.....	124
DAFTAR PUSTAKA	125

Daftar Gambar

Gambar 3. 1 Contoh Media Sosial Instagram	11
Gambar 3. 2 Proses Web Scraping.....	13
Gambar 3. 3 Arsitektur Web Server.....	14
Gambar 3. 4 Single Endpoint (Codelabs Unikom Tioreza Febrian, 2017)	15
Gambar 3. 5 Penggunaan JSON REST	16
Gambar 3. 6 Struktur Flask Framework (PACKT Publishing, Joël Perras, 2015)	17
Gambar 4. 1 Arsitektur Perangkat Lunak SIBM.....	21
Gambar 4. 2 Flow Sistem Super Admin (Scraping) dan Admin (Normalization).....	22
Gambar 4. 3 Flow Sistem Super Admin Verification & Rejected Normalization	25
Gambar 4. 4 Flowchart Crawl Post Link After Add Foodgram User	26
Gambar 4. 5 Gambar HTML Instagram Number Post & Private	27
Gambar 4. 6 Flowchart Scraping GetDataPost	28
Gambar 4. 7 Flowchart Scraping UpdateLikesCount	30
Gambar 4. 8 Use Case Diagram SIBM	35
Gambar 4. 9 Physical Data Model (PDM) SIBM	36
Gambar 4. 10 Package Diagram	37
Gambar 4. 11 Package Diagram Admin	38
Gambar 4. 12 Halaman Login.....	39
Gambar 4. 13 Halaman Dashboard Super Admin.....	39
Gambar 4. 14 Halaman Dashboard Admin.....	40
Gambar 4. 15 Halaman Master Data User	40
Gambar 4. 16 Modal Tambah Data User	41
Gambar 4. 17 Form Edit Data User	41
Gambar 4. 18 Form Reset Password User	42
Gambar 4. 19 Form Hapus Data User.....	42
Gambar 4. 20 Halaman Foodgram.....	42
Gambar 4. 21 Tambah Foodgram	43
Gambar 4. 22 Hapus Foodgram.....	43
Gambar 4. 23 Scrape User Post	43
Gambar 4. 24 Scrape Data Post	44
Gambar 4. 25 Scrape Update Post	44
Gambar 4. 26 Halaman Master Location Data.....	44
Gambar 4. 27 Halaman Food Post Link.....	45
Gambar 4. 28 Halaman Post Data Normalize	45
Gambar 4. 29 Form Food Post Validation	46
Gambar 4. 30 Form Food Post Validation - No	46
Gambar 4. 31 Form Food Normalization Location & Food Category	47
Gambar 4. 32 Form Food Normalization – Choose Available Location	47
Gambar 4. 33 Form Food Normalization - Food	48
Gambar 4. 34 Form Normalization - Add Food.....	48
Gambar 4. 35 Form Normalization - Search Available Food	49
Gambar 4. 36 Form Normalization - Edit Data Food	49
Gambar 4. 37 Halaman – Post Data Verify	50
Gambar 4. 38 Form - Post Data Verify-Valid Normalization.....	50
Gambar 4. 39 Form - Post Data Verify-Invalid Normalization	51
Gambar 4. 40 Post Data Verify - Valid Post Rejected Question	51
Gambar 4. 41 Post Data Verify - Valid Post Rejected Answer No.....	51

Gambar 4. 42 Post Data Verify- Valid Post Rejected Answer Yes	52
Gambar 4. 43 Post Data Verify - Invalid Post Rejected Question	52
Gambar 4. 44 Post Data Verify - Invalid Post Rejected No.....	53
Gambar 4. 45 Halaman – Rejected Normalization	53
Gambar 4. 46 Form - Rejected Normalization.....	54
Gambar 4. 47 Halaman – All Data Verified	54
Gambar 4. 48 SIBM Admin Profile	55
Gambar 4. 49 SIBM Admin – Edit Profile User.....	55
Gambar 4. 50 SIBM Admin – Change Password	55
Gambar 4. 51 SIBM Admin Register.....	56
Gambar 4. 52 SIBM Admin – Reset Password.....	56
Gambar 4. 53 Halaman SIBM Home.....	57
Gambar 4. 54 Halaman Search	57
Gambar 4. 55 Modal Home SIBM Detail Post	58
Gambar 4. 56 Halaman Place Category	58
Gambar 4. 57 Halaman SIBM Hashtag Report.....	59
Gambar 4. 58 Halaman SIBM User/Hashtag Result.....	60
Gambar 5. 1 Struktur HTML Instagram Profile Page.....	64
Gambar 5. 2 Struktur JSON Profile Page	64
Gambar 5. 3 Struktur JSON User Info with Request.....	65
Gambar 5. 4 Struktur HTML Instagram-Post Link.....	67
Gambar 5. 5 Struktur JSON Graphql PostPage	69
Gambar 5. 6 Struktur JSON Graphql Location Request.....	72
Gambar 5. 7 Halaman Log In.....	76
Gambar 5. 8 Halaman Dashboard Super Admin.....	77
Gambar 5. 9 Halaman Dashboard Admin.....	77
Gambar 5. 10 Halaman Profile	78
Gambar 5. 11 Modal Edit Data -User	78
Gambar 5. 12 Modal Edit Data - Change Password	79
Gambar 5. 13 Halaman Register.....	79
Gambar 5. 14 Halaman Forgot Password	80
Gambar 5. 15 Antarmuka Master User Data.....	81
Gambar 5. 16 Antarmuka Location Data	81
Gambar 5. 17 Antarmuka Master Foodgram	82
Gambar 5. 18 Alert Konfirmasi Edit Data	83
Gambar 5. 19 Modal Konfirmasi Delete Data	83
Gambar 5. 20 Master User-Modal Edit Data User.....	84
Gambar 5. 21 Master User-Modal Edit Data Reset Password.....	84
Gambar 5. 22 Master User-Modal Add User	85
Gambar 5. 23 Master Foodgram-Modal Add New Foodgram User	85
Gambar 5. 24 Flash Message dan Pesan Petunjuk Pengisian Modal	86
Gambar 5. 25 Antarmuka Data Scraping	86
Gambar 5. 26 Modal Konfirmasi Script.....	86
Gambar 5. 27 Antarmuka Foodpost Link	87
Gambar 5. 28 Antarmuka All Foodpost.....	87
Gambar 5. 29 Antarmuka Foodpost Link 2	88
Gambar 5. 30 Antarmuka Foodpost Normalize	89
Gambar 5. 31 Normalization - Modal Valid Food Post.....	90
Gambar 5. 32 Normalization - Post Category	90
Gambar 5. 33 Normalization- Location & Category	91

Gambar 5. 34 Normalization-Location Available.....	91
Gambar 5. 35 Normalization-Location Available Click.....	92
Gambar 5. 36 Normalization - Choose Location.....	93
Gambar 5. 37 Alert Location has been selected.....	93
Gambar 5. 38 Normalization - Location Check Maps.....	95
Gambar 5. 39 Normalization - Added Food.....	96
Gambar 5. 40 Normalization - Choose Food.....	96
Gambar 5. 41 Normalization - Choose Food Konfirmasi.....	96
Gambar 5. 42 Normalization - Create and Add Data Food.....	97
Gambar 5. 43 Normalization - Edit Data Food.....	97
Gambar 5. 44 Antarmuka Post Data Verify.....	98
Gambar 5. 45 Konfirmasi Pengalihan antarmuka Verifikasi Foodpost.....	98
Gambar 5. 46 Antarmuka Food Verification – Valid Foodpost.....	99
Gambar 5. 47 Konfirmasi Normalisasi Disetujui.....	99
Gambar 5. 48 Modal Q Apakah Normalisasi Post tersebut seharusnya Invalid?.....	100
Gambar 5. 49 Modal Verifikasi Valid Post A No - Input Rejected Reason.....	100
Gambar 5. 50 Modal Verifikasi Valid Post A Yes - Input Invalid Post Category.....	100
Gambar 5. 51 Antarmuka Food Verification – Invalid Foodpost.....	101
Gambar 5. 52 Modal Pertanyaan Apakah Normalisasi Post tersebut seharusnya Valid?.....	101
Gambar 5. 53 Modal Verifikasi Invalid Post Jawaban No – Input Invalid Category.....	101
Gambar 5. 54 Antarmuka All Data Verified.....	102
Gambar 5. 55 Antarmuka Halaman Rejected.....	103
Gambar 5. 56 Antarmuka Halaman Rejected - Konfirmasi Pindah Halaman.....	103
Gambar 5. 57 Form Food Rejected Normalization.....	104
Gambar 5. 58 Antarmuka Top Food dan Pencarian.....	105
Gambar 5. 59 Antarmuka Hasil Pencarian.....	106
Gambar 5. 60 Antarmuka Hashtag Report.....	111
Gambar 5. 61 Antarmuka Kategori Place Type.....	111
Gambar 5. 62 Antarmuka User/Hashtag Click.....	112
Gambar 5. 63 Antarmuka Modal Detail Post.....	112
Potongan Kode 5. 1 Search Owner Id.....	62
Potongan Kode 5. 2 getGraphQLProfilePage.....	63
Potongan Kode 5. 3 Method Extract User Post Link.....	66
Potongan Kode 5. 4 Method getGraphQLPostPage.....	69
Potongan Kode 5. 5 Extract Post Info Part 1.....	70
Potongan Kode 5. 6 Extract Post Info Part 2.....	71
Potongan Kode 5. 7 Simpan Resource File.....	74
Potongan Kode 5. 8 Regex Address.....	93
Potongan Kode 5. 9 Create Index tbl post dan konfigurasi di flask.....	107
Potongan Kode 5. 10 Add,Remove,Query Search Index.....	108
Potongan Kode 5. 11 SearchableMixin dan pemanggilan class method search.....	110
Grafik 5. 1 Persentase Hasil Pernyataan 1.....	117
Grafik 5. 2 Persentase Hasil Pernyataan 2.....	117
Grafik 5. 3 Persentase Hasil Pernyataan 3.....	118
Grafik 5. 4 Persentase Hasil Pernyataan 4.....	118
Grafik 5. 5 Persentase Hasil Pernyataan 4.....	119
Grafik 5. 6 Persentase Hasil Pernyataan 6.....	119
Grafik 5. 7 Persentase Hasil Pernyataan 7.....	120
Grafik 5. 8 Persentase Hasil Pernyataan 1.....	121

Daftar Tabel

Table 2. 1 Perbandingan Penelitian	9
Tabel 5. 1 Hasil Pengujian Sistem Scraping	113
Tabel 5. 2 Spesifikasi Responden Pencari Kuliner	114
Tabel 5. 3 Spesifikasi Responden Pemilik/Pengelola Bisnis Kuliner.....	114
Tabel 5. 4 Hasil Pengujian Terhadap Pencari Kuliner	115
Tabel 5. 5 Hasil Pengujian Terhadap Pemilik/Pengelola Bisnis Kuliner.....	121
Tabel 5. 6 Kritik dan Saran	122

Intisari

Banyaknya tempat makan atau makanan saat ini, membuat masyarakat menginginkan makanan yang sesuai dengan keinginan mereka. Oleh karena itu biasanya akan melakukan pencarian dan salah satu media yang dapat membantu adalah Instagram, yang menyediakan tempat untuk dapat membagikan *post* antar pengguna. Namun di Instagram tidak bisa melakukan pencarian berdasarkan *caption* dan harus melakukan scroll satu-satu akun *Foodgram*. Sehingga masyarakat akan kesulitan apabila ingin melakukan pencarian makanan tertentu.

Untuk menangani masalah di atas, maka dibuatkan aplikasi yang bertujuan membantu orang-orang dalam mencari dan menemukan makanan yang diinginkan dengan mudah. Langkah pertama adalah pengambilan data yang menggunakan metode *crawling* dan *scraping*. Langkah kedua adalah pembersihan data menggunakan regex dan penyimpanan data ke *database*. Langkah terakhir pelabelan dan pembersihan data secara manual lewat fitur di aplikasi. Selain itu aplikasi ini menyediakan fitur pencarian berdasarkan *caption* yang ditampilkan melalui website. Aplikasi yang dibangun untuk platform website ini dikembangkan menggunakan *tools* Python, DBMS MySQL dan *framework* Flask.

Tersedia suatu portal yang dapat digunakan untuk mencari tempat makan atau makanan di DIY dari data hasil *scraping*. Hasil survey terhadap pengguna sebanyak 53.3% menyatakan sangat setuju dan 43.3% setuju, bahwa dengan dibangunnya SIBM dapat memberikan kemudahan dalam mencari sebuah tempat makan atau makanan dari kumpulan berbagai *Foodgram*.

Kata kunci : instagram, data *scraping*, makanan, python, flask.