



SKRIPSI

**KLASIFIKASI EMOSI KOMENTAR YOUTUBE
PROGRAM MAKAN BERGIZI GRATIS
MENGUNAKAN ALGORITMA SVM DAN
*LOGISTIC REGRESSION***

KARINA DINDA ARTANTI

NPM 22082010221

DOSEN PEMBIMBING

Seftin Fitri Ana Wati, S.Kom., M.Kom.

Abdul Rezha Efrat Najaf, S.Kom., M.Kom.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI SISTEM INFORMASI
SURABAYA
2026**

LEMBAR PENGESAHAN

KLASIFIKASI EMOSI KOMENTAR YOUTUBE PROGRAM MAKAN BERGIZI GRATIS MENGGUNAKAN ALGORITMA SVM DAN *LOGISTIC REGRESSION*

Oleh :

KARINA DINDA ARTANTI

NPM. 22082010221

Telah dipertahankan dihadapan dan diterima oleh Tim Penguji Skripsi Prodi
Sistem Informasi Fakultas Ilmu Komputer Universitas Pembangunan Veteran
Jawa Timur Pada tanggal 25 Juni 2026

Menyetujui,

Seftin Fitri Ana Wati, S.Kom., M.Kom.

NPT. 21219910320267

Abdul Rezha Efrat Najaf, S.Kom., M.Kom.

NIP. 199409292022031008

Amalia Anjani Arifvanti, S.Kom., M.Kom.

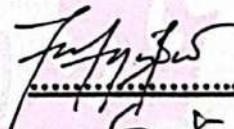
NIP. 199208122018032001

Anindo Saka Fitri, S.Kom., M.Kom.

NIP. 199303252024062001

Muhammad Septama Prasetya, S.Kom., M.Kom.

NIP. 199309282025061002



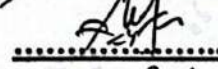
(Pembimbing I)



(Pembimbing II)



(Ketua Penguji)



(Penguji II)



(Penguji III)

Menyetujui,

Dekan Fakultas Ilmu Komputer



Prof. Dr. Ir. Novirina Hendrasarie, MT

NIP. 19681126 199403 2 001

LEMBAR PERSETUJUAN

KLASIFIKASI EMOSI KOMENTAR YOUTUBE PROGRAM MAKAN BERGIZI GRATIS MENGGUNAKAN ALGORITMA SVM DAN *LOGISTIC REGRESSION*

Oleh :

KARINA DINDA ARTANTI

NPM. 22082010221

Menyetujui,

Koordinator Program Studi Sistem Informasi

Fakultas Ilmu Komputer



Siti Mukaromah, S.Kom., M.Kom.

NIP. 198107042021212011

SURAT PERNYATAAN BEBAS PLAGIASI

Saya yang bertanda tangan di bawah ini :

Nama Mahasiswa : Karina Dinda Artanti

NPM : 22082010221

Program : Sarjana (S1)

Program Studi : Sistem Informasi

Fakultas : Ilmu Komputer

Menyatakan bahwa dalam dokumen ilmiah Skripsi ini tidak terdapat bagian dari karya ilmiah lain yang telah diajukan untuk memperoleh gelar akademik di suatu lembaga Pendidikan Tinggi, dan juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang/lembaga lain, kecuali yang secara tertulis disitasi dalam dokumen ini dan disebutkan secara lengkap dalam daftar pustaka.

Dan saya menyatakan bahwa dokumen ilmiah ini bebas dari unsur-unsur plagiasi. Apabila di kemudian hari ditemukan indikasi plagiat pada Skripsi ini, saya bersedia menerima sanksi sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian surat pernyataan ini saya buat dengan sesungguhnya tanpa ada paksaan dari siapapun juga dan untuk dipergunakan sebagaimana mestinya.



Surabaya,

Surat Pernyataan,



KARINA DINDA ARTANTI

NPM. 22082010221

ABSTRAK

Nama Mahasiswa / NPM : Karina Dinda Artanti / 22082010221
Judul Skripsi : Klasifikasi Emosi Komentar YouTube Program Makan Bergizi Gratis Menggunakan Algoritma SVM dan *Logistic Regression*.
Dosen Pembimbing : 1. Seftin Fitri Ana Wati, S.Kom., M.Kom.
2. Abdul Rezha Efrat Najaf, S.Kom., M.Kom.

Komentar publik di YouTube mengandung informasi berharga untuk mengevaluasi respons masyarakat terhadap kebijakan pemerintah, namun sulit dianalisis secara manual. Penelitian ini mengklasifikasikan emosi komentar YouTube terkait Program Makan Bergizi Gratis (MBG) ke dalam delapan kelas emosi Plutchik menggunakan algoritma *Support Vector Machine* (SVM) dan *Logistic Regression* (LR). Sebanyak 13.840 komentar valid (dari 17.959 komentar hasil *scrapping* melalui YouTube Data API v3) dilabeli dengan NRC Emotion *Lexicon* Bahasa Indonesia, lalu diuji pada enam skenario kombinasi algoritma dengan TF-IDF, *Bag of Words*, dan Word2Vec menggunakan rasio 80:20. Kombinasi *Logistic Regression* + *Bag of Words* menghasilkan performa terbaik, dan setelah *hyperparameter tuning* dengan GridSearchCV (C=10) mencapai accuracy 86,24% dan F1-score weighted 86,09%. Uji generalisasi pada 267 komentar MBG baru menghasilkan accuracy 80,86%. Model terbaik diimplementasikan sebagai aplikasi web EmotiSense MBG berbasis *framework* Flask.

Kata Kunci : Klasifikasi Emosi, *Support Vector Machine*, *Logistic Regression*, Youtube, MBG

ABSTRACT

Nama Mahasiswa / NPM : Karina Dinda Artanti / 22082010221
Judul Skripsi : Klasifikasi Emosi Komentar YouTube Program Makan Bergizi Gratis Menggunakan Algoritma SVM dan *Logistic Regression*.
Dosen Pembimbing : 1. Seftin Fitri Ana Wati, S.Kom., M.Kom.
2. Abdul Rezha Efrat Najaf, S.Kom., M.Kom.

Public comments on YouTube contain valuable insights for evaluating public response to government policies but are difficult to analyze manually. This study classifies the emotions of YouTube comments regarding the Program Makan Bergizi Gratis (MBG) into eight Plutchik emotion classes using Support Vector Machine (SVM) and Logistic Regression (LR) algorithms. A total of 13,840 valid comments (from 17,959 scraped through the YouTube Data API v3) were labeled using the Indonesian NRC Emotion Lexicon and evaluated across six scenarios combining the algorithms with TF-IDF, Bag of Words, and Word2Vec on an 80:20 split. The combination of Logistic Regression and Bag of Words achieved the best performance, and after hyperparameter tuning with GridSearchCV (C = 10) reached an accuracy of 86.24% and a weighted F1-score of 86.09%. A generalization test on 267 new MBG comments produced an accuracy of 80.86%. The best model was deployed as a web application named EmotiSense MBG using the Flask framework.

Keywords : *Emotion Classification, Support Vector Machine, Logistic Regression, YouTube, MBG*

KATA PENGANTAR

Puji syukur penulis ucapkan kehadiran Allah SWT atas segala rahmat, karunia, dan hidayah-Nya sehingga penulis dapat menyelesaikan penyusunan skripsi yang berjudul “Klasifikasi Emosi Komentar YouTube Program Makan Bergizi Gratis menggunakan Algoritma *Support Vector Machine* dan *Logistic Regression*” dengan baik dan tepat waktu. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jawa Timur.

Penulis menyadari bahwa penyusunan skripsi ini tidak terlepas dari bantuan, bimbingan, serta dukungan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Ibu Prof. Dr. Ir. Novirina Hendrasarie, MT., selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional "Veteran" Jawa Timur, atas dukungan dan fasilitas yang diberikan selama proses perkuliahan hingga penyusunan skripsi ini.

2. Ibu Siti Mukaromah, S.Kom., M.Kom., selaku Koordinator Program Studi Sistem Informasi Universitas Pembangunan Nasional "Veteran" Jawa Timur, atas dukungan dan arahannya dalam memperlancar proses penyusunan skripsi ini.

3. Penulis menyampaikan rasa terima kasih yang sebesar-besarnya kepada Papa, Mama, Mbak Tika, dan Kasih serta seluruh keluarga tercinta atas doa, kasih sayang, perhatian, dan dukungan yang tidak pernah putus kepada penulis. Terima kasih karena selalu memberikan semangat, menjadi tempat berbagi cerita, dan mendampingi penulis selama menjalani perkuliahan hingga menyelesaikan skripsi ini. Terima kasih telah menciptakan keluarga yang penuh kehangatan dan selalu saling mendukung. Penulis merasa sangat bersyukur dan bangga telah dilahirkan di tengah keluarga yang penuh kasih sayang.

4. Ibu Seftin Fitri Ana Wati, S.Kom., M.Kom. selaku Dosen Pembimbing I, yang telah dengan sabar meluangkan waktu untuk memberikan bimbingan, arahan, serta masukan yang sangat berharga selama proses penyusunan skripsi. Penulis sangat menghargai ketelatenan dan dedikasi beliau dalam mendampingi hingga skripsi ini dapat terselesaikan dengan baik.

5. Bapak Abdul Rezha Efrat Najaf, S.Kom., M.Kom. selaku Dosen Pembimbing II, yang telah memberikan banyak bimbingan, masukan, dan pengarahan. Penulis mengucapkan terima kasih atas nasihat dan perhatian yang diberikan dengan penuh kesabaran selama masa bimbingan.

6. Bapak/Ibu Dosen Penguji yang telah memberikan kritik, saran, dan masukan yang membangun guna penyempurnaan skripsi ini.

7. Seluruh Dosen dan Staf Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jawa Timur, yang telah memberikan ilmu dan pengalaman yang sangat berharga selama masa perkuliahan.

8. Seluruh teman seperjuangan “Galaksi” Sistem Informasi 2022, atas dukungan dan kebersamaan yang telah terjalin selama menempuh perkuliahan.

9. Penulis menyampaikan rasa terima kasih kepada Agung Andhika, yang telah menemani dan memberikan banyak dukungan kepada penulis selama menjalani perkuliahan hingga proses penyusunan skripsi ini. Terima kasih atas doa, perhatian, kesabaran, dan semangat yang selalu diberikan kepada penulis. Terima kasih karena selalu bersedia mendengarkan setiap keluhan, memberikan motivasi di saat penulis merasa lelah, serta selalu mengingatkan penulis untuk tetap semangat dan percaya pada diri sendiri. Kehadiranmu menjadi salah satu penyemangat bagi penulis dalam menghadapi setiap proses dan tantangan selama menyelesaikan skripsi ini. Semoga kita semua senantiasa diberikan kelancaran dalam setiap langkah, dimudahkan dalam meraih cita-cita, serta diberikan kesuksesan dan keberkahan di masa yang akan datang.

10. Penulis menyampaikan rasa terima kasih yang sebesar-besarnya kepada teman-teman "Sekawan", yaitu Nabilah Sahda Firjatullah, Salsa Diah Apriliani, dan Hawa Shabilla Fanfa, yang senantiasa hadir menemani penulis dalam berbagai situasi serta memberikan dukungan moral selama masa perkuliahan. Terima kasih atas kebersamaan, canda tawa, masukan, serta kesediaan untuk menjadi pendengar yang baik dalam setiap suka dan duka yang penulis alami. 4 tahun perjalanan perkuliahan ini menjadi lebih berwarna, bermakna, dan menyenangkan karena kehadiran kalian. Semoga persahabatan yang telah terjalin dapat terus terjaga hingga masa mendatang.

11. Penulis juga mengucapkan terima kasih kepada Ghaitsa, Vina, Ameng, Amru, Fariz, Riznanta, Amel, dan Enjel atas segala dukungan, perhatian, serta semangat yang telah diberikan selama proses penyusunan skripsi ini. Terima kasih karena telah meluangkan waktu untuk mendengarkan keluh kesah, memberikan motivasi ketika penulis merasa lelah dan kehilangan semangat, serta selalu mengingatkan penulis untuk tetap percaya pada kemampuan diri sendiri. Kehadiran dan dukungan kalian menjadi salah satu sumber kekuatan bagi penulis untuk menyelesaikan skripsi ini hingga akhir.

12. Penulis juga menyampaikan rasa terima kasih kepada seluruh pihak yang tidak dapat disebutkan satu per satu yang telah memberikan bantuan, doa, dukungan, maupun semangat kepada penulis selama proses penyusunan skripsi ini. Setiap bantuan dan kebaikan yang diberikan sangat berarti bagi penulis. Penulis berharap skripsi ini dapat memberikan manfaat bagi banyak pihak serta menjadi bentuk apresiasi atas dukungan yang telah diberikan kepada penulis selama ini.

13. Penulis menyampaikan rasa terima kasih kepada diri sendiri, Karina Dinda Artanti, yang telah mampu bertahan dan terus berjuang hingga dapat menyelesaikan skripsi ini. Terima kasih karena tidak memilih untuk menyerah meskipun banyak proses yang terasa berat, melelahkan, dan penuh tantangan. Terima kasih karena selalu berusaha bangkit setiap kali merasa ragu dan terus percaya bahwa semua usaha akan membuahkan hasil. Penulis bangga dapat sampai di titik ini dan membuktikan bahwa kerja keras, doa, serta ketekunan

mampu mengantarkan pada salah satu pencapaian terbesar selama masa perkuliahan. Semoga langkah ini menjadi awal untuk terus berkembang, meraih impian, dan membanggakan orang-orang yang selalu mendukung penulis.

Penulis menyadari bahwa skripsi ini masih jauh dari kata sempurna dan masih terdapat banyak kekurangan, baik dari segi penulisan, isi, maupun analisis yang disajikan. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang bersifat membangun guna perbaikan di masa yang akan datang.

Surabaya, 12 Juni 2026

Penulis

DAFTAR ISI

LEMBAR PENGESAHAN	iii
LEMBAR PERSETUJUAN	v
SURAT PERNYATAAN BEBAS PLAGIASI.....	vii
ABSTRAK	ix
ABSTRACT	xi
KATA PENGANTAR.....	xiii
DAFTAR ISI.....	xvii
DAFTAR GAMBAR.....	xxi
DAFTAR TABEL	xxiii
DAFTAR LAMPIRAN	xxv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan.....	4
BAB II TINJAUAN PUSTAKA.....	7
2.1 Dasar Teori.....	7
2.1.1 Media Sosial Youtube.....	7
2.1.2 Program Makan Bergizi Gratis	8
2.1.3 <i>Text Mining</i>	9
2.1.4 Klasifikasi Emosi.....	10
2.1.5 Pelabelan Emosi.....	11
2.1.6 <i>Preprocessing Text</i>	13
2.1.7 <i>Bag of Words</i> (BoW)	14
2.1.8 TF-IDF	15
2.1.9 <i>Word2Vec</i>	16
2.1.10 <i>Support Vector Machine</i>	17
2.1.11 <i>Logistic Regression</i>	17
2.1.12 <i>Confusion Matrix</i>	18
2.1.13 Flask	19

2.2 Penelitian Terdahulu.....	20
BAB III METODOLOGI PENELITIAN.....	25
3.1 Identifikasi Masalah	25
3.2 Studi Literatur.....	25
3.3 Analisis Kebutuhan.....	26
3.3.1 Kebutuhan Data	26
3.3.2 Kebutuhan <i>software</i> dan <i>hardware</i>	27
3.4 Pengumpulan Data.....	27
3.5 <i>Preprocessing Text</i>	28
3.6 Pelabelan Emosi	29
3.7 <i>Filtering data</i>	30
3.8 Eksplorasi Data.....	31
3.9 Pembagian Data.....	31
3.10 Ekstraksi Fitur.....	32
3.11 Perancangan Model	32
3.11.1 Konfigurasi Model <i>Support Vector Machine</i>	33
3.11.2 Konfigurasi Model <i>Logistic Regression</i>	34
3.11.3 <i>Tuning</i> Parameter	34
3.11.4 Evaluasi Model	35
3.11.5 Penyimpanan Model	35
3.12 <i>Deployment</i>	36
BAB IV HASIL DAN PEMBAHASAN.....	39
4.1 Pengumpulan Data.....	39
4.2 <i>Preprocessing Text</i>	41
4.2.1 <i>Case Folding</i>	41
4.2.2 <i>Cleansing</i>	42
4.2.3 <i>Tokenizing</i>	43
4.2.4 <i>Normalization</i>	43
4.2.5 <i>Stemming</i>	45
4.2.6 <i>Stopword Removal</i>	46
4.3 Pelabelan Emosi	48
4.4 <i>Filtering</i>	49
4.5 Eksplorasi Data.....	50
4.5.1 Distribusi Kelas Emosi	51
4.5.2 Analisis Panjang Komentar.....	51

4.5.3 Analisis <i>Word Cloud</i>	53
4.6 Pembagian Data	54
4.7 Perancangan Model.....	55
4.7.1 Konfigurasi <i>Library</i> dan Persiapan <i>Pipeline</i>	56
4.7.2 Ekstraksi Fitur.....	58
4.7.3 Konfigurasi Model SVM dan <i>Logistic Regression</i>	60
4.7.4 Skenario 1: SVM + TF-IDF.....	61
4.7.5 Skenario 2: SVM + <i>Bag of Words</i>	63
4.7.6 Skenario 3: SVM + <i>Word2Vec</i>	65
4.7.7 Skenario 4: <i>Logistic Regression</i> + TF-IDF	68
4.7.8 Skenario 5: <i>Logistic Regression</i> + <i>Bag of Words</i>	70
4.7.9 Skenario 6: <i>Logistic Regression</i> + <i>Word2Vec</i>	72
4.7.10 Rekapitulasi dan Perbandingan Keenam Skenario	75
4.7.11 <i>Tuning</i> Parameter Model Terbaik	76
4.7.12 Evaluasi Model	79
4.7.13 Perbandingan Hasil Skripsi dengan Penelitian Terdahulu	81
4.7.14 Uji Generalisasi Model pada Data Scraping Baru	82
4.8 <i>Deployment</i>	87
BAB V PENUTUP	93
5.1 Kesimpulan	93
5.2 Saran.....	94
DAFTAR PUSTAKA	95
LAMPIRAN.....	101

DAFTAR GAMBAR

Gambar 2. 1 Roda Emosi Plutchik.....	11
Gambar 3. 1 Alur Penelitian.....	25
Gambar 3. 2 Alur Perancangan Model.....	32
Gambar 3. 3 <i>Wireframe</i> Beranda	37
Gambar 3. 4 <i>Wireframe</i> Input Teks.....	37
Gambar 3. 5 <i>Wireframe</i> Input CSV.....	38
Gambar 4. 1 Potongan Kode <i>Crawling</i> Komentar Youtube.....	39
Gambar 4. 2 <i>Dataset</i> Skripsi	40
Gambar 4. 3 <i>Dataset</i> Format File CSV	40
Gambar 4. 4 Potongan Kode Untuk Proses <i>Case Folding</i>	41
Gambar 4. 5 Potongan Kode Untuk Proses <i>Cleansing</i>	42
Gambar 4. 6 Potongan Kode Untuk Proses <i>Tokenizing</i>	43
Gambar 4. 7 Potongan Kode Untuk Proses <i>Normalization</i>	45
Gambar 4. 8 Potongan Kode Untuk Proses <i>Stemming</i>	46
Gambar 4. 9 Potongan Kode Untuk Proses <i>Stopword Removal</i>	47
Gambar 4. 10 Contoh Kode Pelabelan Data	49
Gambar 4. 11 Distribusi Kelas Emosi <i>Dataset</i>	51
Gambar 4. 12 Boxplot Panjang Komentar per Kelas Emosi.....	52
Gambar 4. 13 Word Cloud <i>Dataset</i> Komentar Program MBG.....	53
Gambar 4. 14 Potongan Kode Untuk Pembagian Data.....	54
Gambar 4. 15 Potongan Kode Import Library	57
Gambar 4. 16 Potongan Kode Ekstraksi Fitur TF-IDF	58
Gambar 4. 17 Potongan Kode Ekstraksi Fitur <i>Bag of Words</i>	59
Gambar 4. 18 Potongan Kode Ekstraksi Fitur Word2Vec	59
Gambar 4. 19 Potongan Kode Konfigurasi Awal Model SVM	60
Gambar 4. 20 Potongan Kode Konfigurasi Awal Model <i>Logistic Regression</i>	60
Gambar 4. 21 Classification Report Skenario 1 SVM + TF-IDF	61
Gambar 4. 22 Confusion Matrix Skenario 1 SVM + TF-IDF.....	62
Gambar 4. 23 Classification Report Skenario 2 SVM + BoW	63
Gambar 4. 24 Confusion Matrix Skenario 2 SVM + BoW	64

Gambar 4. 25 Classification Report Skenario 3 SVM + Word2Vec.....	66
Gambar 4. 26 Confusion Matrix Skenario 3 SVM + Word2Vec	67
Gambar 4. 27 Classification Report Skenario 4 LR + TF-IDF	68
Gambar 4. 28 Confusion Matrix Skenario 4 LR + TF-IDF	69
Gambar 4. 29 Classification Report Skenario 5 LR + BoW	70
Gambar 4. 30 Confusion Matrix Skenario 5 LR + BoW	71
Gambar 4. 31 Classification Report Skenario 6 LR + Word2Vec	73
Gambar 4. 32 Confusion Matrix Skenario 6 LR + Word2Vec.....	74
Gambar 4. 33 Potongan Kode GridSearchCV untuk Tuning <i>Logistic Regression</i>	78
Gambar 4. 34 Confusion Matrix pada Data MBG Baru.....	85
Gambar 4. 35 Halaman Beranda EmotiSense MBG	88
Gambar 4. 36 Halaman Analisis Teks	88
Gambar 4. 37 Halaman Analisis CSV	89
Gambar 4. 38 Halaman Evaluasi Model.....	90
Gambar 4. 39 Halaman Evaluasi Model — Metrik per Kelas & F1-Score per Kelas	90
Gambar 4. 40 Halaman Evaluasi Model — Confusion Matrix	91

DAFTAR TABEL

Tabel 2. 1 Definisi Delapan Kelas Emosi Berdasarkan Teori Plutchik	12
Tabel 2. 2 Tabel Penelitian Terdahulu	20
Tabel 4. 1 Contoh Hasil Dari Tahap <i>Preprocessing</i>	47
Tabel 4. 2 Pembagian Data Latih dan Data Uji per Kelas Emosi	54
Tabel 4. 3 Skenario Model	55
Tabel 4. 4 Grafik Perbandingan Akurasi dan F1-Score 6 Skenario	75
Tabel 4. 5 Hasil <i>Cross-Validation</i> GridSearchCV per Nilai C	78
Tabel 4. 6 Hasil <i>Tuning</i> Parameter Model Terbaik.....	78
Tabel 4. 7 Classification Report Model LR + BoW (Setelah Tuning) pada Data Uji 20%	79
Tabel 4. 8 Ringkasan Performa Model LR + BoW pada Semua Tahap Evaluasi	80
Tabel 4. 9 Perbandingan Hasil Penelitian dengan Penelitian Terdahulu	81
Tabel 4. 10 Distribusi Emosi Data MBG Baru (Hasil <i>Scraping</i>).....	83
Tabel 4. 11 Perbandingan Performa: Data Uji 20% vs Data MBG Baru.....	84
Tabel 4. 12 Classification Report Model LR + BoW pada Data MBG Baru	84
Tabel 4. 13 Contoh Prediksi Benar dan Salah pada Uji Generalisasi Data MBG	86

DAFTAR LAMPIRAN

Lampiran 1 Hasil Scraping Data	101
--------------------------------------	-----