

SCR-Analysis Data

**ASPECT-BASED SENTIMENT ANALYSIS IPHONE 14 PRO
MENGUNAKAN ALGORITMA XGBOOST DAN
METODE PARTICLE SWARM OPTIMIZATION**

SKRIPSI



Oleh :

Galih Arum Prabowo

19081010181

**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"
JAWA TIMUR
2023**

ASPECT-BASED SENTIMENT ANALYSIS IPHONE 14 PRO MENGUNAKAN ALGORITMA XGBOOST DAN METODE PARTICLE SWARM OPTIMIZATION

SKRIPSI

Diajukan untuk Memenuhi Salah Satu Syarat Memperoleh
Gelar Sarjana Komputer Program Studi Informatika



Oleh :

Galih Arum Prabowo

19081010181

**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"
JAWA TIMUR
2023**

LEMBAR PENGESAHAN SKRIPSI

Judul : ASPECT-BASED SENTIMENT ANALYSIS IPHONE 14 PRO
MENGUNAKAN ALGORITMA XGBOOST DAN METODE
PARTICLE SWARM OPTIMIZATION

Oleh : Galih Arum Prabowo

NPM : 19081010181

Telah Diseminarkan Dalam Ujian Skripsi Pada :
Hari Jumat Tanggal 10 November 2023

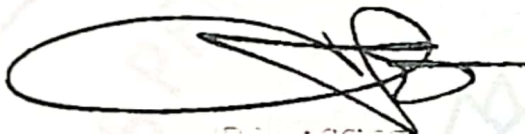
Mengetahui

Dosen Pembimbing

Dosen Penguji

1.

1.



Dr. Basuki Rahmat, S.Si, MT.

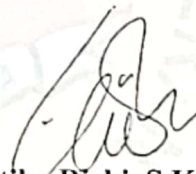
Eva Yulia Puspaningrum, S.Kom., M.Kom

NIP. 19690723 2021211 002

NIP. 19890705 2021212 002

2.

2.



Henni Endah Wahanani, ST, M.Kom

Agung Mustika Rizki, S.Kom., M.Kom

NIP. 19780922 2021212 005

NIP. 19930725 2022031 008

Menyetujui

Dekan

Koordinator Program Studi

Fakultas Ilmu Komputer

Informatika



Prof. Dr. Ir. Novirina Hendrasarie, M.T.

Fetty Tri Anggraeny, S.Kom, M.Kom

NIP. 19681126 199403 2 001

NIP. 19820211 2021212 005

SURAT PERNYATAAN ANTI PLAGIAT

Saya, mahasiswa Program Studi Informatika UPN “Veteran” Jawa Timur,
yang bertandatangan di bawah ini:

Nama : Galih Arum Prabowo

NPM : 19081010181

Menyatakan bahwa judul skripsi yang saya ajukan dan kerjakan, dengan judul :

**“ASPECT-BASED SENTIMENT ANALYSIS IPHONE 14 PRO
MENGUNAKAN ALGORITMA XGBOOST DAN METODE PARTICLE
SWARM OPTIMIZATION”**

Bukan merupakan plagiat dari skripsi/tugas akhir/penelitian orang lain dan juga bukan merupakan produk atau software yang saya beli dari pihak lain. Saya juga menyatakan bahwa skripsi ini adalah pekerjaan saya sendiri, kecuali yang dinyatakan dalam daftar pustaka dan tidak pernah diajukan untuk syarat memperoleh gelar di UPN “Veteran” Jawa Timur maupun di institusi pendidikan lain. Jika ternyata di kemudian hari pernyataan ini terbukti tidak benar, maka Saya siap menerima segala konsekuensinya.

Surabaya, 10 November 2023

Hormat Saya,



Galih Arum Prabowo
NPM : 19081010181

ASPECT-BASED SENTIMENT ANALYSIS IPHONE 14 PRO MENGUNAKAN ALGORITMA XGBOOST DAN METODE PARTICLE SWARM OPTIMIZATION

Nama Mahasiswa : Galih Arum Prabowo

NPM : 19081010181

Program Studi : Informatika

**Dosen Pembimbing : Dr. Basuki Rahmat, S.Si, MT.
Henni Endah Wahanani, ST, M.Kom**

ABSTRAK

Youtube merupakan salah satu aplikasi yang banyak dikunjungi orang Indonesia pada tahun 2022. Hal ini dimanfaatkan banyak pihak sebagai wadah untuk branding dan penyebaran informasi, salah satu channel yang cukup menarik pengguna Youtube khususnya para penggemar gadget yaitu GadgetIn. Hal ini berguna bagi calon user dalam mencari informasi mengenai Gadget yang akan dibeli selain video review yang di sajikan calon user juga dapat melihat ulasan yang di berikan pengguna youtube lainnya untuk mengetahui pengalaman pengguna lain dalam menggunakan produk tersebut. ulasan ini juga berguna bagi developer aplikasi untuk dapat mengetahui aspek-aspek apa saja yang sudah baik dan aspek-aspek apa yang kurang sehingga bisa dijadikan bahan evaluasi untuk kedepannya.

Penelitian ini akan mengimplementasikan algoritma XGBoost dan metode optimasi PSO untuk melakukan analisis sentimen berbasis aspek pada kolom komentar video youtube review iPhone 14 Pro Indonesia pada channel GadgetIn, yang bertujuan untuk mengukur tingkat akurasi algoritma XGBoost dalam mengklasifikasi ulasan ke dalam dua kelas, yaitu kelas sentimen positif-negatif, sekaligus mengetahui kelebihan dan kelemahan Iphone iPhone 14 pro. Dataset yang digunakan sebanyak 1000 data yang di scraping dari kolom komentar Youtube.

Dari hasil pengujian didapat tingkat akurasi yang cukup tinggi, yaitu antara 89% sampai 97%. Selain itu, dari penelitian ini juga dapat diketahui bahwa kelebihan yang dimiliki terdapat pada aspek citra merek. Sedangkan untuk kelemahannya, iPhone 14 pro masih banyak mendapat sentimen negatif terutama pada aspek desain

Kata kunci: Analisis Sentimen, iPhone 14 pro, PSO, XGBoost, Youtube

KATA PENGANTAR

Alhamdulillah, puji syukur kehadiran Allah SWT yang telah memberikan rahmat dan inayah-Nya sehingga penulis dapat menyelesaikan laporan skripsi yang berjudul “Aspect-Based Sentiment Analysis Iphone 14 Pro Menggunakan Algoritma Xgboost Dan Metode Particle Swarm Optimization”. Terima kasih saya ucapkan kepada seluruh pihak yang telah memberikan banyak dukungan kepada penulis sehingga penulis bisa menyelesaikan laporan skripsi ini dengan baik tepat waktu.

Terima kasih saya ucapkan kepada seluruh pihak yang telah memberikan banyak dukungan kepada penulis sehingga penulis bisa menyelesaikan laporan skripsi ini dengan baik tepat waktu.

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

Semoga laporan skripsi ini bisa menambah wawasan para pembaca dan bisa bermanfaat untuk perkembangan dan peningkatan ilmu pengetahuan.

Surabaya, Oktober 2023

Galih Arum Prabowo
19081010181

UCAPAN TERIMA KASIH

Laporan skripsi ini dapat diselesaikan dengan baik atas berkat dan rahmat Allah SWT yang telah memberikan kesehatan dan kelancaran untuk penulis selama proses menyelesaikan skripsi. Proses penyelesaian skripsi ini juga tak lepas dari bantuan dan dukungan berbagai pihak, untuk itu penulis ingin menyampaikan ucapan terima kasih kepada:

1. Bapak Prof. Dr. Ir. Akhmad Fauzi, M.MT selaku Rektor Universitas Pembangunan Nasional “Veteran” Jawa Timur.
2. Ibu Prof. Dr. Ir. Novirina Hendrasarie, M.T. selaku Dekan Fakultas Ilmu Komputer di Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Ibu Fetty Tri Anggraeny S.Kom., M.Kom. selaku Koordinator Program Studi Informatika Fakultas Ilmu Komputer UPN “Veteran” Jawa Timur sekaligus dosen wali yang telah banyak memberikan arahan dan bimbingan kepada penulis selama proses perkuliahan.
4. Bapak Dr. Basuki Rahmat, S.Si, MT. selaku dosen pembimbing I, dan Ibu Henni Endah Wahanani, ST, M.Kom. selaku dosen pembimbing II penulis yang selalu memberikan arahan selama proses penyelesaian skripsi.
5. Staff dan seluruh bapak ibu dosen program studi Informatika UPN “Veteran” Jawa Timur yang telah mengajar dan memberikan ilmu serta pengalaman selama masa perkuliahan.
6. Kedua orang tua penulis, Bapak Agus Budi Widodo dan Ibu Denny Kristanti, serta adik-adik yang selalu memberikan motivasi dan doa untuk penulis.
7. Teman-teman Informatika angkatan 2019, kakak dan adik tingkat, serta seluruh pihak yang telah membantu dan mendukung penulis dalam menyelesaikan skripsi dan perkuliahan ini.

Penulis hanya bisa berharap, semoga Allah SWT selalu memberikan perlindungan dan balasan yang lebih di kemudian hari. Amin

DAFTAR ISI

LEMBAR PENGESAHAN SKRIPSI	ii
SURAT PERNYATAAN ANTI PLAGIAT	Error! Bookmark not defined.
ABSTRAK	iv
KATA PENGANTAR	v
UCAPAN TERIMA KASIH.....	vi
DAFTAR ISI	vii
DAFTAR TABEL	x
DAFTAR GAMBAR.....	xi
DAFTAR KODE PROGRAM.....	xii
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	4
1.3. Tujuan.....	4
1.4. Manfaat.....	5
1.5. Batasan Masalah.....	5
BAB II TINJAUAN PUSTAKA.....	6
2.1. Penelitian terdahulu	6
2.2. Youtube	8
2.3. Analisis sentimen.....	8
2.4. Aspect-Based Sentiment Analysis (ABSA).....	8
2.5. Preprocessing	9
2.5.1 Case-folding	9
2.5.2 Cleaning.....	9
2.5.3 Tokenizing.....	9
2.5.4 Stopword Removal	10
2.5.5 Stemming.....	10
2.6. eXtreme Gradient Boosting (XGBoost).....	10
2.7. Term Frequency-Inverse Document Frequency (TF-IDF)	12

2.8.	Particle Swarm Optimization (PSO)	13
2.9.	Confusion Matrix	15
BAB III METODOLOGI		17
3.1.	Tahapan Penelitian	17
3.2.	Studi Literatur	18
3.3.	Analisa dan Desain	18
3.4.	Pengambilan dan Labelling Data	20
3.5.	Preprocessing Data	22
3.5.1.	Case Folding	23
3.5.2.	Cleaning	23
3.5.3.	Tokenizing	24
3.5.4.	Stopword Removal	25
3.5.5.	Stemming	26
3.6.	Word Embedding	26
3.7.	Particle Swarm Optimization (PSO)	28
3.8.	Extreme Gradient Boosting	34
3.9.	Skenario Pengujian	36
BAB IV HASIL DAN PEMBAHASAN		38
4.1.	Akuisisi Dan Labeling Data	38
4.2.	Preprocessing Data	40
4.2.1	<i>Case-folding</i>	41
4.2.2	<i>Cleaning</i>	42
4.2.3	<i>Tokenizing</i>	43
4.2.4	<i>Stopword Removal</i>	44
4.2.5	<i>Stemming</i>	45

4.3.	Word Embedding.....	46
4.4.	Splitting Data.....	48
4.5.	Particle Swarm Optimization.....	49
4.6.	XGBoost.....	53
4.7.	Hasil Evaluasi Model.....	54
4.8.	Analisis Sentimen Aspek.....	55
BAB V KESIMPULAN DAN SARAN.....		57
5.1	Kesimpulan.....	57
5.2	Saran.....	57
DAFTAR PUSTAKA.....		58

DAFTAR TABEL

Tabel 2. 1 Confusion Matrix	15
Tabel 3. 1 Contoh Case-folding	23
Tabel 3. 2 Contoh Cleaning	23
Tabel 3. 3 Contoh Tokenizing.....	24
Tabel 3. 4 Contoh Stopword Removal.....	25
Tabel 3. 5 Contoh Stemming	26
Tabel 3. 6 Inisialisasi Partikel	32
Tabel 3. 7 Hasil Evaluasi Fungsi Objektif.....	32
Tabel 3. 8 Hasil Evaluasi Pembaruan Posisi dan Kecepatan 1.....	33
Tabel 3. 9 Hasil Evaluasi Pembaruan Posisi dan Kecepatan 2.....	33
Tabel 3. 10 Menentukan Probabilitas.....	34
Tabel 3. 11 Skenario Pengujian.....	36
Tabel 4. 1 Hasil Evaluasi Fungsi Objektif.....	54

DAFTAR GAMBAR

Gambar 2. 1 Flowchart Extreme Gradient Boosting	11
Gambar 2. 2 Ilustrasi Extreme Gradient Boosting	12
Gambar 3. 1 Tahapan Penelitian	17
Gambar 3. 2 Diagram Alur Sistem	18
Gambar 3. 3 Diagram Alur Tahapan Pengambilan dan Labelling Data.....	20
Gambar 3. 4 Scrapping Dataset	21
Gambar 3. 5 Contoh Labelling Dataset	22
Gambar 3. 6 Diagram Alur Tahapan Preprocessing Data	22
Gambar 3. 7 Diagram Alur PSO.....	28
Gambar 3. 8 Diagram Alur Inisialisasi Particle.....	29
Gambar 3. 9 Diagram Alur Update Partikel	30
Gambar 3. 10 Diagram Alur Evaluasi Fitness	31
Gambar 3. 11 Pohon Xgboost	36
Gambar 4. 1 Dataset Hasil Scraping Data	38
Gambar 4. 2 Download Scraping Data	39
Gambar 4. 3 Dataset	39
Gambar 4. 4 Upload File Untuk Preprocessing	41
Gambar 4. 5 Hasil Upload File	41
Gambar 4. 6 Hasil Case-folding	42
Gambar 4. 7 Hasil Cleaning	43
Gambar 4. 8 Hasil Tokenizing.....	44
Gambar 4. 9 Hasil Stopword Removal	45
Gambar 4. 10 Hasil Stemming	46
Gambar 4. 11 Hasil Label Encoding	47
Gambar 4. 12 Hasil TF-IDF	48
Gambar 4. 13 Hasil Best Max_depth	53
Gambar 4. 14 Analisis Sentimen Masing Masing Aspek.....	55
Gambar 4. 15 Analisis Sentimen Aspek Design	56

DAFTAR KODE PROGRAM

Kode Program 4. 1 Import Library Preprocessing Data.....	40
Kode Program 4. 2 Upload File Untuk Preprocessing	40
Kode Program 4. 3 Case-folding	41
Kode Program 4. 4 Cleaning	42
Kode Program 4. 5 Tokenizing	43
Kode Program 4. 6 Stopword Removal.....	44
Kode Program 4. 7 Stemming	45
Kode Program 4. 8 Label Encoding	46
Kode Program 4. 9 TF-IDF	47
Kode Program 4. 10 Splitting Data	48
Kode Program 4. 11 Membuat Fungsi Objective	49
Kode Program 4. 12 Inisialisasi PSO	51
Kode Program 4. 13 Update Partikel.....	52
Kode Program 4. 14 Mencetak Hasil Max Depth	53
Kode Program 4. 15 XGBoost.....	53