



**KLASIFIKASI SENTIMEN FENOMENA *CHILDFREE* PADA
KOMENTAR VIDEO YOUTUBE MENGGUNAKAN METODE *SUPPORT
VECTOR MACHINE (SVM)***

SKRIPSI

BERLIANA SEPTYANI SUGANDA

NIM. 2010512004

PROGRAM STUDI S1 SISTEM INFORMASI

FAKULTAS ILMU KOMPUTER

UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA

2024



**KLASIFIKASI SENTIMEN FENOMENA *CHILDFREE* PADA
KOMENTAR VIDEO YOUTUBE MENGGUNAKAN METODE *SUPPORT
VECTOR MACHINE (SVM)***

SKRIPSI

**Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana
Komputer**

BERLIANA SEPTYANI SUGANDA

NIM. 2010512004

PROGRAM STUDI S1 SISTEM INFORMASI

FAKULTAS ILMU KOMPUTER

UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA

2024

PERNYATAAN ORISINALITAS

PERNYATAAN ORISINALITAS

Skripsi ini merupakan hasil karya sendiri serta semua sumber referensi yang dikutip maupun yang dirujuk telah saya nyatakan benar.

Nama : Berliana Septyani Suganda
NIM : 2010512004
Tanggal : 21 Juni 2024

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan berlaku.

Jakarta, 21 Juni 2024

Yang Menyatakan,



Berliana Septyani Suganda

PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademik Universitas Pembangunan Nasional Veteran Jakarta, saya yang bertanda tangan di bawah ini:

Nama : Berliana Septyani Suganda
NIM : 2010512004
Fakultas : Ilmu Komputer
Program Studi : S1 Sistem Informasi

Demi pembangunan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional Veteran Jakarta Hak Bebas Royalti Eksklusif (*Non-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul:

Klasifikasi Sentimen Fenomena *Childfree* Pada Komentar Video Youtube Menggunakan Metode *Support Vector Machine* (SVM)

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sesungguhnya.

Dibuat di : Jakarta
Pada Tanggal : 21 Juni 2024

Yang menyatakan,



Berliana Septyani Suganda

LEMBAR PENGESAHAN**LEMBAR PENGESAHAN**

Skripsi ini diajukan oleh :

Nama : Berliana Septyani Suganda
NIM : 2010512004
Program Studi : S1 Sistem Informasi
Judul Tugas Akhir : Klasifikasi Sentimen Fenomena *Childfree* Pada Komentar Video Youtube Menggunakan Metode *Support Vector Machine* (SVM)

Telah berhasil dipertahankan dihadapan Tim Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Ilmu Komputer pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta



Iin Ernawati S.Kom., M.Si.

Penguji I



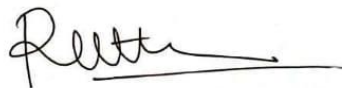
I Wayan Widi Pradnyana S.Kom., M.TI.

Penguji II



Ika Nurlaili Isnainiyah S.Kom., M.Sc.

Pembimbing I



Ruth Mariana Bunga Wadu, S.Kom., MMSI.

Pembimbing II



Prof. Dr. Ir. Supriyanto, ST., M.Sc., IPM.

Dekan



Anita Muliawati, S.Kom., MTI.

Ketua Program Studi S1 Sistem Informasi

Ditetapkan di : Jakarta

Tanggal Ujian : 21 Juni 2024

***KLASIFIKASI SENTIMEN FENOMENA CHILDFREE PADA KOMENTAR
VIDEO YOUTUBE MENGGUNAKAN METODE SUPPORT VECTOR
MACHINE (SVM)***

BERLIANA SEPTYANI SUGANDA

ABSTRAK

Penggunaan internet telah menjadi kebutuhan sehari-hari, terutama dengan meningkatnya penggunaan media sosial. YouTube, sebagai salah satu *platform* terbesar, menjadi tempat bagi pengguna untuk menyampaikan opini tentang berbagai fenomena, termasuk *childfree*. Dalam penelitian ini, analisis sentimen terhadap komentar pengguna YouTube dilakukan untuk memahami respons masyarakat terhadap fenomena *childfree*. Ditemukan mayoritas sentimen pengguna adalah positif, dengan banyaknya dukungan dan semangat. Metode *Support Vector Machine* (SVM) digunakan untuk mengklasifikasikan sentimen, menunjukkan kinerja signifikan. Penggunaan teknik *oversampling* SMOTE (*Synthetic Minority Over-sampling Technique*) membantu meningkatkan kinerja model, terutama dalam mengatasi ketidakseimbangan kelas dalam dataset. Evaluasi menunjukkan hasil akurasi 80% dengan SMOTE dan 79% tanpa SMOTE. Penelitian ini berkontribusi dalam memahami persepsi masyarakat terhadap fenomena sosial *childfree* di media sosial.

Kata kunci: Analisis Sentimen, *Support Vector Machine*, YouTube, *Childfree*.

***KLASIFIKASI SENTIMEN FENOMENA CHILDFREE PADA KOMENTAR
VIDEO YOUTUBE MENGGUNAKAN METODE SUPPORT VECTOR
MACHINE (SVM)***

BERLIANA SEPTYANI SUGANDA

ABSTRACT

The use of the internet has become a daily necessity, especially with the increasing use of social media. YouTube, as one of the largest platforms, is a place for users to express opinions about various phenomena, including childfree. In this study, a sentiment analysis of YouTube user comments was conducted to understand people's responses to the childfree phenomenon. It was found that the majority of user sentiment was positive, with a lot of support and encouragement. The Support Vector Machine (SVM) method was used to classify the sentiments, showing significant performance. The use of the oversampling technique SMOTE (Synthetic Minority Over-sampling Technique) helped improve the performance of the model, especially in overcoming class imbalance in the dataset. The evaluation showed an accuracy result of 80% with SMOTE and 79% without SMOTE. This research contributes to understanding people's perception of the social phenomenon of childfree on social media.

Keywords: Sentiment Analysis, Support Vector Machine, YouTube, Childfree.

KATA PENGANTAR

Segala puji bagi Allah SWT. Karena atas karunia dan rahmatnya sehingga penulis dapat menyelesaikan skripsi tugas akhir dengan judul “**Klasifikasi Sentimen Fenomena *Childfree* Pada Komentar Video YouTube Menggunakan Metode *Support Vector Machine* (SVM)**” dengan baik. Penulisan skripsi ini ditunjukkan sebagai syarat kelulusan dari Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jakarta.

Atas selesainya proposal ini, penulis mengucapkan rasa syukur kepada Allah SWT. Karena tidak lepas dari bantuan banyak pihak yang telah memberikan bantuan dan dukungan. Untuk itu penulis menyampaikan rasa terima kasih kepada:

1. Allah SWT atas segala Rahmat, ridho, dan karunia-Nya.
2. Kedua Orang Tua yang selalu memberikan dukungan secara moral maupun materil.
3. Prof. Dr. Ir. Supriyanto, ST., M.Sc., IPM selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jakarta.
4. Ibu Anita Muliawati, S.Kom., MTI selaku Ketua Program Studi S1 Sistem Informasi.
5. Ibu Ika Nurlaili Isnainiyah, S.Kom., M.Sc., Ibu Helena Helena Nurramdhan Irmada, S.Pd., M.Kom., dan Ibu Ruth Mariana Bunga Wadu, S.Kom., MMSI selaku dosen pembimbing atas kepercayaan, dukungan, bantuan, dan arahnya selama pengerjaan skripsi.
6. Dosen-dosen serta seluruh civitas akademika Fakultas Ilmu Komputer UPN Veteran Jakarta atas seluruh ilmu, nasihat, bantuan, dan dukungannya selama penulis menuntut ilmu disini.
7. Bilqis, Adzra, Tsaabitah, Dinda, dan Rani selaku teman seperjuangan yang selalu memberikan dukungan, bantuan, dan berbagi pengetahuan selama proses penyusunan skripsi.
8. Jaehyun dan Mark yang telah menjadi sumber inspirasi, motivasi, dan semangat bagi penulis dalam menyelesaikan skripsi ini.
9. Diri sendiri yang sudah berproses dan berjuang untuk menyelesaikan penyusunan tugas akhir skripsi.

10. Seluruh pihak yang telah memberikan dukungan serta bantuan yang bersifat langsung maupun tidak langsung dan tidak dapat penulis sebutkan satu persatu tanpa mengurangi rasa hormat.

Penulis menyadari bahwa masih banyak keterbatasan dan kekurangan yang ada dalam penulisan Skripsi ini. Maka dari itu, segala kritik dan saran sangat penulis butuhkan agar dapat menghasikan karya tulis yang lebih baik. Sekian yang dapat penulis sampaikan, semoga skripsi ini dapat berguna dan bermanfaat bagi banyak orang.

Jakarta, 22 April 2024

Berliana Septyani Suganda

DAFTAR ISI

PERNYATAAN ORISINALITAS.....	ii
PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS.....	iii
LEMBAR PENGESAHAN	iv
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR.....	vii
DAFTAR ISI	ix
DAFTAR GAMBAR	xii
DAFTAR TABEL.....	xiii
DAFTAR RUMUS	xiv
DAFTAR LAMPIRAN	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
2.1 Rumusan Masalah	4
2.1 Ruang Lingkup	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	5
1.6 Luaran Penelitian.....	5
1.7 Sistematika Penulisan.....	5
BAB II TINJAUAN PUSTAKA	7
2.1 Analisis Sentimen.....	7
2.2 <i>Childfree</i>	7
2.3 YouTube	8
2.4 <i>Text Mining</i>	8
2.5 <i>Text Preprocessing</i>	8
2.6 <i>Term Weighting</i>	9
2.6.1 <i>Term Frequency (TF)</i>	10
2.6.2 <i>Inverse Document Frequency (IDF)</i>	11
2.6.3 <i>Term Frequency – Inverse Document Frequency (TF-IDF)</i>	11
2.7 SMOTE	11
2.8 <i>Support Vector Machine (SVM)</i>	12

2.9	Evaluasi Performa	12
2.9.1	<i>Confusion Matrix</i>	12
2.9.2	Akurasi	13
2.9.3	Presisi	13
2.9.4	<i>Recall</i>	13
2.9.5	<i>F1-Score</i>	14
2.10	Asosiasi Kata	14
2.11	Bahasa Pemrograman <i>Python</i>	14
2.12	Penelitian Terdahulu	15
BAB III METODOLOGI PENELITIAN.....		20
3.1	Tahapan Penelitian.....	20
3.2	Identifikasi Masalah	21
3.4	Pengumpulan Data	21
3.5	Pelabelan Data	21
3.6	Praproses Data	22
3.6.1	<i>Data Cleaning</i> dan <i>Case Folding</i>	22
3.6.2	<i>Spelling Normalization</i>	23
3.6.3	<i>Tokenizing</i>	23
3.6.4	<i>Stemming</i>	24
3.6.5	<i>Stopword Removal</i>	24
3.7	<i>Term Weighting</i>	25
3.8	Klasifikasi dengan Model.....	25
3.8.1	Split Data.....	25
3.8.2	Klasifikasi <i>Support Vector Machine (SVM)</i>	25
3.8.3	Evaluasi Model.....	26
3.9	Visualisasi dan Asosiasi Hasil Klasifikasi.....	26
3.10	Alat Bantu Penelitian.....	27
3.11	Jadwal Rencana Penelitian	27
BAB IV HASIL DAN PEMBAHASAN.....		29
4.1	Pengumpulan Data	29
4.2	Pelabelan Data	29
4.3	Praproses Data	32
4.3.1	<i>Data Cleaning</i> dan <i>Case Folding</i>	32
4.3.2	<i>Spelling Normalization</i>	33

4.3.3	<i>Tokenizing</i>	35
4.3.4	<i>Stemming</i>	36
4.3.5	<i>Stopword Removal</i>	37
4.4	<i>Term Weighting</i>	39
4.5	Klasifikasi dengan Model.....	46
4.5.1	Split Data.....	46
4.5.2	Klasifikasi <i>Support Vector Machine</i> (SVM)	47
4.5.3	Evaluasi Model.....	49
4.6	Visualisasi dan Asosiasi Hasil Klasifikasi.....	51
4.6.1	Distribusi Sentimen.....	51
4.6.2	Sentimen Positif	51
4.6.3	Sentimen Negatif.....	55
BAB V KESIMPULAN DAN SARAN.....		60
5.1	Kesimpulan.....	60
5.2	Saran	61
DAFTAR PUSTAKA		62
DAFTAR RIWAYAT HIDUP.....		67
LAMPIRAN.....		68

DAFTAR GAMBAR

Gambar 3.1 Tahapan Penelitian	20
Gambar 4.1 Hasil Scraping Data Komentar YouTube	29
Gambar 4.2 Distribusi Sentimen	32
Gambar 4.3 Hasil Komputasi Perhitungan TF-IDF	46
Gambar 4.4 Visualisasi Confusion Matrix	50
Gambar 4.5 Visualisasi Dashboard	51
Gambar 4.6 WordCloud Sentimen Positif.....	52
Gambar 4.7 Visualisasi Word Frequent Sentimen Positif	53
Gambar 4.8 Network Diagram Asosiasi Kata (Sentimen Positif).....	54
Gambar 4.9 WordCloud Sentimen Negatif	56
Gambar 4.10 Visualisasi Word Frequent Sentimen Negatif.....	57
Gambar 4.11 Network Diagram Asosiasi Kata (Sentimen Negatif)	58

DAFTAR TABEL

Tabel 2.1 Model Confusion Matrix	12
Tabel 2.2 Penelitian Terdahulu	15
Tabel 3.1 Proses Pelabelan Data	22
Tabel 3.2 Proses Data Cleaning dan Case Folding	23
Tabel 3.3 Proses Spelling Normalization	23
Tabel 3.4 Proses Tokenizing	23
Tabel 3.5 Proses Stemming	24
Tabel 3.6 Proses Stopword Removal	24
Tabel 3.7 Jadwal Rencana Penelitian	28
Tabel 4.1 Hasil Pelabelan Data	30
Tabel 4.2 Demografi Anotator	31
Tabel 4.3 Hasil Data Cleaning dan Case Folding	32
Tabel 4.4 Hasil Spelling Normalization	34
Tabel 4.5 Hasil Tokenizing	35
Tabel 4.6 Hasil Stemming	36
Tabel 4.7 Hasil Stopword Removal	38
Tabel 4.8 Dokumen Sampel	39
Tabel 4.9 Nilai TF	40
Tabel 4.10 Nilai IDF	42
Tabel 4.11 Nilai TF-IDF	43
Tabel 4.12 Nilai TF-IDF Normalisasi	45
Tabel 4.13 Distribusi Data dalam Penggunaan SMOTE	47
Tabel 4.14 Parameter Terbaik GridSearchCV	47
Tabel 4.15 Perbandingan hasil K-Fold Cross Validation	48
Tabel 4.16 Perbandingan Hasil Evaluasi Model	49
Tabel 4.17 Asosiasi Kata Sentimen Positif	53
Tabel 4.18 Asosiasi Kata Sentimen Negatif	57

DAFTAR RUMUS

Rumus (1) <i>Term Frequency</i> (TF) Murni	10
Rumus (2) <i>Term Frequency</i> (TF) Logaritmik	10
Rumus (3) <i>Term Frequency</i> (TF) Normalisasi	11
Rumus (4) <i>Invers Document Frequency</i> (IDF)	11
Rumus (5) <i>Term Frequency & Inverse Document Frequency</i> (TF-IDF)	11
Rumus (6) Akurasi	13
Rumus (7) Presisi	13
Rumus (8) <i>Recall</i>	14
Rumus (9) <i>F1-Score</i>	14

DAFTAR LAMPIRAN

Lampiran 1. Hasil Turnitin.....	68
Lampiran 2. Kamus Normalisasi	80
Lampiran 3. Source Code.....	81
Lampiran 4. Hasil Pelabelan	89