

**ANALISIS SENTIMEN PUBLIK TERHADAP KEBIJAKAN
EFISIENSI ANGGARAN PENDIDIKAN MENGGUNAKAN
ALGORITME SUPPORT VECTOR MACHINE (SVM)**

Tugas Akhir

Diajukan untuk memenuhi persyaratan mencapai derajat Sarjana Sistem
Informasi



Cecilia Aprillia Nathania

NPM: 211711533

PROGRAM STUDI SISTEM INFORMASI

FAKULTAS TEKNOLOGI INDUSTRI

UNIVERSITAS ATMA JAYA YOGYAKARTA

2025

HALAMAN PENGESAHAN

Tugas Akhir Berjudul

ANALISIS SENTIMEN PUBLIK TERHADAP KEBIJAKAN EFISIENSI ANGGARAN PENDIDIKAN
MENGUNAKAN ALGORITME SUPPORT VECTOR MACHINE (SVM)

yang disusun oleh

Cecilia Aprillia Nathania

211711533

dinyatakan telah memenuhi syarat pada tanggal 11 Juni 2025

Dosen Pembimbing 1	: Generosa Lukhayu Pritalia, S.T., M.Eng.	Keterangan Telah Menyetujui
Dosen Pembimbing 2	: Elisabeth Marsella, S.S., M.Li.	Telah Menyetujui
Tim Penguji		
Penguji 1	: Generosa Lukhayu Pritalia, S.T., M.Eng.	Telah Menyetujui
Penguji 2	: Hendro Gunawan, S.Si., M.T.	Telah Menyetujui
Penguji 3	: Dr. Fl. Spty Rahayu, S.T., M.Kom.	Telah Menyetujui

Yogyakarta, 11 Juni 2025

Universitas Atma Jaya Yogyakarta

Teknologi Industri

Dekan

ttd.

Dr. Ir. Parama Kartika Dewa SP., S.T., M.T.

Dokumen ini merupakan dokumen resmi UAJY yang tidak memerlukan tanda tangan karena dihasilkan secara elektronik oleh Sistem Bimbingan UAJY. UAJY bertanggung jawab penuh atas informasi yang tertera di dalam dokumen ini

LEMBAR PERNYATAAN KEASLIAN PENELITIAN

Saya yang bertanda tangan di bawah ini:

Nama Lengkap : Cecilia Aprillia Nathania

NPM : 211711533

Program Studi : Sistem Informasi

Fakultas : Teknologi Industri

Judul Penelitian : ANALISIS SENTIMEN PUBLIK TERHADAP KEBIJAKAN
EFISIENSI ANGGARAN PENDIDIKAN MENGGUNAKAN ALGORITME SUPPORT
VECTOR MACHINE (SVM)

Menyatakan dengan ini:

1. Skripsi ini adalah benar merupakan hasil karya sendiri dan tidak merupakan salinan sebagian atau keseluruhan dari karya orang lain.
2. Memberikan kepada Universitas Atma Jaya Yogyakarta, berupa Hak Bebas Royalti non eksklusif (*Non-Exclusive-Royalty-Free Right*) atas Penelitian ini, dan berhak menyimpan, mengelola dalam pangkalan data, mendistribusikan, serta menampilkan untuk kepentingan akademis, tanpa perlu meminta izin selama tetap mencantumkan nama peneliti.
3. Bersedia menanggung secara pribadi segala bentuk tuntutan hukum yang mengikuti atas pelanggaran Hak Cipta dalam pembuatan Skripsi ini.

Demikianlah pernyataan ini dibuat dan dapat dipergunakan sebagaimana mestinya.

Yogyakarta, 11 Juni 2025

Yang menyatakan,

Cecilia Aprillia Nathania

211711533

PRAKATA

Puji dan syukur peneliti panjatkan ke hadirat Tuhan Yang Maha Esa atas segala rahmat dan karunia-Nya sehingga peneliti dapat menyelesaikan tugas akhir yang berjudul “Analisis Sentimen Publik Terhadap Kebijakan Efisiensi Anggaran Pendidikan Menggunakan Algoritme Support Vector Machine (SVM)” dengan baik.

Tugas akhir ini disusun sebagai salah satu syarat untuk mencapai derajat Sarjana Sistem Informasi pada Fakultas Teknologi Industri Universitas Atma Jaya Yogyakarta. Dalam penyusunannya, peneliti menyadari bahwa tanpa bantuan, bimbingan, dan dukungan dari berbagai pihak, tugas akhir ini tidak akan terselesaikan dengan baik. Oleh karena itu, dengan segala kerendahan hati peneliti menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Ibu Generosa Lukhayu Pritalia, S.T., M.Eng, selaku dosen pembimbing 1 yang telah memberikan arahan, bimbingan, dan motivasi selama proses penyusunan tugas akhir ini.
2. Ibu Elisabeth Marsella, SS., M.Li, selaku dosen pembimbing 2 yang telah memberikan masukan dan evaluasi yang sangat berarti selama proses penyusunan tugas akhir ini.
3. Kedua orang tua dan keluarga tercinta, yang selalu memberikan doa dan dukungan yang tiada henti selama proses perkuliahan hingga penyusunan tugas akhir ini.
4. Pacar tersayang, yang telah menjadi sumber semangat, pengertian, dan dukungan moral selama proses pengerjaan tugas akhir ini, baik di saat senang maupun saat menghadapi kesulitan.
5. Teman-teman, yang telah memberikan semangat, bantuan, dan kebersamaan selama menjalani masa perkuliahan hingga proses penyusunan tugas akhir ini.

Peneliti menyadari bahwa tugas akhir ini masih jauh dari sempurna. Oleh karena itu, peneliti terbuka menerima kritik dan saran yang membangun dari berbagai pihak. Besar harapan peneliti agar karya ini dapat memberikan manfaat bagi semua pihak yang berkepentingan.

Yogyakarta, 11 Juni 2025

Peneliti,

Cecilia Aprillia Nathania

211711533

INTISARI

Kebijakan efisiensi anggaran pendidikan yang ditetapkan oleh pemerintah menuai beragam respons dari masyarakat di media sosial. Kurangnya pemetaan sentimen yang jelas menyulitkan pemahaman terhadap persepsi publik atas kebijakan tersebut. Penelitian ini bertujuan untuk menganalisis opini publik mengenai kebijakan efisiensi anggaran pendidikan berdasarkan unggahan di platform X. Data dianalisis menggunakan algoritme Support Vector Machine (SVM) untuk mengklasifikasikan sentimen menjadi tiga kategori yaitu positif, negatif, dan netral. Evaluasi model dilakukan menggunakan metrik akurasi, *precision*, *recall*, dan *f1-score*. Hasil penelitian menunjukkan bahwa kernel RBF memberikan performa terbaik dengan akurasi tertinggi sebesar 89.41%. Selain itu, sentimen negatif mendominasi dalam persepsi masyarakat terhadap kebijakan efisiensi anggaran pendidikan.

Kata kunci: Analisis Sentimen, Efisiensi Anggaran Pendidikan, Media Sosial, SVM

ABSTRACT

The government's education budget efficiency policy has elicited various responses from the public on social media. The lack of clear sentiment mapping makes it difficult to understand public perception of the policy. This study aims to analyze public opinion regarding the education budget efficiency policy based on posts on platform X. The data were analyzed using the Support Vector Machine (SVM) algorithm to classify sentiments into three categories: positive, negative, and neutral. Model evaluation was conducted using accuracy, precision, recall, and F1-score metrics. The results showed that the RBF kernel provided the best performance with the highest accuracy of 89.41%. Additionally, negative sentiment dominated public perception of the education budget efficiency policy.

Keywords: Sentiment Analysis, Education Budget Efficiency, Social Media, SVM

DAFTAR ISI

LEMBAR PERNYATAAN KEASLIAN PENELITIAN.....	iii
PRAKATA	iv
INTISARI	vi
ABSTRACT.....	vii
DAFTAR ISI	viii
DAFTAR GAMBAR	xi
DAFTAR TABEL	xi
DAFTAR PERBAIKAN.....	xii
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Pertanyaan Penelitian.....	3
1.4. Batasan Masalah.....	3
1.5. Tujuan Penelitian	4
1.6. Manfaat Penelitian	4
1.7. Bagan Keterkaitan.....	5
BAB II TINJAUAN PUSTAKA.....	6
2.1. Studi Sebelumnya	6
2.2. Dasar Teori.....	10
2.2.1. Anggaran Pendidikan.....	10
2.2.2. Media Sosial	10
2.2.3. X.....	10
2.2.4. <i>Artificial Intelligence</i>	11
2.2.5. <i>Machine learning</i>	11
2.2.6. <i>Text Mining</i>	11
2.2.7. Analisis Sentimen	11
2.2.8. <i>Data Crawling</i>	12
2.2.9. <i>Data Preprocessing</i>	12
2.2.10. Pelabelan Data	12
2.2.11. <i>Term Frequency Inverse Document Frequency (TF-IDF)</i>	13

2.2.12.	Penanganan <i>Imbalanced Data</i>	14
2.2.13.	Klasifikasi	14
2.2.14.	Support Vector Machine	15
2.2.15.	<i>Confusion Matrix</i>	16
BAB III METODOLOGI PENELITIAN		19
3.1.	Tahapan Penelitian	20
3.2.	Metode Penelitian	21
3.2.1.	Studi Literatur.....	21
3.2.2.	Pengumpulan Data	21
3.2.3.	Pelabelan Data	21
3.2.4.	<i>Preprocessing Data</i>	22
3.2.5.	<i>Feature Extraction</i>	23
3.2.6.	Pembagian <i>Dataset</i>	24
3.2.7.	Pemodelan	24
3.2.8.	Evaluasi Model	25
3.2.9.	Interpretasi Hasil	25
3.3.	Alat yang digunakan	26
3.3.1.	Python	26
3.3.2.	Google Colaboratory	26
BAB IV HASIL DAN PEMBAHASAN.....		27
4.1.	Pengumpulan Data	27
4.1.1.	Pelabelan Data	30
4.2.	<i>Preprocessing Data</i>	34
4.2.1.	<i>Case Folding</i>	35
4.2.2.	<i>Punctuation Removal</i>	36
4.2.3.	Normalisasi Teks	37
4.2.4.	<i>Repetition Character Removal</i>	38
4.2.5.	<i>Stopword Removal</i>	39
4.2.6.	<i>Stemming</i>	40
4.3.	Feature Extraction	44
4.4.	Pembagian <i>Dataset</i>	46
4.5.	Pemodelan dan Evaluasi Model	47
4.6.	Interpretasi Hasil	56

BAB V PENUTUP	70
5.1. Kesimpulan	70
5.2. Saran	70
Daftar Pustaka	72

DAFTAR GAMBAR

Gambar 1. 1. Bagan Keterkaitan	5
Gambar 3. 1. Diagram Alur Metodologi Penelitian	20
Gambar 3. 2. Support Vector Machine	25
Gambar 4. 1. Otorisasi Auth Token X	27
Gambar 4. 2. Proses Crawling Data.....	28
Gambar 4. 3. Dataset	28
Gambar 4. 4. Proses filter 1 komentar per 1 akun.....	29
Gambar 4. 5. Import Pustaka	34
Gambar 4. 6. Eksplorasi awal data	35
Gambar 4. 7. Case Folding	36
Gambar 4. 8. Punctuation Removal	37
Gambar 4. 9. Normalisasi Teks.....	38
Gambar 4. 10. Repetition Character Removal	39
Gambar 4. 11. Stopword Removal	40
Gambar 4. 12. Implementasi Fungsi <code>replace_words()</code>	40
Gambar 4. 13. Stemming	41
Gambar 4. 14. Dataset Setelah Preprocessing.....	43
Gambar 4. 15. Kode Implementasi TF-IDF	44
Gambar 4. 16. Kode Implementasi Pembagian Dataset	46
Gambar 4. 17. Kode Implementasi SMOTE.....	48
Gambar 4. 18. Implementasi dan Evaluasi Model dengan Kernel Linear.....	50
Gambar 4. 19. Implementasi dan Evaluasi Model dengan Kernel Polynomial.....	52
Gambar 4. 20. Implementasi dan Evaluasi Model dengan Kernel Sigmoid	53
Gambar 4. 21. Implementasi dan Evaluasi Model dengan Kernel RBF.....	55
Gambar 4. 22. Perbandingan Akurasi Tiap Kernel	56
Gambar 4. 23. Perbandingan Precision Tiap Kelas dan Kernal	57
Gambar 4. 24. Perbandingan Recall Tiap Kelas dan Kernel	58
Gambar 4. 25. Perbandingan F1-Score Tiap Kelas dan Kernel.....	58
Gambar 4. 26. Perbandingan Waktu Pelatihan Model	60
Gambar 4. 27. Word Cloud Sentimen Positif	61
Gambar 4. 28. Word Cloud Sentimen Negatif	62
Gambar 4. 29. Word Cloud Sentimen Netral	63
Gambar 4. 30. Distribusi Kelas Sentimen.....	64
Gambar 4. 31. Distribusi Kelas Sentimen Prediksi	68

DAFTAR TABEL

Tabel 2. 1. Perbandingan Studi Sebelumnya.....	8
Tabel 4. 1. Standar Klasifikasi Sentimen.....	30
Tabel 4. 2. Contoh Hasil Pelabelan Sentimen	32

Tabel 4. 3. Contoh Implementasi Hasil Preprocessing	42
Tabel 4. 4. Top 10 Kata Paling Berpengaruh.....	45
Tabel 4. 5. 5 Kata Paling Tidak Berpengaruh.....	46
Tabel 4. 6. Perbandingan Akurasi Beberapa Algoritme Klasifikasi.....	47
Tabel 4. 7. Contoh Hasil Prediksi Sentimen Pada Data Baru	65

DAFTAR PERBAIKAN

No.	Revisi	Halaman
1	Kata algoritma diganti menjadi algoritme	Telah mengubah kata algoritma menjadi algoritme di seluruh halaman
2	Menambahkan perbandingan algoritme sebelum memilih svm	Telah menambahkan perbandingan algoritme sebelum memilih svm pada halaman 47
3	Memastikan kualitas dari data seperti spam tweet, data yang redundan, atau akun yang <i>organic</i> atau bot dan memastikan satu komentar per satu akun	Telah menambahkan penjelasan mengenai kualitas data pada halaman 29-30 dan halaman 35
4	Menghapus saran untuk pemerintah pada bagian saran	Telah menghapus saran untuk pemerintah pada bagian saran pada halaman 70-71
5	Menambahkan heading untuk daftar pustaka	Telah menambahkan heading untuk daftar pustaka pada halaman 72