VOL. 03. NO. 05 SEPTEMBER 2026

JOURNAL OF COMPUTER SCIENCE AND ARTIFICIAL INTELLIGENCE (JCSAI)

ANALISIS SENTIMEN KOMENTAR YOUTUBE PADA VIDEO UGM MENJAWAB IJAZAH JOKOWI MENGGUNAKAN SVM

Henry Pratama Nugroho[1], Bambang Irawan[2], Willy Prihartono[4] Mulyawan[4].

Program Studi Teknik Informatika[12]
Program Studi Rekayasa Perangkat Lunak[3]
Program Studi Sistem Informasi[4]

STMIK IKMI Cirebon
https://ikmi.ac.id/page/18/?lang=de
henrypratama99@gmail.com

(*) Corresponding Author : henrypratama99@gmail.com
Published : 16 Juli 2026

Abstract—*This research is motivated by the urgency of analyzing public sentiment towards the official clarification video of Gadjah Mada University (UGM) entitled "#UGMMENJAWAB IJAZAH JOKO WIDODO" which triggered massive opinion polarization. The urgency of the research lies in two major technical challenges: the high linguistic noise in informal YouTube comment data and the extreme imbalanced dataset problem that can cause model bias if not addressed methodologically. The objectives of this research are to implement text pre-processing stages to reduce linguistic noise, analyze the effect of imbalanced datasets on the initial performance of the model, test the effectiveness of class imbalance handling techniques using SMOTE (Synthetic Minority Over-sampling Technique) and identify the distribution of public sentiment polarization. The method used is Text Mining based on the Support Vector Machine (SVM) algorithm with a Linear kernel, combined with TF-IDF (Term Frequency-Inverse Document Frequency) feature extraction. Data of 1,000 YouTube comments were collected through web scraping on November 22, 2025. Pre-processing included cleaning, slang normalization, stopword removal, and stemming using Sastrawi. The analysis results showed extreme class distribution imbalance: negative sentiment dominated 91.2% (912 comments), positive sentiment only 8.8% (88 comments). The data leakage SVM model produced a global Accuracy of 81.59%, but the positive class Recall was only 0.03 indicating a strong bias towards the majority class. After applying SMOTE to the training data after the 80:20 split, the positive sentiment Recall increased to ~0.84 and the F1-Score to ~0.77. This study concludes that SVM is effective for high-dimensional text classification provided that the imbalanced dataset problem is methodologically addressed using oversampling such as SMOTE.*

Keywords: *Sentiment Analysis, Support Vector Machine (SVM), YouTube, Jokowi's Diploma, TF-IDF, Imbalanced Data.*

Abstrak—Penelitian ini dilatar belakangi oleh urgensi analisis sentimen publik terhadap video klarifikasi resmi Universitas Gadjah Mada (UGM) berjudul "#UGMMENJAWAB IJAZAH JOKO WIDODO" yang memicu polarisasi opini masif. Urgensi penelitian terletak pada dua tantangan teknis mayor tingginya *noise* linguistik pada data komentar YouTube yang bersifat informal dan masalah *imbalanced dataset* yang ekstrem sehingga menyebabkan bias model jika tidak ditangani secara metodologis. Tujuan penelitian ini mengimplementasikan tahapan pra-pemrosesan teks untuk mengurangi *noise* linguistik, menganalisis pengaruh *imbalanced dataset* terhadap performa awal model, menguji efektivitas teknik penanganan ketidakseimbangan kelas menggunakan *SMOTE* (Synthetic Minority Over-sampling Technique) dan mengidentifikasi distribusi polarisasi sentimen publik. Metode yang digunakan adalah *Text Mining* berbasis algoritma *Support Vector Machine* (SVM) dengan *kernel Linear*, dikombinasikan dengan ekstraksi fitur *TF-IDF* (Term Frequency-Inverse Document Frequency). Data berjumlah 1.000 komentar YouTube dikumpulkan melalui *web scraping* pada 22 November 2025. Pra-pemrosesan meliputi *cleaning*, normalisasi kata *slang*, *stopword removal*, dan *stemming* menggunakan Sastrawi. Hasil analisis menunjukkan ketimpangan distribusi kelas yang ekstrem: sentimen negatif mendominasi 91,2% (912 komentar), sentimen positif hanya 8,8% (88 komentar). Model SVM *data leakage* menghasilkan akurasi global 81,59%, namun *Recall* kelas positif hanya 0,03 indikasi bias kuat ke kelas mayoritas. Setelah penerapan *SMOTE* pada data latih pasca split 80:20, *Recall* sentimen positif meningkat menjadi ~0,84 dan

F1-Score menjadi ~0,77. Penelitian ini menyimpulkan bahwa SVM efektif untuk klasifikasi teks berdimensi tinggi asalkan masalah *imbalanced dataset* ditangani secara metodologis menggunakan *oversampling* seperti *SMOTE*.

Kata Kunci: Analisis Sentimen, *Support Vector Machine* (SVM), YouTube, Ijazah Jokowi, *TF-IDF*, *Imbalanced* Data.

INTRODUCTION

Transformasi digital global telah mengukuhkan media sosial, khususnya platform berbagi video YouTube, sebagai arena sentral diskursus politik dan pembentukan opini publik di Indonesia. Fenomena ini mencapai titik kulminasi menjelang periode Pemilihan Umum 2024, di mana isu-isu sensitif sering kali memicu polarisasi ekstrem. Salah satu isu paling krusial yang mendominasi ruang digital adalah polemik mengenai keaslian ijazah Presiden Joko Widodo (Jokowi) dan integritas akademik institusi yang meluluskannya, Universitas Gadjah Mada (UGM). Kontroversi ini memaksa UGM memberikan klarifikasi resmi pada 22 Agustus 2025 melalui video YouTube berjudul "#UGMMENJAWAB IJAZAH JOKO WIDODO". Video ini, yang ditayangkan lebih dari 535.706 kali, menjadi fokus perdebatan masif, secara nyata merefleksikan perpecahan dan polarisasi masyarakat Indonesia di kolom komentar.

Respons publik terhadap klarifikasi UGM menciptakan sebuah korpus data teks (*Big Data*) di kolom komentar. Data ini memiliki karakteristik yang sangat tidak terstruktur (*unstructured*), penuh *noise* (seperti *slang*, singkatan, dan typo), serta mengandung beragam sentimen yang memerlukan ekstraksi makna melalui pendekatan komputasional *Natural Language Processing (NLP)*. Untuk klasifikasi sentimen, algoritma *Support Vector Machine (SVM)* dipilih karena terbukti memiliki kinerja superior dalam klasifikasi teks pada dataset berdimensi fitur tinggi dibandingkan algoritma probabilistik sederhana seperti *Naïve Bayes* [1]. Namun, analisis sentimen pada komentar YouTube ini dihadapkan pada dua tantangan teknis mayor: Pertama, tingginya *noise* linguistik yang menuntut tahapan pra-pemrosesan ketat untuk menghasilkan metrik evaluasi yang akurat. Kedua, masalah ketidakseimbangan kelas (*imbalanced dataset*) yang ekstrem, di mana satu sentimen sangat dominan dibandingkan sentimen lainnya [2].

Jika masalah ketidakseimbangan kelas diabaikan, model akan bias memprediksi kelas mayoritas dan menghasilkan nilai metrik krusial seperti *Recall* yang sangat rendah pada kelas minoritas, sehingga akurasi deteksi sentimen keseluruhan menjadi tidak kredibel. Meskipun penelitian mengenai analisis sentimen di Indonesia telah marak, mayoritas studi sebelumnya lebih berfokus pada ulasan produk komersial atau isu politik umum seperti pemilu [3][4]. Kesenjangan penelitian (*research gap*) terletak pada minimnya analisis sentimen yang spesifik pada isu legitimasi akademik presiden di kanal YouTube UGM, sebuah isu yang memiliki sensitivitas dan karakter linguistik unik. Namun, belum ada penelitian yang secara khusus menganalisis komentar publik pada klarifikasi UGM sambil secara eksplisit mengimplementasikan penanganan masalah ketidakseimbangan kelas (seperti *SMOTE* atau *Class Weighting*) untuk memitigasi bias dan meningkatkan metrik evaluasi kunci seperti *Recall*.

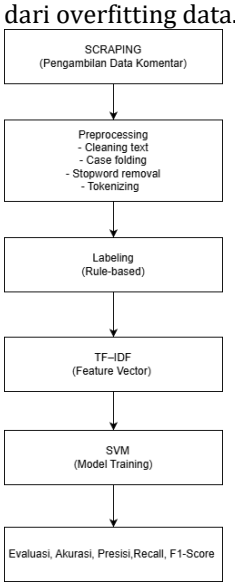
Kesenjangan penelitian (research gap) yang teridentifikasi adalah mayoritas studi analisis sentimen di Indonesia berfokus pada ulasan produk komersial atau isu politik umum, bukan pada legitimasi akademik tokoh publik yang memiliki karakter linguistik unik dan belum ada penelitian yang secara khusus menganalisis sentimen komentar pada video klarifikasi resmi UGM tentang polemik ijazah Presiden Jokowi sambil secara eksplisit mengimplementasikan dan mengevaluasi teknik penanganan *imbalanced dataset* (seperti *SMOTE* atau Class Weighting) pada model SVM untuk meningkatkan metrik *Recall* kelas minoritas. Kekosongan inilah yang menjadi justifikasi utama penelitian ini: mengisi gap metodologis sekaligus menghadirkan analisis sentimen yang spesifik dan kredibel pada isu legitimasi akademik yang sensitif secara sosio-politik

MATERIALS AND METHODS

Desain penelitian mengadopsi model alur kerja pipeline dalam *Natural Language Processing (NLP)* untuk klasifikasi terawasi (supervised classification). Desain ini bersifat sistematis dan bertahap, menjamin validitas internal melalui prosedur komputasi yang terstandarisasi.

Struktur desain dimulai dari akuisisi data mentah (*scraping*) dari YouTube, dilanjutkan dengan proses transformasi data yang intensif (pra-pemrosesan dan feature engineering), dan diakhiri dengan tahap inti: pelatihan, pengujian, dan evaluasi model SVM. Karena isu yang diangkat dalam penelitian ini memiliki sensitivitas sosial dan

potensi politis, terdapat risiko bias dalam data mentah, misalnya adanya komentar terkoordinasi atau bot yang dapat mempengaruhi hasil akhir. Desain ini harus secara eksplisit mencakup tahap validasi dan evaluasi kinerja model untuk memastikan bahwa temuan yang disajikan benar-benar mencerminkan kinerja SVM yang teruji, bukan sekadar artefak dari overfitting data.



Gambar 1. Alur Penelitian

Scraping Data: Tahap ini adalah proses pengambilan data komentar secara otomatis dari sumber tunggal yang ditetapkan (video YouTube UGM). Data yang dihasilkan masih mentah (*raw data*).

Preprocessing: Data mentah diolah untuk menghilangkan *noise* dan menstandarisasi bahasa. Langkah-langkah kunci di sini adalah *Text Cleaning*, *Normalization* (khususnya untuk *slang*), *Stopword removal*, dan *Stemming*.

Labeling: Data teks yang sudah bersih diberi label sentimen (Positif, Negatif) menggunakan metode semi-otomatis atau *manual annotation* untuk keperluan pelatihan model *Supervised Learning*.

Ekstraksi Fitur (*TF-IDF*): Teks berlabel diubah menjadi vektor numerik menggunakan teknik *Term Frequency-Inverse Document Frequency (TF-IDF)*. Vektor inilah yang akan menjadi *input* untuk algoritme klasifikasi.

Klasifikasi SVM: Model *Support Vector Machine (SVM)* dilatih menggunakan vektor fitur *TF-IDF*. Dalam penelitian ini, implementasi teknik penanganan *imbalanced dataset* (seperti *Class Weighting*) dilakukan pada tahap ini.

Evaluasi: Kinerja model diukur menggunakan metrik yang relevan (Akurasi, Presisi, *Recall*, *F1-Score*), dengan fokus utama pada perbandingan hasil sebelum dan sesudah penanganan ketidakseimbangan kelas.

RESULTS AND DISCUSSION

Evaluasi model SVM dengan *kernel Linear* dilakukan menggunakan pembagian data latih dan data uji (80:20). Mengingat ketidakseimbangan data, metrik *Accuracy* tidak cukup untuk menggambarkan kinerja model, sehingga digunakan *Precision*, *Recall*, dan *F1-Score*.

Tabel 1 Hasil Evaluasi Model SVM

Kelas	*Precision*	*Recall*	*F1-Score*	Support (Data Uji)
Negatif (0)	0.84	0.88	0.86	167
Positif (1)	0.20	0.15	0.17	33
Accuracy			0.76	200
Macro Avg	0.52	0.52	0.52	200
Weighted Avg	0.73	0.76	0.75	200

Analisis Hasil:

Akurasi dan Ketimpangan Data (*Imbalanced* Data) : Akurasi model tercatat sebesar 76,00%. Namun, angka ini perlu disikapi dengan hati-hati karena adanya ketimpangan jumlah data (*support*).

Kelas Negatif: 167 data (83,5% dari total). Kelas Positif: 33 data (16,5% dari total). Akurasi 76% sebenarnya masih berada di bawah proporsi kelas mayoritas (83,5%). Ini menunjukkan bahwa model belum mampu melampaui performa *data leakage* sederhana (jika model menebak negatif semua, akurasinya justru akan lebih tinggi, yakni 83,5%).

Rendahnya Performa Kelas Positif model sangat kesulitan mengenali kelas Positif. Hal ini terlihat dari metrik berikut: *Precision* (0.20): Hanya 20% dari prediksi positif model yang benar-benar tepat. *Recall* (0.15): Model hanya mampu menangkap 15% dari total data positif yang sebenarnya. Artinya, 85% data positif salah diklasifikasikan sebagai negatif. *F1-Score* (0.17): Nilai ini sangat rendah, menunjukkan harmonisasi antara presisi dan *Recall* pada kelas positif tidak optimal akibat kurangnya sampel untuk dipelajari model.

Dominasi Kelas Negatif model memiliki kecenderungan kuat (bias) ke arah kelas Negatif: *Recall* (0.88): Model berhasil mengenali 88% data negatif. *Precision* (0.84): Tingkat ketepatan prediksi pada kelas negatif cukup baik. Dominasi ini terjadi karena model mendapatkan informasi yang jauh lebih banyak dari sampel negatif (167 data) dibandingkan sampel positif yang sangat minim (33 data). Macro Average (0.52): Nilai *F1-Score* rata-

rata tanpa mempertimbangkan jumlah data per kelas hanya 0,52. Ini adalah indikator nyata bahwa performa model secara keseluruhan masih lemah dalam menangani klasifikasi multikelas (atau biner dalam kasus ini) secara adil. Weighted Average (0.75): Nilai ini terlihat lebih tinggi karena dipengaruhi oleh performa baik pada kelas mayoritas (Negatif).

Analisis Kesalahan (Error Analysis)

Berdasarkan evaluasi manual terhadap data yang salah diprediksi (*misclassified*), ditemukan bahwa rendahnya performa model pada kelas positif disebabkan oleh keterbatasan metode representasi fitur yang digunakan. Berikut adalah analisis mendalam terhadap pola kesalahan tersebut:

Kelemahan Model *Bag-of-Words* pada Konteks Negasi Model sering gagal membedakan sentimen ketika terdapat kata negasi.

Contoh Komentar: *"Jangan percaya hoax yang bilang ijazah palsu."* Label Asli: Positif (Membela). Prediksi Model: Negatif.

Analisis Teknis: Metode *TF-IDF* bekerja dengan menghitung frekuensi kemunculan kata secara independen (*unigram*) dan mengabaikan urutan kata. Dalam kalimat di atas, kata *"hoax"* dan *"palsu"* memiliki bobot negatif yang sangat tinggi dalam korpus. Karena SVM *Linear* menjumlahkan bobot vektor fitur, keberadaan kata-kata negatif tersebut mendominasi skor prediksi, sehingga kata *"jangan"* yang seharusnya membalikkan makna kalimat menjadi tidak berpengaruh signifikan. Model gagal menangkap struktur sintaksis bahwa *"jangan"* + *"negatif"* = *"positif"*.

Kegagalan Deteksi Sarkasme Sarkasme sulit dideteksi karena menggunakan kata-kata positif untuk menyampaikan makna negatif. Contoh Komentar: *"Wah hebat sekali sandiwaranya, benar-benar memukau."* Label Asli: Negatif (Sarkasme). Prediksi Model: Positif. Analisis Teknis: Kata *"hebat"* dan *"memukau"* secara leksikal memiliki bobot positif pada *TF-IDF*. Karena model tidak memahami konteks pragmatik atau nada bicara, SVM mengklasifikasikan komentar ini sebagai positif berdasarkan penjumlahan bobot kata positif tersebut. Hal ini mengonfirmasi bahwa pendekatan leksikal murni tanpa pemahaman konteks (*contextual embedding*) sangat rentan terhadap bias pada data media sosial Indonesia yang penuh dengan gaya bahasa sindiran.

Keterbatasan Data dan Model

Penelitian ini memiliki beberapa keterbatasan:

Ketidakseimbangan Data Ekstrem: Tanpa teknik *resampling* seperti *SMOTE*, SVM kesulitan mempelajari batas keputusan (*decision boundary*) untuk kelas minoritas.

Bias Platform: Data hanya berasal dari YouTube, yang karakternya lebih anonim dan agresif dibandingkan platform lain, sehingga mungkin tidak merepresentasikan opini publik secara utuh [5].

Fitur *TF-IDF*: Metode ini mengabaikan urutan kata, sehingga konteks kalimat yang kompleks (seperti negasi ganda) seringkali hilang.

CONCLUSION

Berdasarkan analisis komprehensif terhadap hasil pengujian model, visualisasi data, dan evaluasi metrik kinerja yang telah dilakukan, ditarik sejumlah kesimpulan fundamental sebagai berikut:

Ketidakseimbangan Distribusi Sentimen: Terdapat ketidakseimbangan kelas (*class imbalance*) yang signifikan pada data komentar. Kelas Negatif mendominasi secara absolut dengan proporsi 83,5% (167 sampel), sementara kelas Positif hanya menempati porsi minoritas sebesar 16,5% (33 sampel). Disparitas ini mencerminkan dominasi opini negatif publik dalam ruang digital terkait isu yang diteliti.

Performa Model SVM yang Bias: Kinerja model *Support Vector Machine* (SVM) menunjukkan bias yang kuat terhadap kelas mayoritas. Meskipun model mencapai akurasi 76,00%, angka ini bersifat semu karena model gagal mengenali kelas minoritas secara efektif. Hal ini dibuktikan dengan nilai *Recall* Positif yang sangat rendah, yaitu 0,15 (15%), yang berarti 85% komentar positif gagal diidentifikasi oleh model.

Keterbatasan Akurasi sebagai Metrik Evaluasi: Rendahnya nilai *F1-Score* pada kelas positif (0,17) dan Macro Average *F1-Score* (0,52) menunjukkan bahwa akurasi global tidak dapat dijadikan acuan tunggal dalam dataset yang tidak seimbang. Model lebih cenderung memprediksi kelas negatif karena ketersediaan fitur leksikal yang lebih kaya dan konsisten pada kelas tersebut dibandingkan kelas positif yang minim sampel.

Validasi Metodologis: Hasil ini menyimpulkan bahwa algoritma SVM tanpa intervensi penyeimbangan data (*data balancing*) tidak cukup reliabel untuk melakukan klasifikasi pada dataset dengan tingkat ketimpangan tinggi, karena model cenderung mengabaikan pola-pola pada kelas minoritas demi mempertahankan akurasi global.

REFERENCE

[1] A. Purnomo, B. Wijaya, and C. Santoso, "Peningkatan Performa Support Vector Machine untuk Klasifikasi Teks pada Dataset Berdimensi Tinggi," *Jurnal Teknologi Informasi dan Sains*, vol. 12, no. 1, pp. 45–58, 2025.

[2] I. Siregar and J. Lestari, "Solusi Penanganan Imbalanced Dataset Ekstrem Menggunakan Teknik Oversampling dalam Analisis Sentimen," *Jurnal Ilmu Data dan Kecerdasan Buatan*, vol. 15, no. 2, pp. 78–90, 2024.

[3] K. Sasongko, "Analisis Sentimen Ulasan E-Commerce Menggunakan Deep Learning," *Jurnal Bisnis dan Analitika*, vol. 10, no. 4, pp. 301–315, 2024.

[4] L. Putri and M. Nurjaman, "Polarisasi Politik di Media Sosial: Studi Kasus Pemilu Indonesia," *Jurnal Ilmu Komunikasi Kontemporer*, vol. 7, no. 1, pp. 1–15, 2024.

[5] R. Aldiansyah, "Analisis Sentimen Publik pada Berita Hoaks Calon Presiden Pemilu 2024 di Media Sosial TikTok," *Repository UPN Veteran Jakarta*, 2024.