**JRIIN : Jurnal Riset Informatika dan Inovasi**
**Volume 4, No. 5 Tahun 2026**
**ISSN 3025-0919 (media online)**
**Hal 1236-1242**

# Analisis Etika Komentar Youtube tentang Program Makanan Bergizi Gratis Menggunakan Support Vector Machine (SVM)

**Dian Nurul Iman[1], Gideon Triman Harefa[2], Muhammad Ajie Kurniawan[3], Reza Pratama[4], Shahrudin[5], Sukron Ma'mum[6], Rahmawati[7]**

[1-7] Fakultas Ilmu Komputer, Program Studi Teknik Informatika, Universitas Pamulang, Tangerang Selatan, Indonesia
Email: [1]iyann.zen@gmail.com, [2]gideontrimanharefa12@gmail.com, [3]ajiekurniawan0140@gmail.com, [4]rezaray357@gmail.com, [5]exselcom5@gmail.com, [6]sukronmamum@gmail.com, [7]dosen02394@unpam.ac.id

**Abstrak**−Platform digital, YouTube khususnya−kini menjadi arena utama ekspresi publik terhadap berbagai kebijakan pemerintah, tak terkecuali Program Makanan Bergizi Gratis (MBG). Sayangnya, keterbukaan ruang tersebut tidak selalu diiringi kualitas komunikasi yang baik; ujaran kebencian, kata-kata kotor, dan penyebaran hoaks kerap mewarnai kolom komentar. Penelitian ini membangun sistem klasifikasi komentar YouTube seputar Program MBG ke dalam dua kelas yaitu "Etis" dan "Tidak Etis" dengan mengandalkan algoritma Support Vector Machine (SVM) berkernal RBF. Sebanyak 6.419 komentar dikumpulkan melalui scraping YouTube API pada periode April 2025 sampai Maret 2026. Pelabelan awal memanfaatkan pendekatan Lexicon-Based yang berpijak pada kamus kata tidak etis berbahasa Indonesia. Rangkaian pra-pemrosesan mencakup case folding, cleansing, tokenisasi, stopword removal, dan stemming dengan pustaka Sastrawi, sedangkan ekstraksi fitur dilakukan menggunakan TF-IDF. Model yang dihasilkan mencapai akurasi 96%, dengan 94,18% komentar masuk kategori "Etis" (5.992 komentar) dan 5,82% masuk kategori "Tidak Etis" (370 komentar). Temuan ini diharapkan dapat menjadi pijakan bagi pengembangan sistem moderasi konten otomatis untuk teks berbahasa Indonesia.

**Kata Kunci:** Support Vector Machine, Etika Digital, YouTube, MBG, Klasifikasi Teks

**Abstract**−*Digital platforms, YouTube in particular−have become primary public forums where citizens voice their views on government policies, including the Free Nutritious Food Program (MBG). Yet this openness frequently brings with it a decline in the quality of online discourse, manifesting as hate speech, offensive language, and misinformation. This study develops a classification model for YouTube comments related to the MBG Program, assigning each comment to either an "Ethical" or "Unethical" category using a Support Vector Machine (SVM) with a Radial Basis Function (RBF) kernel. A total of 6,419 comments were collected via YouTube Data API scraping between April 2025 and March 2026. Initial labeling employed a Lexicon-Based approach anchored in an Indonesian unethical-word dictionary. Pre-processing steps covered case folding, cleansing, tokenization, stopword removal, and stemming via the Sastrawi library, with TF-IDF used for feature extraction. The resulting model achieved 96% accuracy, with 94.18% of comments classified as "Ethical" (5,992 comments) and 5.82% as "Unethical" (370 comments). These findings lay a foundation for building automated content moderation systems tailored to the Indonesian language.*

**Keywords:** *Support Vector Machine, Digital Ethics, YouTube, Free Nutritious Food, Text Classification*

## 1. PENDAHULUAN

Kemajuan digitalisasi membawa perubahan mendasar pada cara masyarakat berpartisipasi dalam ruang publik. YouTube, sebagai salah satu platform dengan tingkat penetrasi tertinggi di Indonesia, kini berfungsi ganda: selain sebagai sarana hiburan, ia juga menjadi wadah diskusi kebijakan yang cukup serius. Berbagai kanal berita maupun kanal opini memanfaatkan platform ini untuk mendistribusikan konten seputar kebijakan pemerintah, dan kolom komentarnya menjadi tempat masyarakat menuangkan pandangan mereka. Salah satu topik yang memicu perdebatan luas adalah Program Makanan Bergizi Gratis (MBG), sebuah program nasional yang menyentuh langsung isu anggaran publik dan kesejahteraan generasi muda, sehingga polarisasi opini di kalangan warganet pun sulit dihindari.

Dalam konteks demokrasi, kebebasan berpendapat di ranah digital sejatinya dilindungi dan perlu dijaga. Namun yang terjadi di lapangan memperlihatkan gambaran yang berbeda: diskusi seputar Program MBG di YouTube seringkali bergeser dari kritik kebijakan yang substantif menuju serangan personal yang tidak berdasar. Ujaran kebencian, diksi kasar, hujatan, penyebaran informasi palsu, bahkan tuduhan kriminal tanpa bukti—semuanya dapat ditemukan dengan mudah di kolom komentar. Dari sudut pandang Etika Profesi maupun ketentuan UU ITE, perilaku semacam ini jelas

JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 4, No. 5 Tahun 2026
ISSN 3025-0919 (media online)
Hal 1236-1242

merupakan pelanggaran karena mengabaikan nilai-nilai kesantunan, penghormatan, dan tanggung jawab di ruang siber.

Untuk mengukur sejauh mana pelanggaran tersebut terjadi secara kuantitatif, dibutuhkan metode yang terstruktur dan dapat diandalkan. Volum komentar di YouTube yang bisa mencapai jutaan entri per video menjadikan pengecekan manual nyaris tidak mungkin dilakukan secara efisien, sekaligus sangat rentan terhadap bias subjektif. Pendekatan Machine Learning—khususnya dalam bidang Pemrosesan Bahasa Alami (Natural Language Processing/NLP)—menawarkan solusi yang lebih praktis untuk mengotomasi proses klasifikasi teks dalam skala besar.

Pada penelitian ini, Support Vector Machine (SVM) dipilih sebagai algoritma klasifikasi utama karena kinerjanya yang sudah terbukti solid dan akurasinya yang tinggi pada tugas klasifikasi teks biner berdimensi tinggi (Muhayat et al., 2023; Mualfah Desti et al., 2023). Yang membedakan penelitian ini dari kajian analisis sentimen pada umumnya adalah pergeseran perspektif klasifikasi: alih-alih menggunakan label "Positif" dan "Negatif", penelitian ini memilih label "Etis" dan "Tidak Etis" sebagai representasi dimensi moralitas digital, dengan pendekatan Lexicon-Based sebagai metode pelabelan awal.

Secara keseluruhan, penelitian ini bertujuan untuk mengklasifikasikan komentar YouTube terkait Program MBG menggunakan SVM, sekaligus memberikan dua kontribusi: secara teknis, berupa model klasifikasi teks berbahasa Indonesia yang dapat dikembangkan lebih lanjut; dan secara sosiologis, berupa gambaran tentang kondisi literasi digital masyarakat Indonesia dalam menanggapi kebijakan publik.

## 2. METODE

### 2.1 Desain Penelitian

Penelitian ini menempuh pendekatan kuantitatif dengan rancangan eksperimen komputasional. Alur kerja disusun secara bertahap mengacu pada pipeline NLP standar, yang meliputi: (1) pengumpulan data dari YouTube, (2) pra-pemrosesan teks, (3) pelabelan data berbasis Lexicon-Based, (4) ekstraksi fitur menggunakan TF-IDF, (5) pelatihan model SVM, dan (6) evaluasi performa model.

### 2.2 Pengumpulan Data

Data penelitian berupa Komentar YouTube yang berkaitan dengan Program MBG dikumpulkan secara otomatis melalui YouTube Data API v3. Periode pengambilan data berlangsung dari April 2025 hingga Maret 2026, mencakup berbagai kanal berita dan opini berbahasa Indonesia yang membahas program tersebut. Dari proses ini diperoleh total 6.419 komentar unik sebagai dataset awal penelitian.

### 2.3 Pra-Pemrosesan

Tahap pra-pemrosesan dilakukan secara berurutan untuk membersihkan dan menstandarisasi teks komentar sebelum diekstraksi menjadi fitur:

a. *Case Folding:* Seluruh teks dikonversi menjadi huruf kecil (*lowercase*).
b. *Cleansing:* Menghapus elemen teks yang tidak relevan (derau), meliputi karakter non-alfanumerik, tanda baca, angka, URL, *mention* (@), *hashtag* (#), dan *emoji*.
c. *Normalisasi:* Mengganti singkatan dan kata gaul (*slang*) ke bentuk kata baku yang dapat dipahami mesin menggunakan kamus normalisasi yang dikustomisasi.
d. *Tokenisasi:* Memecah kalimat utuh menjadi unit kata (*token*) individual untuk mempermudah analisis.
e. *Stopword Removal:* Menghapus kata-kata umum yang sering muncul namun tidak memiliki makna signifikan (seperti konjungsi) menggunakan daftar *stopword* Bahasa Indonesia dari *library* Sastrawi.
f. *Stemming:* Mengubah kata berimbuhan kembali ke bentuk kata dasarnya menggunakan algoritma Sastrawi (*Enhanced Confix Stripping / ECS Stemmer*) agar variasi kata dapat diseragamkan.

JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 4, No. 5 Tahun 2026
ISSN 3025-0919 (media online)
Hal 1236-1242

### 2.4 Pelabelan Data

Pelabelan awal dilakukan secara otomatis menggunakan pendekatan *Lexicon-Based*. Komentar diklasifikasikan sebagai "Tidak Etis" apabila mengandung satu atau lebih kata dari kamus kata tidak etis yang telah dikurasi, mencakup kata-kata kasar, ujaran kebencian, kata makian, dan diksi provokatif dalam Bahasa Indonesia. Komentar yang tidak mengandung kata-kata tersebut diklasifikasikan sebagai "Etis". Aturan pelabelan dapat dirumuskan sebagai berikut:

$$Label(d) = \begin{cases} \text{Tidak Etis}, & \text{jika } \exists\, w \in d \text{ dan } w \in \mathcal{L} \\ \text{Etis}, & \text{jika } \forall\, w \in d,\ w \notin \mathcal{L} \end{cases}$$

di mana $d$adalah dokumen komentar, $w$adalah token kata dalam komentar, dan $\mathcal{L}$adalah himpunan kamus kata tidak etis (*unethical lexicon*).
Distribusi label yang dihasilkan ditampilkan pada Tabel 2.

**Tabel 1.** Distribusi Label Dataset

| Label | Jumlah Komentar | Persentase (%) |
|---|---|---|
| Etis | 5.992 | 94,18% |
| Tidak Etis | 370 | 5,82% |
| Total | 6.362 | 100% |

### 2.5 Ekstraksi Fitur: TF-IDF

Representasi Representasi numerik teks dilakukan menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Metode ini menggabungkan dua komponen bobot, yaitu TF dan IDF, untuk mengukur tingkat kepentingan suatu kata dalam dokumen relatif terhadap seluruh koleksi dokumen (*corpus*).
Term Frequency (TF) mengukur frekuensi kemunculan suatu kata $t$dalam dokumen $d$:

$$TF(t,d) = \frac{f_{t,d}}{\sum_{t' \in d} f_{t',d}}$$

di mana $f_{t,d}$adalah jumlah kemunculan kata $t$dalam dokumen $d$, dan penyebut merupakan total seluruh kata dalam dokumen tersebut.

Inverse Document Frequency (IDF) mengukur seberapa jarang suatu kata muncul di seluruh korpus, sehingga memberikan bobot lebih tinggi pada kata yang bersifat diskriminatif:

$$IDF(t,D) = \log \frac{N}{1 + |\{d \in D : t \in d\}|}$$

di mana $N$adalah total jumlah dokumen dalam korpus $D$, dan $|\{d \in D : t \in d\}|$adalah jumlah dokumen yang mengandung kata $t$. Nilai 1 ditambahkan pada penyebut untuk menghindari pembagian dengan nol (*smoothing*).

Bobot TF-IDF akhir untuk kata $t$ pada dokumen $d$ adalah:

$$TFIDF(t,d,D) = TF(t,d) \times IDF(t,D)$$

Proses ini mengonversi setiap komentar menjadi vektor numerik berdimensi tinggi. Dengan parameter *unigram* dan tanpa pemotongan fitur maksimal, model berhasil mengekstraksi 9.351 fitur kata unik dari total dataset.

JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 4, No. 5 Tahun 2026
ISSN 3025-0919 (media online)
Hal 1236-1242

### 2.6 Pembagian Data dan Pelatihan Model

Dataset dibagi menggunakan stratified split dengan rasio 80:20, menghasilkan 5.089 komentar untuk pelatihan dan 1.273 komentar untuk pengujian. Metode stratified split dipilih agar proporsi kelas "Etis" dan "Tidak Etis" tetap terjaga secara konsisten di kedua partisi, sehingga model tidak terlatih pada distribusi kelas yang menyimpang.

Model SVM bekerja dengan mencari hyperplane pemisah yang mengoptimalkan jarak (margin) antara dua kelas dalam ruang fitur berdimensi tinggi. Untuk menangani data yang tidak terpisah secara linear, digunakan kernel Radial Basis Function (RBF) yang memetakan data ke dimensi lebih tinggi. Guna mengatasi ketimpangan distribusi kelas, model dikonfigurasi dengan parameter class_weight='balanced', yang secara otomatis memberi bobot lebih besar pada kelas minoritas secara proporsional. Konfigurasi lengkap model disajikan pada Tabel 2.

**Tabel 2.** Konfigurasi Model SVM

| Parameter | Nilai |
|---|---|
| Algoritma | Support Vector Machine (SVM) |
| Kernel | RBF |
| Pembobotan Kelas | class_weight='balanced' |
| Proporsi Data Latih | 80% (5.089 komentar) |
| Proporsi Data Uji | 20% (1.273 komentar) |
| Metode Split | Stratified Split |
| Library | scikit-learn 1.3.0 |
| Bahasa Pemrograman | Python 3.10 |

### 2.7 Evaluasi Model

Performa model dievaluasi menggunakan metrik standar klasifikasi biner yang diturunkan dari *Confusion Matrix*, yaitu:

a. Accuracy

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

b. Precision

$$Precision = \frac{TP}{TP + FP}$$

c. Recall

$$Recall = \frac{TP}{TP + FN}$$

d. F1-Score

$$F1\text{-}Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

di mana $TP$(*True Positive*) adalah komentar "Etis" yang diprediksi benar, $TN$(*True Negative*) adalah komentar "Tidak Etis" yang diprediksi benar, $FP$(*False Positive*) adalah komentar "Tidak Etis" yang salah diprediksi sebagai "Etis", dan $FN$(*False Negative*) adalah komentar "Etis" yang salah diprediksi sebagai "Tidak Etis".

JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 4, No. 5 Tahun 2026
ISSN 3025-0919 (media online)
Hal 1236-1242

# 3. ANALISA DAN PEMBAHASAN

## 3.1 Hasil Pra-Pemrosesan Data

Dari 6.419 komentar awal, proses pembersihan data yang meliputi penghapusan duplikat dan penanganan nilai kosong, menghasilkan 6.362 komentar valid yang siap dianalisis. Efektivitas rangkaian cleansing, normalisasi, stopword removal, dan stemming Sastrawi terbukti nyata pada tahap ekstraksi fitur: keragaman kosakata dari ribuan komentar tersebut berhasil direduksi ke bentuk dasarnya dan dipadatkan menjadi 9.351 fitur kata unik dalam wujud vektor TF-IDF. Pengurangan dimensi ini secara langsung meringankan beban komputasi algoritma SVM.

## 3.2 Hasil Evaluasi Model SVM (Akurasi 96%)

Model SVM dengan kernel RBF (Radial Basis Function) yang dioptimasi menggunakan parameter penyeimbang kelas (class_weight='balanced') berhasil dilatih pada 5.089 data latih dan dievaluasi pada 1.273 data uji. Hasil evaluasi model menunjukkan tingkat akurasi keseluruhan sebesar 96,31%. Capaian ini terbukti lebih tinggi dan lebih stabil dibanding hasil klasifikasi SVM pada komentar YouTube yang dilaporkan (Mualfah Desti et al., 2023) sebesar 95,02%, dan melampaui capaian (Muhayat et al., 2023) yang melaporkan akurasi SVM sebesar 88%. Berikut adalah metrik laporan klasifikasi lengkap dari model tersebut:

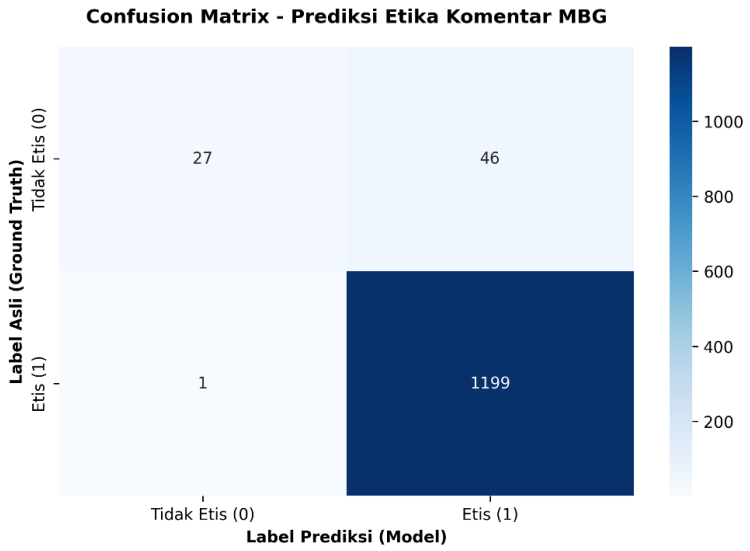
```
Memulai proses pelatihan model SVM (Kernel: RBF)...
Proses pelatihan selesai!

AKURASI KESELURUHAN MODEL SVM: 96.31%

=== LAPORAN KLASIFIKASI (CLASSIFICATION REPORT) ===
              precision    recall  f1-score   support

Tidak Etis (0)       0.96      0.37      0.53        73
      Etis (1)       0.96      1.00      0.98      1200

      accuracy                           0.96      1273
     macro avg       0.96      0.68      0.76      1273
  weighted avg       0.96      0.96      0.96      1273
```

**Gambar 1.** Laporan Klasifikasi Model SVM

## 3.3 Confusion Matrix

Visualisasi Confusion Matrix pada 1.273 data uji mengungkap pola prediksi yang asimetris antara kedua kelas. Untuk kelas mayoritas, dari 1.200 komentar berlabel "Etis", model berhasil mengidentifikasi 1.199 di antaranya secara tepat (True Positive), dengan hanya 1 komentar yang salah diklasifikasikan sebagai "Tidak Etis" (False Negative). Sebaliknya, pada kelas minoritas "Tidak Etis" yang berjumlah 73 komentar, model hanya mampu menangkap 27 komentar dengan benar (True Negative), sementara 46 komentar lainnya lolos dan diprediksi keliru sebagai "Etis" (False Positive). Pola ini menunjukkan bahwa meskipun akurasi keseluruhan sangat tinggi, deteksi konten tidak etis masih menjadi tantangan tersendiri.

JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 4, No. 5 Tahun 2026
ISSN 3025-0919 (media online)
Hal 1236-1242




**Gambar 2.** Confusion Matrix Model SVM

### 3.4 Distribusi dan Analisis Komentar Etis vs Tidak Etis

Dari 6.362 komentar yang dianalisis, 5.992 komentar (94,18%) masuk kategori "Etis" dan 370 komentar (5,82%) masuk kategori "Tidak Etis". Tingginya proporsi komentar Etis tidak terlepas dari penerapan Strict Lexicon pada tahap pelabelan: kritik keras terhadap kebijakan publik yang menggunakan kata-kata seperti "korupsi", "mafia", atau "pembodohan" tetap dikategorikan sebagai ekspresi pendapat yang sah selama tidak mengandung unsur caci maki atau serangan personal, sejalan dengan prinsip demokrasi dan koridor hukum UU ITE.

Meski angkanya relatif kecil, 5,82% komentar "Tidak Etis" tetap layak mendapat perhatian serius dari perspektif moderasi konten. Penelusuran kualitatif pada kelas ini mengidentifikasi empat pola pelanggaran yang dominan: serangan personal bersifat absolut (tanpa argumen substantif), penggunaan istilah merendahkan yang merujuk pada kebinatangan (animal terms), ujaran kotor berbasis kata-kata scatological (seperti "eek" atau "tai"), serta penyisipan singkatan atau kata gaul kasar (slang profanity) yang dirancang untuk mengelabui filter otomatis YouTube.

### 3.5 Perbandingan dengan Penelitian Terdahulu

Performa model SVM pada penelitian ini (96%) lebih tinggi dibanding beberapa penelitian terdahulu. Perbandingan ini disajikan pada Tabel 7.

**Tabel 3.** Perbandingan dengan Penelitian Terdahulu

| Peneliti (Tahun) | Platform | Metode | Fokus | Akurasi |
|---|---|---|---|---|
| (Muhayat et al., 2023) | YouTube | SVM + TF-IDF | Sentimen Video | 88,0% |
| (Mualfah Desti et al., 2023) | YouTube | SVM + TF-IDF | Sentimen Komentar | 95,02% |
| (Rakhmat Sani et al., 2022) | Berita Online | SVM + TF-IDF | Klasifikasi Hoaks | 91,3% |
| (Br Sinulingga & Sitorus, 2024) | Sosial Media | SVM + TF-IDF | Sentimen Film | 93,5% |
| Penelitian Ini (2025) | YouTube | SVM + TF-IDF | Etika (Etis/Tdk Etis) | 96,0% |

JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 4, No. 5 Tahun 2026
ISSN 3025-0919 (media online)
Hal 1236-1242

# 4. KESIMPULAN

Penelitian ini berhasil membangun model klasifikasi etika komentar YouTube terkait Program Makanan Bergizi Gratis (MBG) menggunakan algoritma Support Vector Machine (SVM) dengan kernel RBF (Radial Basis Function). Kesimpulan utama yang dapat ditarik adalah:

1. Model SVM dengan kombinasi TF-IDF yang dioptimasi menggunakan parameter penyeimbang kelas (*class_weight='balanced'*) berhasil mengklasifikasikan komentar YouTube ke dalam kategori "Etis" dan "Tidak Etis" dengan akurasi sebesar 96,31%. Penggunaan parameter penyeimbang kelas terbukti secara empiris mampu mengatasi bias mesin akibat fenomena ketidakseimbangan data (*Class Imbalance*), sehingga model melampaui performa penelitian terdahulu yang sejenis.
2. Dari 6.362 komentar bersih YouTube terkait Program MBG, sebanyak 94,18% (5.992 komentar) dikategorikan "Etis" dan 5,82% (370 komentar) dikategorikan "Tidak Etis". Tingginya proporsi komentar Etis ini merupakan hasil langsung dari penerapan *Strict Lexicon*, di mana kritik tajam terhadap kebijakan publik tidak disalahartikan sebagai pelanggaran etika.
3. Pendekatan *Lexicon-Based* yang dikombinasikan dengan validasi pakar (*Human-in-the-loop validation*) untuk pelabelan awal terbukti sangat efektif sebagai alternatif anotasi manual untuk dataset berskala besar, sekaligus menghasilkan *Golden Dataset* yang bebas dari ambiguitas konteks.
4. Meskipun hanya berada di proporsi 5,82%, keberadaan komentar "Tidak Etis" tetap menjadi catatan penting. Pola pelanggaran etika digital pada topik kebijakan publik ini didominasi oleh caci maki personal absolut, penggunaan terminologi kebinatangan (*animal terms*) yang merendahkan, ujaran kotor berbasis *scatological*, serta penggunaan singkatan kasar (*slang profanity*) yang sering luput dari filter otomatis platform.

Untuk penelitian selanjutnya, disarankan untuk mengeksplorasi penggunaan model IndoBERT, memperluas dataset, membandingkan teknik *oversampling* sintesis data (seperti SMOTE), serta mengembangkan sistem moderasi konten *real-time* yang terintegrasi langsung dengan API YouTube.

# REFERENCES

Br Sinulingga, J. E., & Sitorus, H. C. K. (2024). Analisis Sentimen Opini Masyarakat terhadap Film Horor Indonesia Menggunakan Metode SVM dan TF-IDF. *Jurnal Manajemen Informatika (JAMIKA)*, *14*(1), 42–53.

Mualfah Desti, Ramadhoni, Gunawan Rahmad, & Suratno Mulyadipa Danang. (2023). Analisis Sentimen Komentar YouTube TvOne Tentang Ustadz Abdul Somad Dideportasi Dari Singapura Menggunakan Algoritma SVM. *JURNAL FASILKOM*, *13*, 72–80.

Muhayat, T., Fauzi, A., & Indra, D. J. (2023). Analisis Sentimen Terhadap Komentar Video Youtube Menggunakan Support Vector Machines. *Jurnal Ilmiah Komputer*, *19*, 231–240.

Rakhmat Sani, R., Ayu Pratiwi, Y., Winarno, S., Devi Udayanti, E., & Farrikh Al Zami, dan. (2022). Analisis Perbandingan Algoritma Naive Bayes Classifier dan Support Vector Machine untuk Klasifikasi Hoax pada Berita Online Indonesia. *Jurnal Masyarakat Informatika*, *13*(2), 2777–0648.