

ANALISIS SENTIMEN YOUTUBE TERHADAP PEMBATAHAN MEDIA SOSIAL BAGI ANAK DI BAWAH 16 TAHUN MENGGUNAKAN MACHINE LEARNING

Bagas Daniswara¹

Universitas Muhammadiyah Kota Sukabumi, Jl.Syamsudin. SH No.50 S.H, Cikole, Kec. Cikole, Kota Sukabumi, Jawa Barat 43113, Telp. (0266) 218345

Keywords:

Sentiment Analysis, YouTube, Machine Learning, Support Vector Machine, CRISP-DM.

Correspondent Email:

bagasdnwr@gmail.com

Abstrak. YouTube menjadi salah satu media sosial yang banyak digunakan masyarakat untuk menyampaikan opini terhadap berbagai isu, termasuk kebijakan pembatasan media sosial bagi anak di bawah usia 16 tahun. Banyaknya komentar yang dihasilkan memerlukan analisis sentimen berbasis machine learning agar proses klasifikasi lebih efektif. Penelitian ini bertujuan menganalisis distribusi sentimen masyarakat serta membandingkan kinerja algoritma Naive Bayes, Support Vector Machine (SVM), dan K-Nearest Neighbor (K-NN). Penelitian menggunakan pendekatan kuantitatif dengan metodologi CRISP-DM. Data diperoleh melalui proses crawling komentar YouTube, kemudian diproses melalui tahapan preprocessing, pembobotan TF-IDF, split validation, dan klasifikasi menggunakan ketiga algoritma. Hasil penelitian menunjukkan bahwa dari 2.646 komentar terdapat 811 sentimen positif, 622 sentimen netral, dan 1.213 sentimen negatif, sehingga sentimen negatif menjadi kategori yang dominan. Berdasarkan hasil evaluasi, algoritma SVM memberikan performa terbaik dengan accuracy 72,08%, precision 71,85%, recall 72,08%, dan F1-score 71,30%, lebih baik dibandingkan Naive Bayes dan K-NN. Hasil penelitian menunjukkan bahwa SVM merupakan algoritma yang paling efektif untuk klasifikasi sentimen komentar YouTube pada topik ini.



Copyright © 2026 The Author(s), Published by Jurnal Informatika dan Teknik Elektro Terapan. This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

Abstract. YouTube has become one of the most widely used social media platforms for the public to express opinions on various issues, including the policy restricting social media access for children under the age of 16. The large volume of comments generated necessitates sentiment analysis based on machine learning to ensure a more effective classification process. This study aims to analyze the distribution of public sentiment and compare the performance of the Naive Bayes, Support Vector Machine (SVM), and K-Nearest Neighbor (K-NN) algorithms. This research employs a quantitative approach using the CRISP-DM methodology. The data were collected through crawling YouTube comments, followed by preprocessing, TF-IDF weighting, split validation, and classification using the three algorithms. The results show that out of 2,646 comments, there are 811 positive sentiments, 622 neutral sentiments, and 1,213 negative sentiments, indicating that negative sentiment is the dominant category. Based on the evaluation results, the SVM algorithm demonstrates the best performance with an accuracy of 72.08%, precision of 71.85%, recall of 72.08%, and an F1-score of 71.30%, outperforming Naive Bayes and K-NN. These findings indicate that SVM is the most effective algorithm for classifying sentiment in YouTube comments on this topic.

1. PENDAHULUAN

Perkembangan teknologi informasi telah mendorong media sosial menjadi ruang publik digital yang memfasilitasi pertukaran informasi dan opini secara luas [1]. Salah satu platform yang paling banyak digunakan di Indonesia adalah YouTube dengan tingkat penggunaan mencapai sekitar 88%, sehingga menjadi media yang merepresentasikan opini masyarakat terhadap berbagai isu melalui kolom komentar [2]. Tingginya aktivitas pengguna menjadikan komentar YouTube sebagai sumber data yang potensial untuk mengidentifikasi sentimen publik secara real-time.

Meningkatnya penggunaan media sosial oleh anak di bawah usia 16 tahun memunculkan kekhawatiran mengenai dampaknya terhadap perkembangan psikologis, sosial, dan kognitif. Kondisi tersebut mendorong munculnya wacana pembatasan akses media sosial bagi anak sebagai bentuk perlindungan. Kebijakan ini memicu beragam tanggapan dari masyarakat. Sebagian mendukung kebijakan tersebut demi melindungi anak, sementara sebagian lainnya menilai kebijakan tersebut dapat membatasi kebebasan berekspresi dan akses terhadap informasi. Perbedaan pandangan tersebut tercermin pada komentar pengguna YouTube yang dapat dimanfaatkan sebagai sumber analisis opini publik.

Analisis komentar dalam jumlah besar tidak dapat dilakukan secara manual karena data bersifat tidak terstruktur dan terus bertambah [3]. Oleh sebab itu, penerapan *text mining* dan analisis sentimen berbasis *machine learning* menjadi pendekatan yang efektif untuk mengolah serta mengklasifikasikan sentimen masyarakat secara otomatis [4].

Berbagai penelitian sebelumnya telah menerapkan analisis sentimen pada media sosial dengan objek penelitian yang beragam. Putro dan Hendrawan (2026) menganalisis sentimen terhadap kasus kebocoran data Bjorka menggunakan data komentar YouTube [5], sedangkan Rizqilla (2025) mengkaji sentimen masyarakat terhadap kinerja pemerintah melalui platform yang sama [6]. Penelitian Franko (2024) juga menunjukkan bahwa algoritma Naive Bayes mampu menghasilkan akurasi klasifikasi sentimen sebesar 85,4% [7]. Meskipun demikian, penelitian mengenai sentimen masyarakat terhadap kebijakan pembatasan media sosial bagi anak di bawah

usia 16 tahun masih terbatas, terutama yang memanfaatkan data komentar YouTube sebagai sumber data *real-time* sekaligus membandingkan beberapa algoritma klasifikasi.

Berdasarkan kondisi tersebut, penelitian ini bertujuan menganalisis sentimen masyarakat terhadap kebijakan pembatasan media sosial bagi anak di bawah usia 16 tahun menggunakan data komentar YouTube. Penelitian ini menerapkan algoritma *Naive Bayes*, *Support Vector Machine* (SVM), dan *K-Nearest Neighbor* (KNN) untuk membandingkan kinerja klasifikasi. Hasil penelitian diharapkan dapat memberikan gambaran mengenai opini publik terhadap kebijakan tersebut sekaligus menjadi referensi bagi pengembangan penelitian analisis sentimen dan pertimbangan dalam penyusunan kebijakan publik.

2. TINJAUAN PUSTAKA

Teori-teori yang digunakan berkaitan dengan platform media sosial sebagai sumber data, metodologi data mining, serta algoritma machine learning yang digunakan dalam proses klasifikasi sentimen. Dalam penelitian ini, landasan teori meliputi:

2.1. YouTube

YouTube merupakan platform media sosial berbasis video yang memungkinkan pengguna mengunggah, menonton, berbagi, dan memberikan komentar terhadap suatu konten. Fitur komentar pada YouTube menjadi sumber data yang relevan untuk analisis sentimen karena mampu merepresentasikan opini masyarakat terhadap berbagai isu, termasuk kebijakan publik [8].

2.2. CRISP-DM

Cross Industry Standard Process for Data Mining (CRISP-DM) merupakan metodologi standar dalam proses data mining yang terdiri atas enam tahapan, yaitu *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, *Evaluation*, dan *Deployment*. Metodologi ini digunakan sebagai pedoman penelitian karena menyediakan alur kerja yang sistematis dalam pengolahan dan analisis data [9].

2.3. Naïve Bayes Classifier

Naïve Bayes Classifier merupakan algoritma klasifikasi berbasis Teorema Bayes yang banyak digunakan pada analisis sentimen karena memiliki proses komputasi yang cepat, sederhana, dan efektif dalam mengolah data teks berdimensi tinggi [10].

2.4. Support Vector Machine (SVM)

Support Vector Machine (SVM) merupakan algoritma *supervised machine learning* yang mampu menghasilkan klasifikasi dengan tingkat akurasi tinggi, terutama pada data teks hasil pembobotan TF-IDF. Oleh karena itu, algoritma ini banyak diterapkan dalam penelitian analisis sentimen sebagai metode klasifikasi maupun pembandingan [11].

2.5. K-Nearest Neighbor (KNN)

K-Nearest Neighbor (K-NN) merupakan algoritma *supervised machine learning* yang mengklasifikasikan data berdasarkan kedekatan dengan sejumlah tetangga terdekat. Algoritma ini banyak digunakan pada analisis sentimen karena sederhana, mudah diimplementasikan, dan mampu memberikan performa yang baik setelah data melalui proses preprocessing dan pembobotan TF-IDF [12].

2.6. Klasifikasi

Klasifikasi merupakan teknik dalam *machine learning* yang digunakan untuk mengelompokkan data ke dalam kelas tertentu berdasarkan karakteristiknya. Dalam analisis sentimen, klasifikasi bertujuan menentukan polaritas teks, seperti positif, negatif, atau netral, melalui pembelajaran pola dari data latih yang kemudian digunakan untuk memprediksi data baru. Pada penelitian ini, metode klasifikasi digunakan untuk mengelompokkan komentar YouTube terkait kebijakan pembatasan media sosial bagi anak di bawah usia 16 tahun secara otomatis dan terukur. Pendekatan berbasis *machine learning* dinilai efektif dalam mengidentifikasi opini publik dengan tingkat akurasi yang baik, khususnya pada data media sosial berskala besar [13].

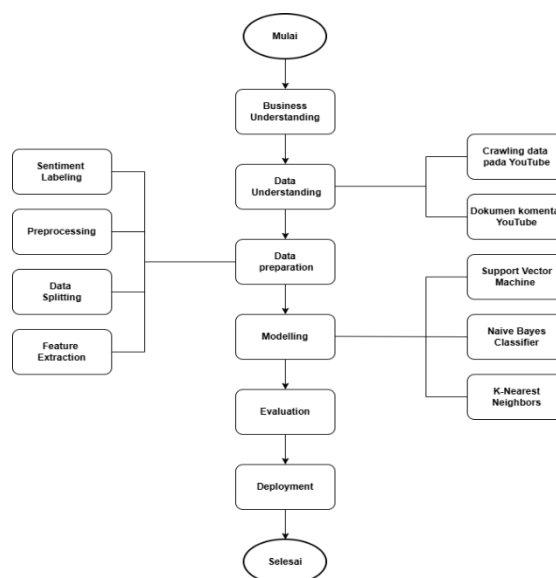
2.7. Evaluasi

Evaluasi model merupakan tahap penting dalam *machine learning* untuk mengukur kinerja dan kemampuan model dalam melakukan klasifikasi secara akurat. Pada

penelitian ini, evaluasi dilakukan untuk membandingkan kinerja algoritma *Naïve Bayes*, *Support Vector Machine*, dan *K-Nearest Neighbor* dalam mengklasifikasikan sentimen komentar YouTube ke dalam kategori positif, negatif, dan netral. Tahap ini memastikan validitas dan reliabilitas hasil klasifikasi sehingga model yang digunakan memiliki performa yang baik dan mampu memberikan prediksi yang konsisten pada data baru [14].

3. METODE PENELITIAN

Makalah Metode Penelitian ini menggunakan pendekatan kuantitatif dengan metodologi *Cross Industry Standard Process for Data Mining (CRISP-DM)*. Metodologi ini dipilih karena menyediakan tahapan yang sistematis dalam proses analisis sentimen, mulai dari *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, *Evaluation*, hingga *Deployment* [15].



Gambar 1. Alur Penelitian

Data penelitian berupa komentar pengguna YouTube mengenai kebijakan pembatasan media sosial bagi anak di bawah usia 16 tahun. Pengumpulan data dilakukan melalui teknik *web crawling* menggunakan Apify, sedangkan studi pustaka dilakukan untuk memperoleh referensi yang berkaitan dengan analisis sentimen, *text mining*, dan algoritma klasifikasi yang digunakan.

Proses pengolahan data dilakukan menggunakan bahasa pemrograman Python pada lingkungan Google Colaboratory.

Tahapan penelitian meliputi *text preprocessing*, pembobotan fitur menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF), pembagian data latih dan data uji (*split validation*), klasifikasi menggunakan algoritma *Naïve Bayes*, *Support Vector Machine* (SVM), dan *K-Nearest Neighbor* (K-NN), serta evaluasi model menggunakan *accuracy*, *precision*, *recall*, dan *F1-score*.

Penelitian dilaksanakan dengan objek berupa komentar pengguna YouTube yang berkaitan dengan kebijakan pembatasan media sosial bagi anak di bawah usia 16 tahun.

4. HASIL DAN PEMBAHASAN

Implementasi penelitian dilakukan menggunakan bahasa pemrograman Python pada lingkungan Google Colaboratory. Dataset yang diperoleh dari proses *web crawling* kemudian melalui tahapan *preprocessing*, pelabelan sentimen, pembagian data, ekstraksi fitur menggunakan TF-IDF, serta klasifikasi menggunakan algoritma *Naïve Bayes*, *Support Vector Machine* (SVM), dan *K-Nearest Neighbor* (K-NN). Selanjutnya, model terbaik diimplementasikan ke dalam aplikasi berbasis Streamlit.

4.1. Business Understanding

Tahap *Business Understanding* diawali dengan identifikasi meningkatnya penggunaan media sosial oleh anak di bawah usia 16 tahun yang memunculkan wacana pembatasan akses sebagai upaya perlindungan digital. Kebijakan ini menimbulkan beragam tanggapan masyarakat yang tercermin dalam komentar di YouTube. Oleh karena itu, penelitian ini menganalisis komentar tersebut untuk memperoleh gambaran objektif mengenai kecenderungan sentimen publik terhadap kebijakan yang diusulkan.

4.2. Data Understanding

Tahap *Data Understanding* dilakukan dengan mengumpulkan data melalui proses scraping menggunakan Apify pada platform YouTube. Data diperoleh berdasarkan kata kunci "Kebijakan media sosial untuk anak" dan "Pembatasan media sosial untuk anak di bawah 16 tahun" selama periode Maret hingga Mei 2026, sehingga terkumpul sebanyak 2.647 komentar. Selanjutnya, atribut yang tidak relevan, seperti *cid*, *replyToCid*, *type*,

published, *videoId*, *pageUrl*, *CommentsCount*, dan *title*, dihapus pada tahap *Data Preparation*.

Gambar 2. Dataset

4.3. Data Preparation

Data Preparation dalam penelitian ini meliputi:

4.3.1. Preprocessing

Tahap *Preprocessing* terdiri dari *case folding*, *cleansing*, *normalization*, *tokenizing*, *stopword removal*, dan *stemming*. Berikut hasil dari 6 tahapan tersebut:

Tabel 1. Case Folding

Komentar	Case Folding
Ibarat anak kita mau makan ikan pasti kita ajarin jangan makan durinya bukan berarti gk boleh makan ikan. Nah jadi sama seharusnya dibatasi bukan di nonaktifkan	ibarat anak kita mau makan ikan pasti kita ajarin jangan makan durinya bukan berarti gk boleh makan ikan. nah jadi sama seharusnya dibatasi bukan di nonaktifkan

Tabel 2. Cleansing

Case Folding	Cleansing
syarat medsos kita gak boleh dibatasi kalau gak mau gak usah dipaksakan 🤔🤔🤔	syarat medsos kita gak boleh dibatasi kalau gak mau gak usah dipaksakan

Tabel 3. Normalization

Cleansing	Normalization
iya betul banyak kasus perundungan sama penipuan yang ngarah ke anak anak di media sosial. Pembatasan usia bisa menjadi salah satu langkah pencegahan yang masuk akal.	iya betul banyak kasus perundungan sama penipuan yang ngarah ke anak anak di media sosial. Pembatasan usia bisa menjadi salah satu langkah pencegahan yang masuk akal.

Tabel 4. Tokenization

Normalization	Tokenization
tapi kalau sudah lama terus sudah di atas tahun boleh enggak maksud nya pas sudah	['tapi', 'kalau', 'sudah', 'lama', 'terus', 'sudah', 'di', 'atas', 'tahun', 'boleh', 'enggak', 'maksud', 'nya', 'pas', 'sudah']

lewat maret semoga di baca	'maksud', 'nya', 'pas', 'sudah', 'lewat', 'maret', 'semoga', 'di', 'baca']
----------------------------	--

Tabel 5. *Stopword Removal*

Tokenization	Stopword Removal
['saya', 'mendukung', 'kebijakan', 'ini', 'karena', 'anak', 'di', 'bawah', 'tahun', 'perlu', 'masih', 'perlu', 'pengawasan', 'dalam', 'menggunakan', 'media', 'sosial', 'dampak', 'konten', 'yang', 'tidak', 'sesuai', 'usia', 'memang', 'perlu', 'diantisipasi', 'sejak', 'dini']	['mendukung', 'kebijakan', 'anak', 'bawah', 'tahun', 'perlu', 'pengawasan', 'menggunakan', 'media', 'sosial', 'dampak', 'konten', 'tidak', 'sesuai', 'usia', 'memang', 'perlu', 'diantisipasi', 'sejak', 'dini']

Tabel 6. *Stemming*

Stopword Removal	Stemming
['banyak', 'kasus', 'perundungan', 'terjadi', 'media', 'sosial', 'pembatasan', 'usia', 'menjadi', 'salah', 'satu', 'langkah', 'pencegahan', 'masuk', 'akal']	['banyak', 'kasus', 'rundung', 'tipu', 'jadi', 'media', 'sosial', 'batas', 'usia', 'jadi', 'salah', 'satu', 'langkah', 'cegah', 'masuk', 'akal']

4.3.2. Sentiment Labeling

Proses pelabelan menghasilkan 1.213 komentar negatif, 811 komentar positif, dan 622 komentar netral dari 2646 komentar.

```
sentimen
negatif    1213
positif     811
netral      622
Name: count, dtype: int64
```

Gambar 3. *Sentimen Labeling*

4.3.3. Data Splitting

Pembagian data dilakukan dengan proporsi 80% data latih dan 20% data uji.

```
Jumlah data training: 2116
Jumlah data testing: 530
```

Gambar 4. *Data Splitting*

4.3.4. Feature Extraction

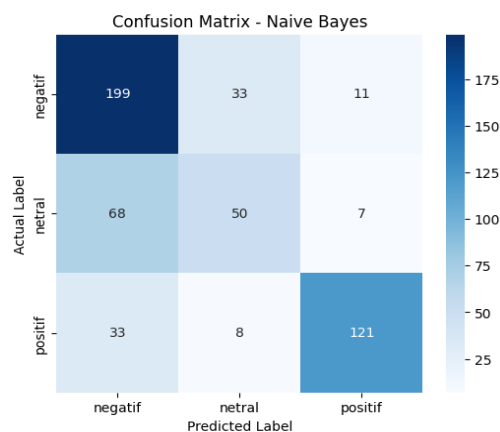
Pada tahap *feature extraction*, data komentar yang telah melalui proses *preprocessing* diubah menjadi representasi numerik menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF).

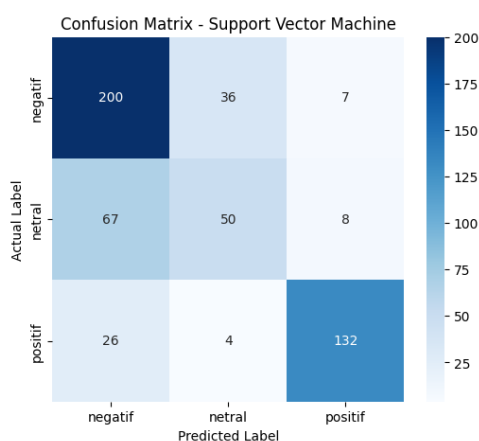
4.4. Modeling

Tahap *modeling* dilakukan setelah proses *feature extraction* menggunakan metode TF-IDF selesai. Pada tahap ini, data hasil transformasi TF-IDF digunakan sebagai masukan untuk membangun model klasifikasi sentimen. Penelitian ini menggunakan tiga algoritma klasifikasi, yaitu *Naive Bayes Classifier* (MultinomialNB), *Support Vector Machine* (SVM), dan *K-Nearest Neighbor* (KNN).

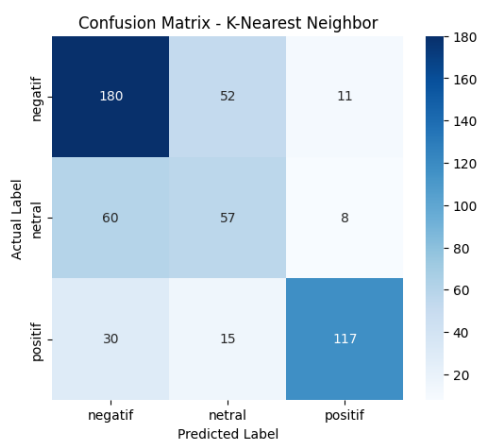
4.5. Evaluation

Tahap evaluasi dilakukan untuk mengukur performa model klasifikasi sentimen yang telah dibangun menggunakan algoritma *Naive Bayes Classifier* (MultinomialNB), *Support Vector Machine* (SVM), dan *K-Nearest Neighbor* (KNN). Pengukuran kinerja model terdiri dari metrik *Accuracy*, *Precision*, *Recall*, dan *F1-Score* yang dihitung dari *multiclass confusion matrix* 3x3. Berikut confusion matrix dari ketiga algoritma yang digunakan.

Gambar 5. *Confusion Matrix NB*



Gambar 6. Confusion Matrix SVM



Gambar 7. Confusion Matrix KNN

Berdasarkan perbandingan dari ketiga model yang dilakukan, didapatkan hasil Berikut.

Tabel 6. Hasil Perbandingan

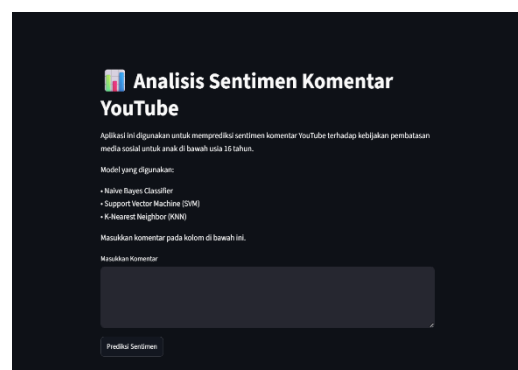
Algoritma	Accuracy	Precision	recall	F1-Score
Naive Bayes	0,69	0,69	0,69	0,69
Support Vector Machine (SVM)	0,72	0,71	0,72	0,71
K-Nearest Neighbor (KNN)	0,66	0,67	0,66	0,66

Pada hasil perbandingan kinerja ketiga algoritma dapat dilihat bahwa, *Support Vector Machine* (SVM) memberikan performa terbaik dengan *Accuracy* 72,08%, *Precision* 71,85%,

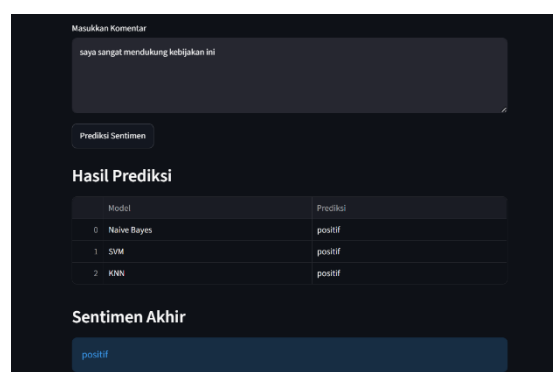
Recall 72,08%, dan *F1-Score* 71,30%. *Naive Bayes* berada pada posisi kedua dengan *Accuracy* 69,81%, sedangkan *K-Nearest Neighbor* (KNN) memperoleh performa terendah dengan *Accuracy* 66,79%. Hasil ini menunjukkan bahwa SVM lebih efektif dalam mengklasifikasikan sentimen, sementara KNN kurang optimal dalam menangani data teks berdimensi tinggi yang direpresentasikan menggunakan TF-IDF.

4.6. Deployment

Tahap *deployment* dilakukan untuk mengimplementasikan model analisis sentimen yang telah dibangun ke dalam bentuk aplikasi berbasis web menggunakan Streamlit. Tujuan dari tahap ini adalah agar model yang telah melalui proses *preprocessing*, ekstraksi fitur, pemodelan, dan evaluasi dapat digunakan untuk melakukan prediksi sentimen secara langsung terhadap komentar yang dimasukkan oleh pengguna.



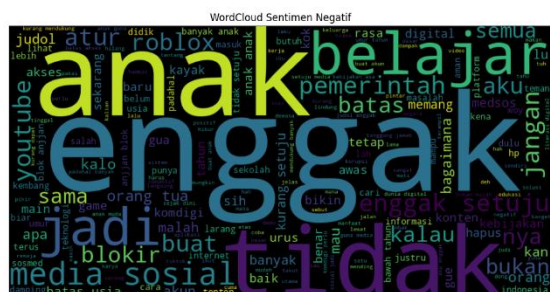
Gambar 8. Tampilan Awal pada Website



Gambar 9. Tampilan hasil prediksi pada Website

Setelah itu, Pada tahap deployment ini peneliti ingin mengetahui juga kata-kata yang banyak diperbincangkan, *Word Cloud*

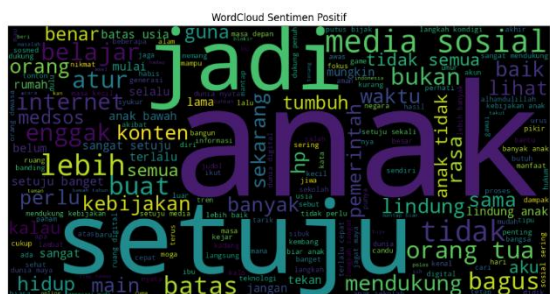
digunakan untuk memvisualisasikan kata-kata yang paling sering muncul pada setiap kelas sentimen. Berikut *Word Cloud* dari setiap kelas sentimen.



Gambar 10. *Wordcloud* pada sentimen *Negatif*



Gambar 11. *Wordcloud* pada sentimen *Netral*



Gambar 10. *Wordcloud* pada sentimen *Positif*

5. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, diperoleh kesimpulan sebagai berikut.

- Analisis sentimen terhadap 2.646 komentar YouTube mengenai kebijakan pembatasan media sosial bagi anak di bawah usia 16 tahun menunjukkan distribusi sentimen sebanyak 1.213 komentar negatif, 811 komentar positif, dan 622 komentar netral. Hasil tersebut menunjukkan bahwa sentimen negatif menjadi kategori yang paling dominan.
- Hasil perbandingan algoritma *Naïve Bayes*, *Support Vector Machine* (SVM), dan *K-*

Nearest Neighbor (K-NN) menunjukkan bahwa algoritma SVM memberikan performa terbaik dengan nilai *accuracy* sebesar 72,08%, *precision* 71,85%, *recall* 72,08%, dan *F1-score* 71,30%. Dengan demikian, SVM merupakan algoritma yang paling efektif untuk mengklasifikasikan sentimen komentar YouTube pada penelitian ini.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Allah SWT, dengan pembimbing, serta keluarga dan semua pihak yang telah memberikan dukungan, bimbingan, dan motivasi sehingga penelitian ini dapat diselesaikan dengan baik.

DAFTAR PUSTAKA

- [1] D. F. Zhafira, B. Rahayudi, and I. Indriati, "Analisis Sentimen Kebijakan Kampus Merdeka Menggunakan Naive Bayes dan Pembobotan TF-IDF Berdasarkan Komentar pada Youtube," *Jurnal Sistem Informasi, Teknologi Informasi, Dan Edukasi Sistem Informasi*, vol. 2, no. 1, 2021.
- [2] S. M. Harahap and R. Kurniawan, "Analisis sentimen komentar Youtube terhadap food vlogger dengan menggunakan metode Naive Bayes," *MEANS (Media Informasi Analisa dan Sistem)*, pp. 87–96, 2024.
- [3] K. Azmi, "Fenomena viral di media sosial: Analisis pola komunikasi dan pengaruhnya terhadap opini publik," *Journal of Science and Social Research*, vol. 8, no. 2, pp. 1814–1821, 2025.
- [4] M. Monica and A. Purwanto, "MALCOM: Indonesian Journal of Machine Learning and Computer Science Comparative Performance Evaluation of Naive Bayes and Logistic Regression for Indonesian YouTube Comment Sentiment Classification Evaluasi Komparatif Kinerja Algoritma Naive Bayes dan Logistic Regression dalam Klasifikasi Sentimen Komentar YouTube Berbahasa Indonesia," vol. 6, no. 2, pp. 934–943, 2026, doi: 10.57152/malcom.v6i2.2639.
- [5] R. H. Putro and A. Hendrawan, "Analisis Sentimen Komentar Youtube Terhadap Kasus Bjorka Dengan Membandingkan Naive Bayes dan SVM," *KETIK: Jurnal Informatika*, vol. 3, no. 03, pp. 137–147, 2026.
- [6] M. Rizqilla, D. Nurani, A. Rahmi, S. Supriatin, and M. Maemunah, "Analisis Sentimen Kinerja Jokowi Menggunakan Metode Naive Bayes," *REMIK: Riset dan E-Jurnal Manajemen*

- Informatika Komputer*, vol. 9, no. 2, pp. 550–561, 2025.
- [7] B. Franko, N. Wilyanto, and H. Irsyad, “Analisis Sentimen terhadap Naturalisasi Pemain pada Youtube Menggunakan Decision tree dan Naive bayes,” *Software Development, Digital Business Intelligence, and Computer Engineering*, vol. 3, no. 1, pp. 8–16, 2024.
- [8] K. A. Ulumiyah, “Analisis Sentimen Komentar Youtube Terkait Vidio Demo DPR Menggunakan Algoritma Naïve Bayes,” *Jurnal Mahasiswa Teknik Informatika*, vol. 5, no. 1, pp. 388–394, 2026.
- [9] Y. A. Singgalen, “Penerapan Metode CRISP-DM dalam Klasifikasi Data Ulasan Pengunjung Destinasi Danau Toba Menggunakan Algoritma Naïve Bayes Classifier (NBC) dan Decision Tree (DT),” *J. Media Inform. Budidarma*, vol. 7, no. 3, pp. 1551–1562, 2023.
- [10] O. M. Wulandari, I. Maulana, F. Syamsudin, and R. Waluyo, “Perbandingan Algoritma Naive Bayes dan SVM dalam Analisis Sentimen Twitter terhadap Isu Ijazah Jokowi Palsu,” *Jurnal Manajemen Informatika, Sistem Informasi dan Teknologi Komputer (JUMISTIK)*, vol. 4, no. 1, pp. 392–400, 2025.
- [11] S. Styawati, N. Hendrastuty, and A. R. Isnain, “Analisis sentimen masyarakat terhadap program kartu prakerja pada twitter dengan metode support vector machine,” *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 6, no. 3, pp. 150–155, 2021.
- [12] I. H. Kusuma and N. Cahyono, “Analisis Sentimen Masyarakat Terhadap Penggunaan E-Commerce Menggunakan Algoritma K-Nearest Neighbor,” *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 8, no. 3, pp. 302–307, 2023.
- [13] S. Fitrianti and A. Yudhistira, “ANALISIS SENTIMEN MEDIA SOSIAL TERHADAP CALON PILKADA 2024 DENGAN METODE NAÏVE BAYES,” *JURNAL PENDIDIKAN*, vol. 5, no. 1, pp. 167–176, 2025.
- [14] N. Hidayah and D. Dodiman, “Implementasi algoritma multinomial naïve bayes, TF-IDF dan confusion matrix dalam pengklasifikasian saran monitoring dan evaluasi mahasiswa terhadap dosen teknik informatika Universitas Dayanu Ikhsanuddin,” *Jurnal Akademik Pendidikan Matematika*, pp. 8–15, 2024.
- [15] H. Ardiansya, “Analisis Sentimen Pendapat Masyarakat Terhadap PPKM DKI Jakarta Dengan Metode Naïve Bayes,” *JATISI*, vol. 10, no. 2, pp. 60–70, 2023.