
ANALISIS OPINI PUBLIK MENGENAI BATAS USIA PENGGUNAAN SOSMED ANAK INDONESIA MENGGUNAKAN PENDEKATAN MACHINE LEARNING

Prima Lydia¹, Muhammad Irfan Sarif², Nauval Alfarizi³, Sri Hidayati⁴, Adi Putra⁵
Universitas Panca Budi, Medan

e-mail: ¹2anpu3@gmail.com, ²irfanberbagi@gmail.com, ³nauvalalfarizi026@gmail.com,
⁴srihidayati08@guru.smp.belajar.id, ⁵myreal.autobot@gmail.com

Abstract: *This research looks at how people feel about government rules that limit children under 16 from using social media. This study uses data consisting of 4,651 social media posts obtained through scraping, as well as questionnaire responses from 338 participants at a public junior high school in the Deli Serdang region. In this study, two types of classification models were used: Support Vector Machine (SVM) and Random Forest. These models were compared to see how well they perform. This method serves as a means of achieving relatively similar levels of accuracy. The results indicate that both models perform at comparable accuracy levels. However, SVM demonstrates superior performance in maintaining class balance, as indicated by a higher macro F1-score of 0.85, compared to Random Forest, which tends to be more biased toward the neutral class. Further analysis shows that the neutral class is the most common in the sentiment distribution. Additionally, the models often have difficulty accurately recognizing positive sentiment, probably because they tend to be cautious and classify uncertain inputs as neutral. This study doesn't use special methods to balance the data, so it can check how well the model works when the data is naturally unbalanced. This method helps us understand how public opinion reacts to government policies, especially those affecting children, as the main focus of the analysis.*

Keywords: *Access Restriction, Sentiment Analysis, Public Opinion, Children, Social Media*

Abstrak: Penelitian ini bertujuan menganalisis perasaan masyarakat umum terhadap kebijakan yang dibuat pemerintah yang membatasi akses anak di bawah usia 16 tahun ke media sosial. Data yang digunakan berasal dari hasil pengumpulan data melalui scraping media sosial sebanyak 4651 data dan data dari kuesioner yang diisi oleh 338 responden dari salah satu SMP Negeri Wilayah Deli Serdang. Model klasifikasi yang digunakan dalam penelitian ini adalah Support Vector Machine (SVM) dan Random Forest sebagai perbandingan dalam melihat performa. Hasil pengujian menunjukkan kedua model memiliki tingkat akurasi relatif sama. Namun, SVM menunjukkan hasil yang lebih baik dalam menjaga keseimbangan antar kelas dengan skor F1 makro sebesar 0,85, dibandingkan dengan Random Forest yang cenderung lebih condong ke kelas netral. Analisis lebih lanjut menunjukkan bahwa kelas netral mendominasi dalam penyebaran sentimen. Kemudian, model juga sering kesulitan dalam mengenali perasaan positif dikarenakan model cenderung konservatif dengan memilih netral. Dalam penelitian ini tidak mengklasifikasi secara khusus seperti model imbalance data untuk melihat performa tanpa perlu Teknik sampling guna melihat seberapa dan kenapa opini publik bereaksi terhadap suatu kebijakan pemerintahan khususnya terhadap anak-anak sebagai output utama

Kata Kunci: Pembatasan, Sentimen, Opini, Anak-Anak, Media Sosial

terus berkembang memberikan dampak langsung khususnya pada individu, baik dalam pola perilaku (budaya, etika, moral), kreativitas, maupun pembentukan karakter. Sebagai contoh, media sosial berperan penting dalam membentuk identitas diri, terutama pada remaja yang berada dalam fase pencarian jati diri [1]. Akses internet di Indonesia terus meningkat setiap tahunnya dengan peningkatan yang cukup signifikan. Data

BPS Indonesia melalui hasil survei Susenas menunjukkan bahwa pada tahun 2021, sebanyak 62,10% dari penduduk Indonesia sudah bisa menggunakan *internet*. Peningkatan penggunaan akses *internet* terjadi karena jumlah pengguna *internet* semakin bertambah, yang tidak terlepas dari perkembangan cepat teknologi telepon seluler, yang juga dikenal sebagai *smartphone*. Hal itu ditunjukkan oleh peningkatan persentase jumlah rumah tangga yang memiliki perangkat telepon seluler. Dalam data dari BPS (Badan Pusat Statistik) Indonesia, tercatat bahwa 90,54% rumah tangga di Indonesia sudah memiliki atau menguasai telepon seluler pada tahun 2021. Angka ini naik dibandingkan tahun 2018 yang hanya mencapai 88,46%. Masyarakat [2]. Opini publik merupakan sebuah pandangan, sikap, atau penilaian yang berkembang di tengah masyarakat terhadap suatu isu, kebijakan, atau fenomena tertentu yang menjadi perhatian bersama. Melalui opini ini terbentuk melalui proses interaksi sosial, pertukaran informasi, serta pengaruh media massa dan media digital, sehingga mencerminkan persepsi kolektif yang dinamis dan dapat berubah [3].

Sekarang ini di masa digital, teknologi informasi dan media sosial sudah mengubah cara masyarakat terlibat dalam urusan khususnya anak dalam mengambil keputusan untuk bermedia sosial [4]. Gadget memberi dampak baik yang penting bagi anak-anak dan remaja dalam membantu berbagai kegiatan yang mereka lakukan. Salah satu manfaat terpentingnya adalah membuat

komunikasi dan berinteraksi dengan keluarga atau teman jadi lebih mudah, tanpa terbatas oleh jarak atau waktu. Dengan gadget dapat memberikan dampak berupa penggunaan media sosial yang yang bisa diakses melalui berbagai jenis perangkat elektronik [5]. Adapun juga dibalik dampak positif dan media sosial terdapat juga dampak negatif yang ada yaitu: kenakalan remaja yang meliputi perkelahian antar pelajar, penyalahgunaan narkoba, penipuan, kebebasan seks, dan sebagainya [6]. Menurut Laporan Komdigi Tahun 2024, tercatat 1.450.403 kasus yang secara khusus membahas mengenai eksploitasi seksual anak daring sepanjang tahun tersebut [7]. Peran orang tua memiliki arti yang sangat penting dalam proses pertumbuhan anak. Oleh sebab itu, pendidikan anak sebaiknya dilakukan dengan memberikan pengalaman yang bermakna serta menghargai setiap usaha yang mereka lakukan. Selain itu, waktu belajar anak tidak seharusnya terlalu banyak tersita oleh berbagai tugas lain, karena kondisi tersebut dapat menimbulkan rasa malas yang pada akhirnya berdampak pada penurunan prestasi belajar. Dalam hal ini, kerja sama antara orang tua dan guru di sekolah perlu menjadi prioritas utama dalam membimbing anak, terutama dalam upaya menumbuhkan dan meningkatkan minat belajar mereka [8].

Adapun regulasi yang mengatur mengenai kaitan bagaimana seseorang dilindungi dalam Undang-Undang Nomor 1 tahun 2024 tentang Perubahan Kedua atas Undang- Undang Nomor 11 Tahun 2008 Pasal 45 Ayat (1) mengatur mengenai hukuman terhadap pelaku yang tergolong sebagai tindakan yang menyangkut Informasi dan Transaksi Elektronik [9]. Setiap individu memiliki hak fundamental untuk menyampaikan pendapat dan memperoleh informasi sebagai bagian dari kebebasan dasar. Hak tersebut tidak hanya berkaitan dengan ekspresi diri, tetapi juga berperan dalam pembentukan opini serta partisipasi dalam kehidupan social khususnya media sosial. Oleh karena itu, keberadaan hak ini perlu

dijamin dan dilindungi oleh negara sebagai entitas hukum dan politik. Negara tidak hanya berfungsi sebagai pemegang kekuasaan tertinggi, tetapi juga sebagai lembaga yang bertanggung jawab dalam melindungi warga [10].

Pola dalam machine learning menitikberatkan pada sejumlah aspek utama, seperti perumusan pertanyaan yang berkaitan dengan prediksi perilaku pengguna untuk memperoleh nilai ekonomi, serta pengembangan sistem yang mengintegrasikan berbagai algoritma pembelajaran mesin dan metodologi teknik yang digunakan dalam proses analisis [11]. Dengan menggunakan machine learning kita dapat melakukan bagaimana cara mengelola opini public yang ada dalam mempelajari pola yang ada. Bentuk implementasi yang ada salah satunya menggunakan algoritma yang cukup klasik seperti SVM dan Random Forest. SVM memiliki performa hasil diatas rata rata dengan menggunakan hyperlane sebagai pemisah antar variabel yang ada dengan performa akurasi 85% dengan penelitian yang dilakukan oleh I. Siti Aisyah 2024 dengan judul Algoritma Support Vector Machine (Svm) Untuk Analisis Sentimen Ulasan Aplikasi Al Qur'an Digital [12].

Penelitian berikutnya yang membahas mengenai bagaimana implementasi dari algoritma Random Forest memiliki hasil evaluasi kinerja yang baik dengan nilai R^2 sebesar 76,66% dan MAPE sebesar 25,27%. Hal ini mengindikasikan bahwa model mampu menjelaskan sebagian besar variasi harga properti dengan tingkat kesalahan yang relatif moderat [13]. Dalam penelitian yang dilakan selanjutnya Adalah bagaimana data mining digunakna untuk menentukan pola terhadap suatu analisis yang akan dilakukan [14] dengan penelitian dengan judul Application of Data Mining to analyze the Best Suppliers yang dilakukan pada tahun 2024.

Dalam menentukan analisis opini public berdasarkan pendekatan *machine learning* diperlukan sebuah cara untuk mengatasi

penaganan data tidak seimbang khususnya data dengan sentiment positif yang dalam analisis sentiment jauh lebih dominant. Maka diperlukan sebuah Teknik penanganan data seperti Random Oversampling dan Smote dalam menentukan bagaimana cara penanganan data yang ada agar dapat memberikan hasil performa yang jauh lebih baik [15].

METODE

Penelitian yang dilakukan kali ini adalah dengan menerapkan analisis kuantitatif berbasis *machine learning* untuk mengkaji opini publik mengenai batas usia penggunaan media sosial pada anak di Indonesia. Data yang digunakan berupa teks opini yang dikumpulkan dari platform media sosial dengan dataset utama yang dikumpul melalui penyebaran kusioner pada salah satu smp negeri di deli Serdang dan juga dataset dari scrapping youtube. Data diproses melalui tahapan preprocessing yang meliputi pembersihan data (cleaning), case folding, tokenizing, stopword removal, dan stemming untuk menghasilkan data yang siap dianalisis.

Dalam proses analisis, penelitian ini menerapkan teknik klasifikasi untuk mengelompokkan opini publik ke dalam kategori tertentu, seperti positif, negatif, atau netral. Fitur teks direpresentasikan menggunakan metode pembobotan seperti TF-IDF. Model machine learning yang digunakan, adalah Support Vector Machine (SVM) dan Random Forest. ketidakseimbangan data dilakukan dengan cara teknik penanganan imbalanced data seperti SMOTE atau Random Oversampling. Evaluasi model dilakukan menggunakan metrik akurasi, precision, recall, dan F1-score untuk mengukur kinerja model dalam mengklasifikasikan opini publik secara akurat dengan hasil akhir berupa perbandingan performa.

Dataset

Dataset yang ada pada penelitian kali ini merupakan sebuah dataset yang

didapat dengan 2 cara . melalui form kusioner terkait larangan platform media sosial khusus anak dibawah 16 tahun dan juga *web scrapping* dari sumber komentar youtube dengan dataset lebih dari >4000 data.

id	user	text	time
1	@mbal@u-ki	Meluang paku... Jaman dulu saat ini masih dominan...	1 month ago
2	@On_Pubi	Tembak, tak ada baik daripada bisa sama...	1 month ago
3	@CicoMarunggi-gy	Jangan kasih hp sebelum anak masuk SMA Saya...	1 month ago
4	@santiamy7E40	HP memang tak penting kepada anak-anak...	1 month ago

Gambar 1 Dataset Gform

id	user	text	time
4848	@santiamy7E40	Masih P. Dulu anak-anak IT. Kalau masih...	3 weeks ago
4847	@santiamy7E40	Auto pengaliran dibelikan. Dapat sampai beta...	3 weeks ago
4846	@santiamy7E40	Aku sendiri kamu kamu... kamu kamu 2 hal ya...	3 weeks ago
4845	@Fayuliana	Hehe... Udah... udah... udah... udah... udah... udah...	3 weeks ago
4844	@santiamy7E40	Hehe... Tak ada masalah... Ya... Ya... Ya... Ya... Ya... Ya...	3 weeks ago

Gambar 2 Dataset Web Scrapping

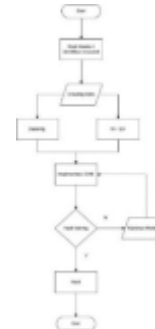
Transformasi Data

Dalam melakukan analisis yang diperlukan , dataset yang ada dapat dilihat pada gambar 1-2 akan kita lakukan proses cleaning , dikarenakan sebelum diperlukan model learning diperlukan sebuah mekanisme yaitu melakukan proses cleaning. Proses cleaning yang dilakukan bertujuan untuk memberikan hasil performa baik dan akurat.

Setelah proses cleaning dilakukan, data yang telah dibersihkan kemudian memasuki tahap preprocessing lanjutan yang meliputi normalisasi teks, tokenisasi, penghapusan stopwords, serta proses stemming atau lemmatization. teks yang telah diproses akan dikonversi menjadi bentuk numerik menggunakan metode pembobotan seperti TF-IDF agar dapat dipahami oleh algoritma klasifikasi. Dengan representasi fitur yang baik, model machine learning seperti Support Vector Machine (SVM) dapat bekerja secara optimal dalam mengidentifikasi pola dan melakukan klasifikasi. Oleh

karena itu, rangkaian proses preprocessing ini memiliki peranan yang sangat penting dalam meningkatkan akurasi dan performa model yang dihasilkan.

Flowchart Penelitian



Gambar 3 Flowchart Penelitian

Flowchart Preprocessing



Gambar 4 Tahapan Preprocessing

TF-IDF

Text preprocessing merupakan tahapan awal dalam pengolahan data teks yang bertujuan untuk membersihkan dan menormalisasi data agar lebih terstruktur dan siap digunakan pada proses analisis selanjutnya. Tahapan ini dimulai dengan cleaning, yaitu menghapus karakter yang tidak diperlukan seperti tanda baca, angka, URL, dan simbol lainnya. Selanjutnya dilakukan *case folding* untuk mengubah seluruh huruf menjadi huruf kecil agar konsisten.

Algoritma SVM

Dalam melakukan klasifikasi regresi diperlukan sebuah algoritma untuk menentukan optimalisasi yang diinginkan , SVM merupakan algoritma yang cukup solid dalam menentukan Batasan dalam tiap proses yang dibuktikan dengan

fungsi dasar (3) *hyperlane* sebagai batas Keputusan terhadap input yang dimasukkan.

Algoritma Random Forest

Random forest merupakan sebuah gabungan dari layer dalam mendukung Keputusan , algoritma ini menggunakan tree sebagai input dalam melakukan visualisasi terhadap output yang ada. Output yang ada akan diterapkan dan digunakan untuk mendukung klasifikasi seperti regresi.

HASIL DAN PEMBAHASAN

Pada penelitian hasil dataset yang didapat dengan scrapping memberikan hasil output berupa 4651 data dengan select proses kolom text menggunakan tahapan *text processing* dapat dilihat pada gambar 5.



Gambar 5 Hasil dari Preprocessing



Gambar 6. Preprocessing G-form

Pada gambar 5 data diberi label sentiment dalam bentuk data object list yang masing sentiment bernilai 150 kosa kata kamus dalam menentukan labeling yang ada untuk menemukan performa mengenai bagaimana opini public terbentuk menanggapi mengenai kebijakan kementerian dalam melarang dan memblokir kegiatan akses media sosial dibawah umur 16 tahun. Juga pada gambar 6 merupakan hasil dari preprocessing yang dilakukan pada

kusioner form resmi mengenai opini public terkait pemerintah mengenai pembatasan akses media sosial yang menasar salah satu smp negeri yang ada di wilayah deli Serdang , Sumatera Utara. Data terkait ini mempunyai sample kusioner dengan jumlah data sample sebanyak 338 orang.

Labeling Sentimen

Pada dataset dari platform media scrapping , didapatkan hasil labeling berupa netral 2959 data , positif 610 data dan negative 1082 data. Pada dataset form yang disebar terdapat hasil labeling berupa netral 202 data , positif 108 data dan juga negative 21 data. Labeling ini didapat dari hasil preprocessing dan juga lexicon kamus data Dimana setiap data memiliki kamus 150 data untuk tiap sentiment yang ada. Penggunaan atribut data sebanyak 450 data yang mewakili 3 sentimen utama negatif, positif dan netral.

Perbandingan Hasil Algoritma

Algoritma yang dibandingkan dalam pengujian kali ini Adalah membandingkan hasil dari SVM dan juga Random Forest untuk dapat melihat performa dalam menganalisis sentiment. Dataset yang ada didapat dari 2 bentuk scrapping dan juga kusioner yang ada pada salah satu SMP Negeri yang ada di wilayah Deli Serdang.

Intepretasi dari Algoritma SVM

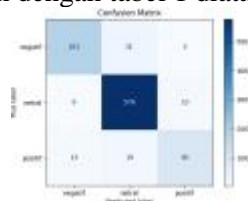
Dalam analisis yang dilakukan pada peneliti berikut ini perlu dipastikan bagaimana hasil tabel ini dapat dijelaskan

Tabel 1 SVM data Scrrapping

Kelas	Precision	Re call	F1- Score	Sup port
Negatif	0.91	0.81	0.85	217
Netral	0.90	0.98	0.94	592
Positif	0.86	0.67	0.76	122
Accuracy			0.90	931
Macro Avg	0.89	0.82	0.85	931
Weighted Avg	0.90	0.90	0.89	931

Model yang digunakan menunjukkan performa yang baik dengan tingkat akurasi sebesar 89.7%. Pada kelas netral, model memiliki performa terbaik dengan nilai recall sebesar 0.98 dan f1-score sebesar 0.94, yang menunjukkan bahwa model sangat efektif dalam mengidentifikasi data yang bersifat netral. Untuk kelas negatif, model juga menunjukkan performa yang cukup baik dengan precision sebesar 0.91 dan recall sebesar 0.81, yang berarti sebagian besar data negatif dapat terklasifikasi dengan benar.

Namun, pada kelas positif, performa model relatif lebih rendah dibandingkan kelas lainnya, dengan recall sebesar 0.67 dan f1-score sebesar 0.76. Hal ini terjadi dikarenakan simpulan dari algoritma yang menegaskan bahwa dataset yang ada memiliki bias dikarenakan model cenderung malas (*lazy load*) dalam menebak prediksi yang sesuai. Hampir klasifikasi yang didapatkan dari model netral terklasifikasi sebagai netral meskipun terdapat label yang memiliki nilai positif namun terkluster menjadi model netral hal ini dibuktikan dengan tabel 1 diatas.



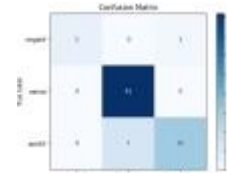
Gambar 7 C – Matrix SVM Scapping

Pada gambar 7 dijelaskan juga dalam bentuk visual mengenai matriks, model menunjukkan hasil baik pada kelas netral dengan jumlah hasil prediksi yang benar mencapai 576 data. Kesalahan klasifikasi terjadi paling banyak pada positif yang sering kali dikategorikan sebagai netral. Model ini cenderung memilih opini netral ketika tidak yakin cukup kuat terhadap pendapat yang diberikan, sehingga menunjukkan sikap konservatif dalam memberikan label sentimen.

Tabel 2 C-Matrix SVM form

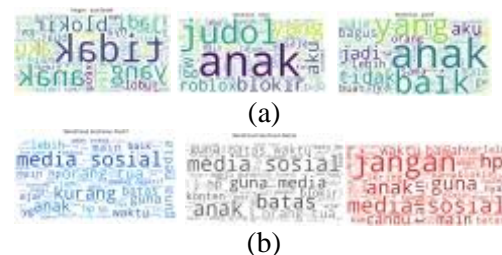
Kelas	Preci	Re	F1-	Sup
	sion	call	Score	port
Negatif	1.00	0.75	0.86	4
Netral	0.85	1.00	0.92	41
Positif	0.94	0.68	0.79	22
Accuracy			0.88	67
Macro Avg	0.93	0.81	0.86	67
Weighted Avg	0.89	0.88	0.87	67

Berdasarkan hasil data dari confusion matrix ini, didapatkan sebuah interpretasi berupa f1-score yang ada pada kelas netral lebih tinggi 0,92 dibandingkan setiap label yang ada.



Gambar 8 C – Matrix SVM Form

Pada gambar 9 menjelaskan data memiliki variasi dalam mendeteksi isi dari tiap sentiment dikarenakan data memuat unsur yang relevant dalam konteks Kata yang cenderung berkaitan yang paling unik memiliki relevansi seperti bagus memiliki bobot nilai 0.055902, judul 0.021492 netral dan hoax untuk negative 0.053091 diambil data datasample scrapping dengan populasi sample porsi lebih besar dibandingkan form.



Gambar 9 (a) Wordlist Scapping (b) Wordlist Form

Intepretasi Algoritma Random Forest

Dalam analisis berikutnya dengan konteks sampel data yang sama memberikan hasil intepretasi berupa nilai accuracy yang sedikit meningkat meski

bentuk model cenderung bias terhadap sentimen dengan label netral tabel 3 sebesar 0.9054779806659506 yang dapat dilihat pada tabel 4 perbandingan dengan nilai bobot sebesar 0.8805970149253731

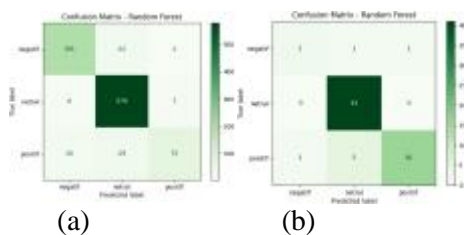
Tabel 3 C-Matrix SVM Scraping

Kelas	Preci	Recall	F1- sion	Sup port
Negatif	0.88	0.88	0.88	217
Netral	0.92	0.98	0.95	592
Positif	0.89	0.60	0.72	122
Accuracy			0.91	931
Macro Avg	0.90	0.82	0.85	931
Weighted Avg	0.90	0.91	0.90	931

Tabel 4 C-Matrix SVM form

Kelas	Preci	Recall	F1- sion	Sup port
Negatif	0.67	0.50	0.57	4
Netral	0.87	1.00	0.93	41
Positif	0.94	0.73	0.82	22
Accuracy			0.88	67
Macro Avg	0.83	0.74	0.77	67
Weighted Avg	0.88	0.88	0.87	67

Untuk mendukung data pendukung sebagai bentuk visualisasi lebih jelas juga dapat dilihat pada gambar 10 (a) scraping dan (b) form model cenderung memprediksi pada hasil labeling netral sama seperti model mempelajari hasil dari algoritma sebelumnya SVM.



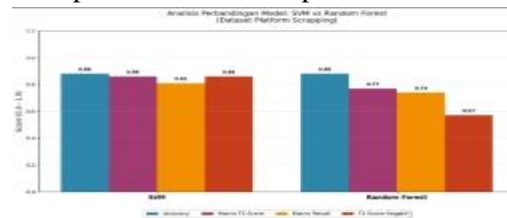
Gambar 10 Confusion Matrix SVM

Hasil visualiasi yang ada cenderung memperlihatkan sebuah model cenderung lazy load meskipun dataset labeling cukup kuat dikarenakan model tetap memprediksi netral lebih dominan

Perbandingan Performa Algoritma

Dalam pengujian yang dilakukan

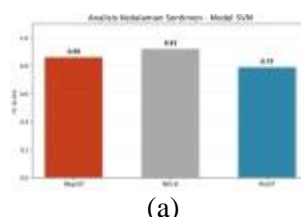
pada penelitian kali ini , penulis mengambill satu output opini analisis sentimen publik mengenai pembatasan pada media sosial anak hasil *scrapping* mengenai seberapa baik model dalam memprediksi sebuah output.



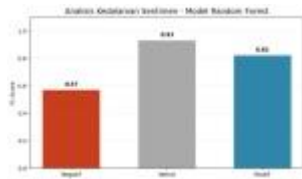
Gambar 12. Perbandingan Performa Algoritma Scrapper

Algoritma Support Vector Machine (SVM) dan Random Forest memiliki nilai akurasi yang sama 0,88. Hal ini menunjukkan bahwa secara umum kedua model mampu melakukan klasifikasi dengan tingkat ketepatan yang setara. Namun, jika dilihat dari perpektif metrik lebih komprehensif SVM menunjukkan performa yang lebih unggul dibandingkan Random Forest. Nilai macro F1-score pada SVM mencapai 0,86, lebih tinggi dibandingkan Random Forest yang hanya sebesar 0,77. Hal ini mengindikasikan bahwa SVM memiliki kemampuan yang lebih baik dalam menjaga keseimbangan performa klasifikasi di seluruh kelas.

Selain itu, pada metrik macro recall, SVM juga memperoleh nilai lebih tinggi yaitu 0,81 dibandingkan Random Forest yang menunjukkan SVM lebih solid. pada gambar 13 juga menjelaskan yang dapat dilihat berikut ini , dalam implementasi algoritma seperti SVM dan Random Forest tetap kelas netral yang memiliki dominasi dalam mendeteksi opini publik terkait larangan pembatasan media sosial pada anak dibawah umur 16 tahun. Namun dalam Random Forest terdapat penurunan yang cukup signifikan pada kelas negatif sebesar 0.57



(a)



(b)

Gambar 13 Perbandingan Performa Algoritma Opini Publik Form

Dalam sampling ini ditampilkan opini public langsung kepada para guru dan siswa mengenai bagaimana kebijakan yang ada dan memberikan sampling langsung dengan total kusioner 335 data. Hasil cenderung relatif sama dengan dominan pada kelas netral

SIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa proses analisis sentimen terhadap opini publik terkait kebijakan pembatasan akses media sosial bagi anak di bawah usia 16 tahun berhasil diimplementasikan melalui pipeline Natural Language Processing (NLP).

Distribusi data menunjukkan ketidakseimbangan kelas Dimana cenderung dominan di kelas netral, dimana kelas netral mendominasi jumlah data dibandingkan kelas positif dan negatif.

Hasil evaluasi menunjukkan bahwa kedua model memiliki performa akurasi yang relatif sama, yaitu sekitar 90%. Namun, jika ditinjau dari metrik evaluasi yang lebih komprehensif seperti macro F1-score dan recall, model SVM menunjukkan performa yang lebih unggul dibandingkan Random Forest. SVM menghasilkan nilai macro F1-score sebesar 0.85–0.87, sedangkan Random Forest berada pada kisaran 0.85 dengan kecenderungan bias terhadap kelas mayoritas.

Analisis confusion matrix menunjukkan bahwa model cenderung memiliki performa terbaik pada kelas netral dengan tingkat recall yang sangat tinggi (0.97–0.98). Namun, pada kelas

positif, nilai recall relatif lebih rendah (0.67–0.74), yang menunjukkan bahwa model masih mengalami kesulitan dalam mendeteksi sentimen positif, terutama yang bersifat implisit. Pola kesalahan menunjukkan bahwa banyak data positif dan negatif diklasifikasikan sebagai netral, yang mengindikasikan bahwa model bersifat konservatif dalam pengambilan keputusan.

DAFTAR PUSTAKA

- A. Amalia, A. L. Ramadhani, F. G. Vitacheria, and I. Azizah, "Pendidikan Karakter dan Teknologi: Pengaruh Penggunaan Media Sosial Terhadap Pembentukan Karakter Remaja," *J. Ilmu Pendidik. dan Pembelajaran*, vol. 3, no. 1, pp. 32–39, 2024, doi: 10.58706/jipp.v3n1.p32-39.
- D. T. Hunaifi, "Analisis Tren Penggunaan Filter Digital pada Pengguna Media Sosial Instagram di Indonesia," *Al Qalam J. Ilm. Keagamaan dan Kemasyarakatan*, vol. 17, no. 6, p. 4541, 2023, doi: 10.35931/aq.v17i6.2831.
- T. Arwindarti, E. I. Setiawan, and S. Imron, "Klasifikasi Sentimen Opini Publik Pada Instagram Pemerintah Kabupaten Bojonegoro Menggunakan LSTM," *Teknika*, vol. 13, no. 1, pp. 1–9, 2023, doi: 10.34148/teknika.v13i1.699.
- D. Fitriani, Y. Budiyan, A. R. Hardika, and M. Choerunissa, "Partisipasi Masyarakat Dalam Proses Demokrasi Di Indonesia: Analisis Peran Teknologi Dan Media Sosial," *Adv. Soc. Humanit. Res.*, vol. 1, no. 4, pp. 362–371, 2023, doi: 10.46799/adv.v1i4.43.
- E. Setyaningsih and D. Setyowatie, "E Sosialisasi Dampak Positif dan Negatif Penggunaan Gadget Serta Media Sosial di Kalangan Anak-anak dan Remaja," *IJCOSIN Indones. J. Community Serv. Innov.*, vol. 3, no. 1, pp. 64–71, 2023, doi:

- 10.20895/ijcosin.v3i1.919.
- M. Yasin and Siti Sri Fattul Jannah, “Penanggulangan Dampak Negatif Media Sosial Melalui Peran Guru dan Masyarakat di Sekolah,” *DIAJAR J. Pendidik. dan Pembelajaran*, vol. 1, no. 3, pp. 250–258, 2022, doi: 10.54259/diajar.v1i3.916.
- I. Pebriani, “Analisis Hukum Terhadap Perlindungan Korban Sextortion dan Eksploitasi Seksual Anak di Ranah Digital: Studi Kasus Laporan Komdigi Tahun 2024,” vol. 02, no. 2, pp. 16–30, 2025.
- N. Lailatul’Izza, “Upaya Penanaman Penggunaan Media Sosial dalam Melindungi Anak-Anak dari Dampak Negatif Media Sosial,” *JOIES (Journal Islam. Educ. Stud.)*, vol. 8, no. 2, pp. 232–254, 2023, [Online]. Available: <https://jurnalpps.uinsa.ac.id/index.php/joies/article/view/505%0Ahttps://jurnalpps.uinsa.ac.id/index.php/joies/article/download/505/276>
- Salsabila Amilda, Yasmin Luthfiah Sutari, Muhammad Arief Aqil Audi, Annisa Hafizhah, and Rosmalinda, “Perlindungan Hukum Terhadap Anak Korban Child Grooming Akibat Keingintahuan yang Salah dalam Penggunaan Media Sosial,” *Cendekia J. Hukum, Sos. dan Hum.*, vol. 3, no. 1, pp. 918–928, 2025, doi: 10.70193/cendekia.v3i1.173.
- A. Rahmadani, M. L. Paramita, S. Haura, and F. Firman, “Regulasi Digital Dan Implikasinya Terhadap Kebebasan Berpendapat (Studi Kasus: Uu Ite Pada Platform Media Sosial Di Indonesia),” *J. Soc. Contempl.*, vol. 2, no. 1, pp. 1–18, 2024, doi: 10.61183/jsc.v2i1.75.
- N. Alfarizi, S. Sinurat, A. Putra, M. Amin, P. Lydia, and U. P. Budi, “ANALISIS PENANGANAN DATA TIDAK SEIMBANG TERHADAP ULASAN MARKETPLACE TOKOPEDIA,” vol. 4307, no. 1, pp. 474–482, 2026.
- I. Siti Aisah, B. Irawan, and T. Suprapti, “Algoritma Support Vector Machine (Svm) Untuk Analisis Sentimen Ulasan Aplikasi Al Qur’an Digital,” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 6, pp. 3759–3765, 2024, doi: 10.36040/jati.v7i6.8263.
- S. Y. F. Noorihsan, T. D. A. Widhianingsih, and H. Kuswanto, “Predictive Analytics for Property Valuation Using Random Forest in Malang City,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 6, no. 1, pp. 186–192, 2026, [Online]. Available: <https://doi.org/10.57152/malcom.v6i1.2411>
- H. Haryanto, M. I. Sarif, and R. S. Hardinata, “Application of Data Mining to analyze the Best Suppliers Hendrie,” vol. 16, pp. 379–387, 2024.
- E. Erlin, Y. Desnelita, N. Nasution, L. Suryati, and F. Zoromi, “Dampak SMOTE terhadap Kinerja Random Forest Classifier berdasarkan Data Tidak seimbang,” *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 21, no. 3, pp. 677–690, 2022, doi: 10.30812/matrik.v21i3.1726.