BAB V
KESIMPULAN DAN SARAN

5.1. Kesimpulan

Penelitian ini menganalisis opini publik terhadap kebijakan pembatasan usia pengguna media sosial di Indonesia melalui analisis sentimen dengan IndoBERT dan pemodelan topik dengan BERTopic atas 33.910 komentar dari YouTube dan Instagram. Model IndoBERT hasil *fine-tuning* mencapai akurasi sebesar 0,9155 dan *Macro F1* sebesar 0,8744 pada data validasi, kemudian dievaluasi pada data test yang tidak pernah digunakan selama pelatihan maupun pemilihan model dan memperoleh akurasi sebesar 0,8500 serta *Macro F1* sebesar 0,8479. Hasil tersebut menunjukkan bahwa model memiliki kemampuan yang baik dalam mengklasifikasikan sentimen sekaligus mampu melakukan generalisasi terhadap data yang belum pernah dilihat sebelumnya. Berdasarkan hasil dan pembahasan, diperoleh kesimpulan sebagai berikut:

1. Distribusi sentimen publik terhadap kebijakan pembatasan usia pengguna media sosial cenderung positif. Sentimen positif mendominasi dengan proporsi 52,5 persen, diikuti sentimen negatif 26,7 persen dan netral 20,7 persen. Dominasi sentimen positif tersebut menunjukkan bahwa kebijakan ini lebih banyak memperoleh dukungan daripada penolakan pada ruang komentar kedua platform.
2. Pemodelan topik membentuk 26 topik yang menggambarkan berbagai pembahasan masyarakat terkait kebijakan pembatasan usia pengguna media sosial. Sepuluh topik dengan jumlah komentar terbanyak mencakup isu mekanisme pembuatan akun dan verifikasi usia, perlindungan anak di ruang digital, penggunaan platform digital seperti YouTube dan Roblox, penanganan judi *online*, peran orang tua, sosok menteri, serta perlindungan generasi muda. Temuan tersebut menunjukkan bahwa perhatian masyarakat tidak hanya tertuju pada tujuan kebijakan, tetapi juga pada mekanisme implementasi serta berbagai isu digital yang berkaitan dengan kebijakan tersebut.
3. Distribusi sentimen dan topik dominan berbeda secara nyata antara kedua platform. Sentimen pada Instagram cenderung positif sebesar 62,3 persen, sedangkan sentimen pada YouTube cenderung negatif sebesar 56,6 persen.

Selain itu, tema pembahasan pada kedua platform juga berbeda. Pembahasan di YouTube lebih banyak berkaitan dengan tontonan YouTube, permainan Roblox, dan judi *online*, sedangkan pada Instagram lebih banyak membahas perlindungan anak, ruang digital, peran orang tua, serta masa depan generasi. Perbedaan tersebut sebagian dipengaruhi oleh komposisi sumber komentar pada masing-masing platform, sehingga pola opini publik tidak dapat digeneralisasi tanpa memperhatikan platform asal komentar.

Seluruh kesimpulan tersebut dimaknai dalam lingkup komentar publik pada YouTube dan Instagram selama periode penelitian ini dan bukan merupakan representasi sikap seluruh masyarakat Indonesia.

5.2. Saran

Berdasarkan keterbatasan yang ditemui selama penelitian, beberapa saran yang dapat dipertimbangkan untuk penelitian selanjutnya adalah sebagai berikut.

1. Pemodelan topik menghasilkan proporsi *outlier* sebesar 21,0 persen akibat banyaknya komentar pendek dan tidak terstruktur. Penelitian selanjutnya dapat mengeksplorasi penyetelan parameter maupun pendekatan pemodelan topik lain agar cakupan komentar yang berhasil dikelompokkan menjadi lebih luas.
2. Komposisi sumber pada kedua platform tidak seimbang, yaitu komentar pada Instagram lebih banyak berasal dari akun pemerintah, sedangkan komentar pada YouTube lebih banyak berasal dari kanal berita. Penelitian selanjutnya disarankan menyeimbangkan komposisi sumber agar perbedaan pola antarplatform dapat dianalisis secara lebih terkendali.
3. Penelitian ini hanya mencakup dua platform, yaitu YouTube dan Instagram. Penelitian ini juga hanya mencakup komentar berbahasa Indonesia. Penelitian selanjutnya dapat memperluas cakupan ke platform lain, seperti TikTok atau Facebook, sehingga gambaran opini publik dapat diperoleh secara lebih menyeluruh.
4. Analisis pada penelitian ini terbatas pada sentimen dan topik. Penelitian selanjutnya dapat menambahkan dimensi analisis lain, seperti analisis emosi

atau pemetaan opini berdasarkan waktu, untuk memperoleh pemahaman yang lebih mendalam terhadap dinamika opini publik.

DAFTAR PUSTAKA

[1] H. K. Azzaakiyyah, "The Impact of Social Media Use on Social Interaction in Contemporary Society," *Technology and Society Perspectives (TACIT)*, vol. 1, no. 1, pp. 1–9, Aug. 2023, doi: 10.61100/tacit.v1i1.33.

[2] F. Ma'ruf, P. J. Pattiasina, R. Setiawati, B. C. F. Camerling, and P. E. Tuasela, "The Influence of Social Media Usage, Internet Access, and Mobile Device Penetration on Social Interaction Quality among Adolescents in Indonesia," *The Eastasouth Journal of Social Science and Humanities*, vol. 1, no. 03, Jun. 2024, doi: 10.58812/esssh.v1i03.275.

[3] Asosiasi Penyelenggara Jasa Internet Indonesia, "Profil Internet Indonesia 2026: Survei Penetrasi Internet dan Perilaku Penggunaan Internet 2026," APJII, Jakarta, Indonesia, 2026. [Online]. Available: https://survei.apjii.or.id

[4] S. Kemp, "Digital 2026: Indonesia," DataReportal, Nov. 2025. [Online]. Available: https://datareportal.com/reports/digital-2026-indonesia

[5] G. Cohen, E. Medina, D. Handysides, H. Shah, A. Arechiga, and W. Shih, "The impact of social media usage on lifestyle behaviors and health," *Lifestyle Medicine*, vol. 5, no. 1, p. e94, Jan. 2024, doi: 10.1002/lim2.94.

[6] E. R. Purboningsih, K. Massar, Z. R. Hinduan, H. Agustiani, R. A. C. Ruiter, and P. Verduyn, "Perception and use of social media by Indonesian adolescents and parents: A qualitative study," *Frontiers in Psychology*, vol. 13, p. 985112, Jan. 2023, doi: 10.3389/fpsyg.2022.985112.

[7] Badan Pusat Statistik, "Profil Anak Usia Dini 2025," Jakarta, Indonesia, 04200.25023, Dec. 2025. [Online]. Available: https://www.bps.go.id/id/publication/2025/12/05/0aa15c4692d4dbc27049aa33/profil-anak-usia-dini-2025.html

[8] B. Hayes, A. James, R. Barn, and D. Watling, "Children's risk and benefit behaviours on social networking sites," *Computers in Human Behavior*, vol. 130, p. 107147, May 2022, doi: 10.1016/j.chb.2021.107147.

[9] A. A. Auliya, A. B. Yahya, and F. K. Hurryos, "Pengaruh Penggunaan Media Sosial Terhadap Perilaku Remaja di Indonesia," *Jurnal Harmoni Nusa Bangsa*, vol. 1, no. 1, p. 57, Oct. 2023, doi: 10.47256/jhnb.v1i1.297.

[10] E. Bozzola *et al.*, "The Use of Social Media in Children and Adolescents: Scoping Review on the Potential Risks," *International Journal of Environmental Research and Public Health*, vol. 19, no. 16, p. 9960, Aug. 2022, doi: 10.3390/ijerph19169960.

[11] R. A. Fitri and N. F. Rahmadhani, "Kebijakan perlindungan Anak di ruang Digital: Perspektif Hukum Terhadap Cyberbullying," *Jurnal Analisis Hukum*, vol. 7, no. 2, pp. 192–199, Sep. 2024, doi: 10.38043/jah.v7i2.5186.

[12] K. L. McAlister, C. C. Beatty, J. E. Smith-Caswell, J. L. Yourell, and J. L. Huberty, "Social Media Use in Adolescents: Bans, Benefits, and Emotion Regulation Behaviors," *JMIR Mental Health*, vol. 11, pp. e64626–e64626, Nov. 2024, doi: 10.2196/64626.

[13] Parliament of Australia, *Online Safety Amendment (Social Media Minimum Age) Act 2024*. 2024. [Online]. Available: https://www.legislation.gov.au/C2024A00127/latest/details

[14] K. Prendergast and E. Dyer, "Youth social media age restrictions: examining Trans-Tasman media coverage of Australia's social media 'ban,'" *Journal of Youth Studies*, pp. 1–17, Feb. 2026, doi: 10.1080/13676261.2026.2626495.

[15] A. Humain-Lescop, “Towards harmonised online age verification? A comparative study of French and EU legal frameworks,” 2025, *SSRN*. doi: 10.2139/ssrn.5123971.
[16] A. Jamaludin, “Age Verification Regulation in Social Media Platform Usage: Preventive Measures Against Online Child Sexual Violence,” *Jurnal Bina Mulia Hukum*, vol. 8, no. 2, pp. 276–295, Mar. 2024, doi: 10.23920/jbmh.v8i2.1311.
[17] *Peraturan Pemerintah Nomor 17 Tahun 2025 tentang Tata Kelola Penyelenggaraan Sistem Elektronik dalam Pelindungan Anak*. 2025. [Online]. Available: https://peraturan.bpk.go.id/Details/316698/pp-no-17-tahun-2025
[18] Kementerian Komunikasi dan Digital Republik Indonesia, “Tentang PP Tunas – Tunas Digital,” Tunas Digital. [Online]. Available: https://tunasdigital.id/tentang-pp-tunas/
[19] Kementerian Komunikasi dan Digital Republik Indonesia, *Peraturan Menteri Komunikasi dan Digital Nomor 9 Tahun 2026 tentang Peraturan Pelaksanaan Peraturan Pemerintah Nomor 17 Tahun 2025 tentang Tata Kelola Penyelenggaraan Sistem Elektronik dalam Pelindungan Anak*, vol. Tahun 2026 Nomor 163. 2026, p. 29. [Online]. Available: https://jdih.komdigi.go.id/produk_hukum/view/id/1007/t/peraturan+menteri+komunikasi+dan+digital+nomor+9+tahun+2026
[20] Kementerian Komunikasi dan Digital Republik Indonesia, “PP TUNAS Berlaku, Platform Wajib Batasi Akses Anak,” Kementerian Komunikasi dan Digital. [Online]. Available: https://www.komdigi.go.id/berita/siaran-pers/detail/pp-tunas-berlaku-platform-wajib-batasi-akses-anak
[21] A. Fathuurrahmaan, “Analisis Filsafat Kebijakan Peraturan Menteri Komdigi No. 9 Tahun 2026 Tentang Pembatasan Penggunaan Media Sosial Pada Anak,” *Arus Jurnal Sosial dan Humaniora*, vol. 6, no. 1, pp. 827–833, Apr. 2026, doi: 10.57250/ajsh.v6i1.2468.
[22] I. Fitriyani, “Pemerintah Batasi Akses Media Sosial Anak: Langkah Progresif atau Pengekangan?,” *Sultra Media*, Mar. 07, 2026. [Online]. Available: https://www.sultramedia.id/pemerintah-batasi-akses-media-sosial-anak-langkah-progresif-atau-pengekangan
[23] Y. Gorodnichenko, T. Pham, and O. Talavera, “Social media, sentiment and public opinions: Evidence from #Brexit and #USElection,” *European Economic Review*, vol. 136, p. 103772, Jul. 2021, doi: 10.1016/j.euroecorev.2021.103772.
[24] A. M. A. Ausat, “The Role of Social Media in Shaping Public Opinion and Its Influence on Economic Decisions,” *Technology and Society Perspectives (TACIT)*, vol. 1, no. 1, pp. 35–44, Aug. 2023, doi: 10.61100/tacit.v1i1.37.
[25] F. Guo, Z. Liu, Q. Lu, S. Ji, and C. Zhang, “Public Opinion About COVID-19 on a Microblog Platform in China: Topic Modeling and Multidimensional Sentiment Analysis of Social Media,” *Journal of Medical Internet Research*, vol. 26, p. e47508, Jan. 2024, doi: 10.2196/47508.
[26] K. Nandini and M. Rahardi, “Sentiment Analysis of Economic Policy Comments on YouTube Using Ensemble Machine Learning,” *Journal of Applied Informatics and Computing*, vol. 9, no. 5, pp. 2607–2615, Oct. 2025, doi: 10.30871/jaic.v9i5.10897.
[27] M. R. Manoppo *et al.*, “ANALISIS SENTIMEN PUBLIK DI MEDIA SOSIAL TERHADAP KENAIKAN PPN 12% DI INDONESIA MENGGUNAKAN INDOBERT,” *Jurnal Kecerdasan Buatan dan Teknologi Informasi*, vol. 4, no. 2, pp. 152–163, May 2025, doi: 10.69916/jkbti.v4i2.322.

[28] V. P. Kalanjati *et al.*, “Sentiment analysis of Indonesian tweets on COVID-19 and COVID-19 vaccinations,” *F1000Research*, vol. 12, p. 1007, Apr. 2024, doi: 10.12688/f1000research.130610.4.
[29] L. Septian, T. Aljauza, and C. Juliane, “Analisis Sentimen Putusan Mahkamah Konstitusi terhadap Batas Usia Capres dan Cawapres Menggunakan IndoBERT,” *The Indonesian Journal of Computer Science*, vol. 12, no. 6, Jan. 2024, doi: 10.33022/ijcs.v12i6.3614.
[30] A. Mar’atusholihat, N. Ghurroh Setyoningrum, and D. Rizal Nursamsi, “Perbandingan Kinerja Algoritma Decision Tree Dan Naïve Bayes Dalam Analisis Sentimen Publik Terhadap Kebijakan Pemerintah Mengenai Batasan Penggunaan Media Sosial Anak,” *Informatics and Digital Expert (INDEX)*, vol. 7, no. 2, pp. 128–134, Nov. 2025, doi: 10.36423/index.v7i2.2407.
[31] W. M. Sihombing and T. Widiyaningtyas, “Sentiment Analysis and Topic Modelling Using IndoBERTweet and BERTopic for Public Health Issues: Analisis Sentimen dan Pemodelan Topik Menggunakan IndoBERTweet dan BERTopic untuk Isu Kesehatan Publik,” *Indonesian Journal of Innovation Studies*, vol. 26, no. 4, Nov. 2025, doi: 10.21070/ijins.v26i4.1833.
[32] F. N. Khaqqi, L. Alfat, and I. Nurhaida, “Sentiment Analysis of US-China Tariffs using IndoBERT and Economic Impact on Indonesia,” *Journal of Applied Informatics and Computing*, vol. 9, no. 6, pp. 3000–3011, Dec. 2025, doi: 10.30871/jaic.v9i6.11544.
[33] M. W. R. Stevens *et al.*, “Beyond the debate: toward pragmatic evaluation of Australia’s social media age restrictions,” *The Lancet Regional Health - Western Pacific*, vol. 68, p. 101818, Mar. 2026, doi: 10.1016/j.lanwpc.2026.101818.
[34] V. D. Setiawan, D. U. Iswavigra, and E. Anggiratih, “Implementation of IndoBERT for Sentiment Analysis of the Constitutional Court’s Decision Regarding the Minimum Age of Vice Presidential Candidates,” *Scientific Journal of Informatics*, vol. 12, no. 3, pp. 397–406, Aug. 2025, doi: 10.15294/sji.v12i3.26320.
[35] S. I. Ishak, O. Arnilia, T. Widodo, and I. G. N. A. B. Tatwa, “Analisis Sentimen terhadap Pemerintahan Prabowo–Gibran menggunakan IndoBERT dan LDA,” *Jambura Journal of Informatics*, vol. 1, no. 2, pp. 72–82, Oct. 2025, doi: 10.37905/jji.v1i2.34895.
[36] M. R. Nur, Y. Wibisono, and R. Megasari, “Analisis Sentimen dan Pemodelan Topik pada Post tentang Merek Teknologi di X Menggunakan Fine-tuning IndoBERT dan BERTopic,” *Jurnal Komputer Teknologi Informasi Sistem Informasi (JUKTISI)*, vol. 4, no. 2, pp. 743–750, Jul. 2025, doi: 10.62712/juktisi.v4i2.508.
[37] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, Minneapolis, Minnesota: Association for Computational Linguistics, 2019, pp. 4171–4186. doi: 10.18653/v1/N19-1423.
[38] B. Wilie *et al.*, “IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding,” in *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing*, Suzhou, China: Association for Computational Linguistics, 2020, pp. 843–857. doi: 10.18653/v1/2020.aacl-main.85.
[39] M. Grootendorst, “BERTopic: Neural topic modeling with a class-based TF-IDF procedure,” 2022, *arXiv*. doi: 10.48550/ARXIV.2203.05794.

[40] N. Newman, “Digital News Report 2025,” Reuters Institute for the Study of Journalism, University of Oxford, Oxford, U.K., 2025. [Online]. Available: https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2025
[41] Katadata Insight Center, “Politik di Mata Anak Muda: Persepsi dan Kecenderungan Gen Z & Milenial terhadap Capres, Parpol, dan Kampanye Pemilu 2024,” Katadata Insight Center, Jakarta, Indonesia, Survei Nasional, Nov. 2023. [Online]. Available: https://databoks.katadata.co.id/publikasi/2023/11/14/politik-di-mata-anak-muda
[42] D. Bai and Y. Gu, “Harnessing Big Data, Hindered by Bias: Evaluating TikTok Research API for Fair and Optimal Social Sciences,” *Social Science Computer Review*, Jan. 2026, doi: 10.1177/08944393251413277.
[43] Dewan Pers, “Data Perusahaan Pers,” Dewan Pers. [Online]. Available: https://dewanpers.or.id/data
[44] I. Balazevic, M. Braun, and K.-R. Müller, “Language Detection For Short Text Messages In Social Media,” 2016, *arXiv*. doi: 10.48550/ARXIV.1608.08515.
[45] V. Batanović, M. Cvetanović, and B. Nikolić, “A versatile framework for resource-limited sentiment articulation, annotation, and analysis of short texts,” *PLOS ONE*, vol. 15, no. 11, p. e0242050, Nov. 2020, doi: 10.1371/journal.pone.0242050.
[46] X. Liu *et al.*, “Developing Multi-Labelled Corpus of Twitter Short Texts: A Semi-Automatic Method,” *Systems*, vol. 11, no. 8, p. 390, Aug. 2023, doi: 10.3390/systems11080390.
[47] C. A. Melton, B. M. White, R. L. Davis, R. A. Bednarczyk, and A. Shaban-Nejad, “Fine-tuned Sentiment Analysis of COVID-19 Vaccine–Related Social Media Data: Comparative Study,” *Journal of Medical Internet Research*, vol. 24, no. 10, Oct. 2022, doi: 10.2196/40408.
[48] A. Karanikola, C. M. Liapis, and S. Kotsiantis, “Reducing labeled data requirements in text classification with active learning and BERT-based transformers,” *Neural Computing and Applications*, vol. 38, no. 2, p. 26, Jan. 2026, doi: 10.1007/s00521-025-11756-8.

LAMPIRAN

PANDUAN ANOTASI SENTIMEN

Analisis Sentimen dan Pemodelan Topik Opini Publik terhadap Kebijakan Pembatasan Usia Pengguna Media Sosial di Indonesia

Menggunakan IndoBERT dan BERTopic

Platform Data: YouTube & Instagram

Data: 3.000 komentar | Label: POSITIVE | NEGATIVE | NEUTRAL

UNIVERSITAS ATMA JAYA YOGYAKARTA

1. Tujuan dan Ruang Lingkup

Dokumen ini menjadi acuan tunggal bagi anotator dalam melakukan pelabelan sentimen pada data komentar media sosial. Panduan disusun secara operasional agar label dapat diberikan secara konsisten, terutama pada kasus yang berpotensi ambigu.

Empat tujuan panduan ini:

- Memberikan definisi operasional yang jelas untuk setiap label sentimen
- Menjamin konsistensi penilaian, baik saat anotasi mandiri maupun bersama teman
- Menyediakan aturan keputusan untuk kasus sulit, khususnya sarkasme dan komentar campuran
- Menghasilkan data anotasi yang reliabel sebagai data latih model IndoBERT

Ruang lingkup data:

- Platform: YouTube (1.035 komentar) dan Instagram (1.965 komentar)
- Topik: opini publik terhadap kebijakan pembatasan usia pengguna media sosial di Indonesia
- Total data: 2.000 komentar
- Tiga kelas label: POSITIVE, NEGATIVE, NEUTRAL (single-label per komentar)

---

2. Definisi Operasional Label Sentimen

Label sentimen mengacu pada penilaian penulis komentar terhadap kebijakan pembatasan usia pengguna media sosial secara keseluruhan. Sentimen dinilai berdasarkan isi teks komentar, bukan asumsi atau latar belakang penulis.

Label	Definisi Operasional	Kondisi Terpenuhi
POSITIVE	Komentar yang secara eksplisit atau implisit mengungkapkan dukungan, persetujuan, pujian, harapan positif, atau apresiasi terhadap kebijakan pembatasan usia pengguna media sosial di Indonesia.	• Menyatakan setuju, mendukung, atau memuji kebijakan • Menyambut baik atau mengungkapkan harapan positif • Menilai kebijakan tepat, perlu, atau bermanfaat • Mengajak orang lain mendukung kebijakan

Label	Definisi Operasional	Kondisi Terpenuhi
NEGATIVE	Komentar yang mengungkapkan penolakan, kritik, keraguan serius, kekecewaan, atau ketidaksetujuan terhadap kebijakan pembatasan usia pengguna media sosial.	• Menyatakan tidak setuju, menolak, atau mengecam kebijakan • Meragukan efektivitas atau kelayakan kebijakan • Mengkritik cara implementasi atau dampaknya • Mengungkapkan frustrasi atau kekecewaan terhadap kebijakan
NEUTRAL	Komentar yang menyampaikan fakta, pertanyaan, informasi, atau pernyataan deskriptif tanpa evaluasi emosional yang jelas terhadap kebijakan.	• Mengajukan pertanyaan teknis/praktis tanpa nada emosional • Menyebutkan fakta atau informasi tanpa evaluasi • Membandingkan tanpa menyimpulkan mana yang lebih baik • Komentar tidak relevan dengan kebijakan

Catatan penting:

- NEUTRAL bukan label default saat ragu. Gunakan NEUTRAL hanya jika komentar memang tidak mengandung evaluasi emosional yang jelas.
- Label dinilai dari perspektif penulis komentar, bukan sudut pandang anotator.
- Sentimen dinilai terhadap kebijakan ini secara spesifik, bukan terhadap pemerintah atau media sosial secara umum.

3. Contoh Komentar Berlabel dari Data Nyata

Seluruh contoh berikut diambil langsung dari dataset yang akan dianotasi (sample_for_labeling.json). Komentar ditampilkan dalam bentuk yang sudah dibersihkan; emoji digambarkan dalam kurung siku.

Komentar (dari dataset)	Label	Alasan
kalau kebijakan pemerintah yg satu ini saya sangat se7 (setuju) 👏	**POSITIVE**	"se7" = setuju. Persetujuan eksplisit dan antusias.
sangat sangat setuju	**POSITIVE**	Persetujuan eksplisit meski sangat singkat.
akhirnya ada berita baik dari nkri	**POSITIVE**	"Akhirnya" menunjukkan sambutan baik dan apresiasi.
peraturan bagus rakyat sudah mendukung tinggal implementasi dan pengawasan	**POSITIVE**	Penilaian utama positif. Catatan implementasi adalah harapan, bukan penolakan.
setuju tapi terlambat, harusnya dari dulu	**POSITIVE**	Dominan positif. Kekecewaan atas keterlambatan tidak mengubah dukungan inti.
pemerintah h4r4m (haram) aku benci pemerintah	**NEGATIVE**	Kemarahan eksplisit terhadap pemerintah. Karena komentar ini berasal dari konteks unggahan tentang kebijakan pembatasan usia, kemarahan dapat diatribusikan pada kebijakan tersebut — bukan kemarahan umum.
masih ga efektive percayalah tgl lahir bisa diganti haha	**NEGATIVE**	Meragukan efektivitas. "haha" di sini bernada sinis (sarkasme), bukan tawa tulus.
emang bisa? orang daftar akun medsos bisa akalin tgl lahir	**NEGATIVE**	Pertanyaan retoris yang secara implisit mengkritik kelemahan kebijakan.
bagaimana teknis menjaringnya teknis verifikasinya	**NEUTRAL**	Pertanyaan teknis murni tanpa nada emosional.
dengerin membaca berita di tv	**NEUTRAL**	Sepenuhnya off-topic, tidak menyinggung kebijakan.

4. Aturan Pengambilan Keputusan (Decision Rules)

4.1 Alur Keputusan Utama

Gunakan alur berikut secara berurutan untuk setiap komentar. Berhenti di langkah pertama yang menghasilkan label yang jelas.

No.	Pertanyaan	Tindakan
1	Apakah komentar berkaitan dengan kebijakan pembatasan usia pengguna media sosial (bukan topik lain sepenuhnya)?	Jika Ya -> lanjut ke No. 2. Jika Tidak -> NEUTRAL.
2	Apakah ada tanda-tanda sarkasme atau ironi? (Lihat Bagian 4.2)	Jika Ya -> baca makna yang dimaksud. Jika ragu -> NEUTRAL + tandai ulang. Jika Tidak -> lanjut ke No. 3.
3	Apakah komentar mengandung unsur pro DAN kontra sekaligus?	Jika Ya -> tentukan sentimen dominan (No. 4). Jika Tidak -> lanjut ke No. 5.
4	Manakah sentimen yang lebih dominan?	Dominan positif -> POSITIVE. Dominan negatif -> NEGATIVE. Seimbang -> NEUTRAL.
5	Apakah komentar secara keseluruhan mendukung/memuji/menyetujui kebijakan?	Jika Ya -> POSITIVE.
6	Apakah komentar secara keseluruhan menolak/mengkritik/meragukan kebijakan?	Jika Ya -> NEGATIVE. Jika Tidak -> NEUTRAL.

4.2 Panduan Khusus: Sarkasme dan Ironi

Sarkasme adalah penggunaan kata-kata yang berlawanan dengan makna yang sebenarnya dimaksud, biasanya untuk mengkritik atau menyindir. Berdasarkan inspeksi dataset, sarkasme ditemukan pada sekitar 35 komentar dan merupakan kasus yang paling sering menyebabkan ketidakkonsistenan antar-anotator. Bagian ini memberikan panduan langkah demi langkah untuk mengenali dan melabelinya dengan benar.

Langkah Mengenali Sarkasme

1. Baca seluruh komentar dari awal sampai akhir, jangan berhenti di kalimat pertama.
2. Cari satu atau lebih dari empat sinyal sarkasme berikut:

Sinyal	Penjelasan	Contoh dari Dataset
Sinyal 1: Tawa (wkwk/haha) + kalimat kritik	Tawa muncul setelah atau bersamaan dengan pernyataan yang meragukan atau mengkritik kebijakan. Ini hampir selalu sarkasme.	"masih ga efektive percayalah tgl lahir bisa diganti haha" (id:502) "Wkwkwk tinggal d setting ke usia atas" (id:213)
Sinyal 2: Pertanyaan retoris meremehkan	Pertanyaan yang jawabannya sudah jelas-jelas negatif, diajukan untuk menyindir bukan sungguh bertanya.	"Emang bisa? Orang daftar akun medsos bisa akalin tgl lahir" (id:331) "mana mungkin bisa diatasi, anak2 tidak bodoh" (id:71)
Sinyal 3: Pujian berlebihan + konteks negatif	Kata "luar biasa", "keren", "mantap" disertai konteks atau kalimat lain yang sinis atau menghina.	"sekarang tau kan kenapa Indo selalu ketinggalan" (id:1880) "biar kelihatan kerja yakk" (id:53)
Sinyal 4: Saran mustahil / hiperbolis	Menyarankan solusi yang jelas berlebihan sebagai sindiran bahwa kebijakan tidak akan berhasil.	"kalau mampuh sekalian aja putus internet" (id:71) "dipikir rakyat itu bodoh wkwkwk" (id:548)

3. Jika satu atau lebih sinyal ditemukan, tanyakan: apakah makna yang DIMAKSUD penulis berbeda dari makna LITERAL kata-katanya?
4. Jika ya, sarkasme terkonfirmasi -> gunakan makna yang dimaksud untuk menentukan label (hampir selalu NEGATIVE).
5. Jika masih ragu setelah memeriksa semua sinyal -> beri label NEUTRAL dan tandai untuk ditinjau ulang.

Perbandingan: Positif Tulus vs Sarkasme dari Dataset

Kedua kolom di bawah ini sama-sama menggunakan kata pujian ("luar biasa", "keren", "mantap"), tetapi labelnya berbeda. Perhatikan perbedaan pada ada atau tidaknya sinyal sarkasme:

POSITIF TULUS -> POSITIVE	SARKASME / IRONI -> NEGATIVE
"Luar biasa nih ide nya! Di era yang serba digital sekarang, perlindungan anak-anak di ranah ini jadi sangat penting." (id:511) Kenapa positif: kalimat panjang dengan argumen konkret tentang manfaat, tidak ada sinyal sinis.	"Wkwkwkwkwkwk sekarang tau kan kenapa Indo selalu ketinggalan?" (id:1880) Kenapa negatif: tawa mengawali pertanyaan retoris yang menghina kualitas kebijakan. Sinyal 1 + Sinyal 2 muncul bersamaan.

POSITIF TULUS -> POSITIVE	SARKASME / IRONI -> NEGATIVE
"Mantap sekali pakm mensos.. Kemensos benar-benar hadir melindungi anak-anak." (id:918) Kenapa positif: menggunakan argumen konkret, tidak ada kontradiksi atau sinyal sinis dalam komentar.	"Cara nya gmana cuy wkwkwkwkwkwk judol aja ngak diapa apain biar kelihatan kerja yakk" (id:53) Kenapa negatif: "biar kelihatan kerja" = sindiran bahwa kebijakan hanya pencitraan. Sinyal 3 jelas.
"Kemensos luar biasa. Pengawasan aktifitas digital anak untuk melindungi masa depan." (id:1184) Kenapa positif: frasa mendukung, tidak ada kontradiksi, argumen substantif tentang perlindungan anak.	"dipikir rakyat itu bodoh? wkwkwk konsep aja kau ga ngerti, yg dibatasi itu sosmed bukan hpnya" (id:548) Kenapa negatif: pertanyaan retoris + tawa + kalimat mengkritik secara langsung. Sinyal 1 + Sinyal 2.

Aturan Cepat Sarkasme (untuk referensi saat anotasi)

- Tawa (wkwk/haha) + kalimat kritik = hampir selalu NEGATIVE.
- Tawa + kalimat positif dengan argumen konkret dan substantif = kemungkinan tawa tulus -> POSITIVE.
- Pujian berlebihan TANPA argumen substantif + ada konteks negatif = sarkasme -> NEGATIVE.
- Pujian + argumen konkret tentang manfaat kebijakan = positif tulus -> POSITIVE.
- Pertanyaan retoris meremehkan ("emang bisa?", "mana mungkin?") = NEGATIVE.
- Jika masih ragu setelah memeriksa empat sinyal -> NEUTRAL + tandai untuk ditinjau ulang.

4.3 Penanganan Kasus Sulit Lainnya

Jenis Kasus	Aturan Keputusan	Contoh Komentar	Label
Komentar campuran (pro + kontra)	Tentukan sentimen DOMINAN. Jika "setuju" adalah posisi utama dan kritik hanya pada teknis implementasi -> POSITIVE. Jika inti argumen adalah penolakan -> NEGATIVE.	"setuju tapi bagaimana kalau mereka mendaftar dengan tanggal lahir yang tidak sebenarnya"	**POSITIVE (dominan setuju)**
Campuran seimbang	Jika kekuatan argumen pro dan kontra benar-benar setara -> NEUTRAL.	"oke dibatasi 16 tahun, tapi jika orang tua	**NEUTRAL**

Jenis Kasus	Aturan Keputusan	Contoh Komentar	Label
		buatkan akun justru berbahaya"	
Komentar ambigu	Jika tidak cukup konteks -> NEUTRAL. Jangan menebak.	"hmm menarik" / "liat aja nanti"	**NEUTRAL**
Komentar 1-3 kata	Terapkan aturan normal. Kata bermakna jelas tetap dilabeli sesuai makna.	"sangat setuju" / "percuma"	**POSITIVE / NEGATIVE**
Off-topic / tidak relevan	NEUTRAL. Tidak menyinggung kebijakan sama sekali.	"dengerin membaca berita di tv"	**NEUTRAL**
Bahasa kasar / makian	Tentukan TARGET makian. Marah pada kebijakan -> NEGATIVE. Marah pada netizen lain -> NEUTRAL.	"aku benci pemerintah" (konteks kebijakan)	**NEGATIVE**
Kritik satu aspek kecil	Menolak satu aspek kecil tapi menerima kebijakan secara umum -> POSITIVE.	"setuju tapi terlambat, harusnya dari dulu"	**POSITIVE**
Pertanyaan retoris negatif	Pertanyaan yang implisit mengkritik kebijakan -> NEGATIVE.	"emang bisa orang akalin tgl lahir?"	**NEGATIVE**
Slang & obfuskasi	Normalisasi dulu: se7->setuju, h4r4m->haram, s3tuju->setuju, trs->terus.	"saya sangat se7" / "pemerintah h4r4m"	**POSITIVE / NEGATIVE**
Emoji sebagai sentimen	Emoji ikut menentukan sentimen. 👏🙏✅ cenderung positif; 🙄😒❌ cenderung negatif; 😂🤣 perlu dilihat konteksnya (tawa gembira vs sinis).	"setuju banget 🔥" / "ya bodo amat 🙄"	**POSITIVE / NEGATIVE**

4.4 Kata Kunci dan Frasa Panduan per Label

Tabel berikut memuat sinyal linguistik khas per label. Gunakan sebagai panduan cepat, tetapi selalu periksa konteks keseluruhan komentar karena satu kata bisa bermakna berbeda tergantung konteks.

Label	Kata Kunci & Frasa	Frasa Khas
POSITIVE	setuju, se7, dukung, mendukung, bagus, benar, perlu, penting, tepat, mantap, keren, akhirnya, alhamdulillah, bismillah, harapan, semoga, bravo, gebrakan	"sudah seharusnya", "langkah yang tepat", "harus segera diterapkan", "sangat setuju", "ini yang ditunggu-tunggu"
NEGATIVE	percuma, gagal, tidak efektif, omong kosong, tidak setuju, sia-sia, salah, benci, bohong, tidak mungkin, susah, akal-akalan, h4r4m (kemarahan)	"gampang diakali", "tidak ada gunanya", "cuma wacana", "tgl lahir bisa diganti", "emang bisa?"
NEUTRAL	apa, bagaimana, kapan, di mana, siapa, berapa, entah, mungkin, menurut, katanya, konon	"ada yang tahu?", "bagaimana cara verifikasinya?", "di negara lain sudah...", "liat aja nanti"

5. Analisis Dataset dan Rekomendasi Aturan Tambahan

5.1 Statistik Deskriptif Dataset

- Total komentar: 3.000
- Platform: Instagram (1.965 komentar, 65,5%) dan YouTube (1.035 komentar, 34,5%)
- Sumber: 9 akun, terbanyak kemkomdigi (262), tvOneNews (184), teknokompas/kemendikdasmen/detikcom (masing-masing 131)
- Panjang komentar: 3-322 kata, rata-rata 15,1 kata per komentar
- Komentar sangat pendek (1-3 kata): sekitar 60 komentar

5.2 Pola Ambiguitas dan Rekomendasi

Pola Komentar	Est. Jumlah	Risiko Anotasi	Rekomendasi Aturan
Komentar campuran (setuju + kritik implementasi)	~200	**TINGGI**	Dominansi sentimen: jika "setuju" adalah posisi utama +

Pola Komentar	Est. Jumlah	Risiko Anotasi	Rekomendasi Aturan
			argumen positif lebih banyak -> POSITIVE.
Sarkasme dengan wkwk/haha/lol	~35	**TINGGI**	Gunakan empat sinyal di Bagian 4.2. Cek konteks sebelum dan sesudah tawa.
Komentar sangat pendek (1-3 kata)	~60	**SEDANG**	Kata bermakna jelas tetap dilabeli. "setuju" = POSITIVE. "percuma" = NEGATIVE. Tanpa makna jelas = NEUTRAL.
Pertanyaan teknis verifikasi	~80	**SEDANG**	Murni bertanya -> NEUTRAL. Pertanyaan dengan kritik implisit ("emang bisa?") -> NEGATIVE.
Komentar tentang game bukan kebijakan	~120	**SEDANG**	Off-topic murni -> NEUTRAL. Menyebut game dalam konteks menolak kebijakan -> NEGATIVE.
Obfuskasi karakter (h4r4m, se7)	~30	**RENDAH**	Normalisasi dulu: se7=setuju, h4r4m=haram, 1tu=itu, d1=di, trs=terus.
Komentar anak/remaja kecewa tidak bisa main medsos	~150	**SEDANG**	Kekecewaan karena tidak bisa main game/medsos yang merujuk kebijakan -> NEGATIVE.
"luar biasa", "keren", "mantap" tanpa ironi jelas	~50	**SEDANG**	Jika tidak ada sinyal sarkasme -> POSITIVE. Bandingkan dengan tabel di Bagian 4.2.

5.3 Aturan Tambahan Spesifik untuk Dataset Ini

1. Komentar tentang game (Roblox, Free Fire, Minecraft) yang tidak mengacu kebijakan -> NEUTRAL. Jika menyebut game dalam konteks menolak kebijakan -> NEGATIVE.
2. Komentar anak/remaja yang kecewa tidak bisa main medsos karena kebijakan -> NEGATIVE.

3. Komentar mendukung kebijakan tapi meminta cakupan lebih luas ("sekalian judol diblokir") -> POSITIVE.
4. Ekspresi "akhirnya", "sudah lama", "telat tapi lebih baik" -> POSITIVE.
5. Komentar tentang teknis orang tua membuatkan akun: nilai berdasarkan arah argumen. Mendukung -> POSITIVE; mengkritik celah -> NEGATIVE.

---

6. Alur Kerja Teknis Anotasi

6.1 Langkah-Langkah Anotasi

1. Buka antarmuka anotasi (Label Studio atau platform lainnya).
2. Baca komentar secara keseluruhan tanpa langsung memberi label.
3. Ikuti Alur Keputusan Utama (Bagian 4.1).
4. Jika ada tanda sarkasme, gunakan Panduan Sarkasme (Bagian 4.2).
5. Jika ada opini campuran, tentukan sentimen dominan.
6. Tentukan label final: POSITIVE, NEGATIVE, atau NEUTRAL.
7. Jika tidak dapat diputuskan dengan yakin, tandai untuk ditinjau ulang. Jangan menebak.
8. Lakukan pengecekan akhir menggunakan Checklist (Bagian 8) sebelum submit.

6.2 Panduan Konsistensi untuk Anotasi Mandiri (Solo Annotator)

Karena sebagian besar anotasi akan dikerjakan sendiri, berikut strategi untuk menjaga konsistensi selama proses pelabelan 2.000 data:

Strategi	Cara Pelaksanaan
Calibration awal	Labeli 50-100 komentar pertama, lalu istirahat minimal 1 hari. Kembali dan tinjau ulang tanpa melihat label sebelumnya. Hitung berapa persen yang konsisten. Target: >= 80%.
Spot-check berkala	Setiap 200-300 komentar selesai, ambil 20 komentar acak dari batch sebelumnya dan labeli ulang tanpa melihat label lama. Jika ada perbedaan, tinjau aturan panduan yang relevan.
Catat edge case	Buat catatan komentar yang sulit. Setelah satu sesi, putuskan secara konsisten dan simpan keputusan agar sesi berikutnya mengikuti pola yang sama.
Batasi sesi anotasi	Anotasi lebih dari 2-3 jam berturut-turut menurunkan konsistensi. Pecah menjadi sesi 200-300 komentar per sesi dengan jeda.

Strategi	Cara Pelaksanaan
Kolaborasi minimal (sangat disarankan)	Ajak 1 teman melabeli 100 komentar yang sama secara mandiri. Hitung Cohen's Kappa (target k >= 0.70). Ini memperkuat validitas panduan di hadapan dosen pembimbing saat sidang.

6.3 Strategi Semi-Otomatis (Active Learning)

Pre-labeling dengan model IndoBERT:

- Gunakan model IndoBERT yang sudah dilatih (checkpoint-v1, Macro F1 = 0.675) untuk membuat prediksi awal.
- Anotator cukup memverifikasi prediksi, bukan membuat label dari nol. Estimasi penghematan waktu: 40-60%.

Prioritasi komentar sulit:

- Label dahulu komentar dengan confidence rendah (probabilitas tertinggi di bawah 0.70).
- Komentar dengan confidence tinggi (di atas 0.90) dapat divalidasi lebih cepat dengan spot-check 10-20%.

7. Kendali Kualitas (Quality Control)

- Perbarui panduan jika ditemukan pola ambiguitas baru yang belum tercakup.

8. Checklist Anotator (Sebelum Submit Setiap Komentar)

Gunakan checklist ini untuk setiap komentar. Jika satu item tidak terpenuhi, tinjau ulang label yang diberikan.

No.	Item Checklist	Centang
1	Saya sudah membaca komentar secara keseluruhan sebelum memberi label.	[ ]
2	Saya memeriksa apakah komentar berkaitan dengan kebijakan pembatasan usia pengguna media sosial.	[ ]
3	Saya memeriksa empat sinyal sarkasme (Bagian 4.2) sebelum menentukan label.	[ ]

No.	Item Checklist	Centang
4	Jika komentar mengandung unsur pro dan kontra, saya telah menentukan sentimen yang dominan.	[ ]
5	Saya tidak menggunakan NEUTRAL sebagai label default saat ragu — NEUTRAL hanya untuk komentar faktual atau pertanyaan murni.	[ ]
6	Untuk komentar sangat pendek (1-3 kata), saya tetap melabeli berdasarkan makna jelas.	[ ]
7	Saya telah menormalisasi slang/obfuskasi (se7->setuju, h4r4m->haram) sebelum menentukan label.	[ ]
8	Emoji tertawa saya nilai berdasarkan konteks (tawa tulus vs sinis/meremehkan).	[ ]
9	Komentar yang sepenuhnya off-topic saya beri label NEUTRAL.	[ ]
10	Kasus yang tidak dapat saya putuskan dengan yakin sudah saya tandai untuk ditinjau ulang.	[ ]