BAB I
PENDAHULUAN

1.1 Latar Belakang

YouTube telah berkembang menjadi platform berbagi video terbesar di dunia dengan lebih dari 2,7 miliar pengguna aktif bulanan secara global pada tahun 2024. Di Indonesia, YouTube memiliki 139 juta pengguna aktif, menjadikannya salah satu pasar YouTube terbesar di Asia Tenggara (Kemp, 2024). Ekosistem kreator Indonesia tumbuh secara substansial, dengan lebih dari 2.200 kanal YouTube Indonesia yang telah melampaui satu juta subscriber hingga akhir tahun 2023, meningkat 34 persen dibandingkan tahun 2021 (Social Blade LLC, 2023). Pertumbuhan ini menempatkan komunitas kreator konten Indonesia sebagai salah satu yang paling dinamis di Asia.

YouTube Partner Program (YPP) merupakan program resmi YouTube yang memungkinkan kreator konten menghasilkan pendapatan melalui berbagai mekanisme, termasuk iklan (AdSense), Channel Membership, Super Chat, dan Super Thanks. Untuk bergabung dengan YPP, kreator harus memenuhi ambang batas minimum 1.000 subscriber dan 4.000 jam watch time dalam 12 bulan terakhir. Kebijakan monetisasi YouTube mengalami perubahan signifikan secara berturut-turut dalam periode 2022 hingga 2025, termasuk diperkenalkannya monetisasi YouTube Shorts pada 2022, pembaruan jalur akselerasi YPP pada Juli 2023, dan pembatasan monetisasi konten berbantuan kecerdasan buatan pada Juli 2025 (Google LLC, 2025). Setiap perubahan kebijakan ini memicu respons yang masif dari komunitas kreator Indonesia yang tercermin dalam kolom komentar video YouTube.

Respons komunitas kreator terhadap kebijakan tersebut kerap mengandung opini multidimensional dalam satu komentar. Misalnya, seorang kreator dapat mengekspresikan kekecewaan terhadap syarat kelayakan monetisasi, sekaligus menyampaikan apresiasi terhadap konten edukasi yang dibagikan sesama kreator, dalam satu komentar yang sama. Analisis sentimen konvensional yang hanya menghasilkan satu label sentimen per dokumen tidak mampu menangkap

Dipindai dengan CamScanner

perbedaan polaritas antar topik semacam ini, sehingga informasi mengenai aspek kebijakan mana yang sebenarnya memicu sentimen negatif menjadi kabur. Kelemahan inilah yang mendasari penggunaan Aspect-Based Sentiment Analysis (ABSA), yaitu pendekatan analisis sentimen yang mengidentifikasi dan mengklasifikasikan sentimen secara terpisah untuk setiap aspek atau topik spesifik yang dibahas dalam satu teks, sehingga menghasilkan profil opini yang lebih granular dan informatif dibandingkan analisis sentimen pada level dokumen Do et al. (2019) Zhang et al. (2022).

Dalam konteks penelitian ini, "aspek" merujuk pada dimensi spesifik dari kebijakan monetisasi YPP yang menjadi objek opini kreator, yang mencakup persyaratan dan kebijakan monetisasi, performa channel dan konten, engagement penonton, apresiasi terhadap konten edukasi, dukungan teknis dan akun, hingga dukungan serta semangat antar sesama komunitas kreator. Dengan mengidentifikasi sentimen pada masing-masing aspek tersebut, penelitian ini dapat memetakan aspek mana yang paling banyak memicu keluhan (sentimen negatif) dan aspek mana yang justru direspons positif oleh komunitas, sebuah informasi yang tidak dapat diperoleh melalui analisis sentimen konvensional berbasis dokumen.

Penelitian ini mengimplementasikan ABSA menggunakan dua model klasifikasi. Pertama, IndoBERT sebagai model utama varian BERT yang dilatih khusus menggunakan korpus empat miliar kata Bahasa Indonesia, sehingga mampu memahami konteks kalimat informal, slang, dan singkatan yang dominan dalam komentar YouTube berbahasa Indonesia (Wilie, et al., 2020). Kedua, Support Vector Machine (SVM) dengan representasi Bag-of-Words (CountVectorizer) sebagai model baseline pembanding algoritma machine learning konvensional yang telah terbukti efektif untuk klasifikasi teks sebelum era deep learning (Vapnik, 1995) Azzouzy et al (2024). Data penelitian berupa komentar YouTube yang dikumpulkan menggunakan YouTube Data API v3 dari video-video yang membahas kebijakan monetisasi YPP.

Penelitian ini mengisi tiga research gap yang ada: (1) belum ada studi ABSA pada opini kreator Indonesia terhadap kebijakan monetisasi YouTube; (2) tidak ada

Dipindai dengan CamScanner

studi yang membandingkan IndoBERT vs SVM pada domain ini; dan (3) tidak ada studi yang menggunakan data komentar YouTube dari periode kebijakan YPP terkini 2024–2026 (Bastian & Perdana, 2025) (Firdaus & Trisnawarman, 2025).

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah dalam penelitian ini adalah:

1. Bagaimana distribusi sentimen (positif, negatif, netral) pada komentar kreator Indonesia terhadap kebijakan monetisasi YouTube Partner Program?
2. Aspek-aspek apa saja yang teridentifikasi dalam komentar tersebut, dan bagaimana distribusi frekuensinya?
3. Bagaimana profil sentimen pada tiap aspek yang teridentifikasi, dan aspek mana yang memiliki sentimen negatif tertinggi?
4. Apakah model IndoBERT menghasilkan performa klasifikasi yang lebih tinggi dibandingkan SVM sebagai baseline pada tugas ABSA ini?

1.3 Batasan Masalah

Untuk memfokuskan penelitian dan memastikan kedalaman analisis, penelitian ini dibatasi pada:

1. Data yang digunakan adalah komentar publik dari video YouTube yang membahas kebijakan monetisasi YouTube Partner Program, dikumpulkan melalui YouTube Data API v3 dalam rentang waktu Agustus 2024 hingga April 2026.
2. Hanya komentar berbahasa Indonesia termasuk bahasa gaul, singkatan, dan campuran Indonesia-Inggris yang diikutsertakan dalam analisis sentimen.
3. Penelitian ini hanya membandingkan dua model klasifikasi: IndoBERT sebagai model utama dan SVM dengan Bag-of-Words (CountVectorizer) sebagai baseline. Arsitektur lain seperti BiLSTM, CNN, atau XLM-RoBERTa berada di luar cakupan penelitian.

Dipindai dengan CamScanner

4. Polaritas sentimen diklasifikasikan ke dalam tiga kelas: Positif, Negatif, dan Netral. Analisis emosi multidimensional atau deteksi sarkasme berada di luar cakupan penelitian.
5. Aspek yang dianalisis dibatasi pada enam kategori: (1) Dukungan Komunitas & Semangat, (2) Syarat & Kebijakan Monetisasi, (3) Apresiasi Konten Edukasi, (4) Channel & Performa Konten, (5) Engagement Penonton, dan (6) Bantuan Teknis & Akun.
6. Penelitian hanya mencakup platform YouTube dan tidak membahas opini dari platform media sosial lain seperti Twitter, Instagram, atau TikTok.

1.4 Tujuan penelitian

Tujuan dari penelitian ini adalah:

1. Mengidentifikasi distribusi sentimen publik pada komentar kreator Indonesia terkait kebijakan monetisasi YouTube Partner Program.
2. Mengekstraksi aspek-aspek spesifik kebijakan monetisasi YouTube yang dominan dalam komentar beserta distribusi frekuensinya.
3. Memetakan profil sentimen berbasis aspek, termasuk aspek dengan sentimen negatif tertinggi.
4. Membandingkan performa model IndoBERT dan SVM pada tugas ABSA komentar berbahasa Indonesia terkait kebijakan monetisasi YouTube.

1.5 Manfaat penelitian

1. Manfaat Teoritis
    a. Kontribusi Akademik: Penelitian ini berkontribusi pada pengembangan literatur tentang Aspect-Based Sentiment Analysis (ABSA) pada teks informal Bahasa Indonesia, khususnya dalam konteks opini komunitas kreator konten digital terhadap kebijakan platform.
    b. Pengembangan Metodologi: Penelitian ini memberikan kontribusi metodologis berupa pipeline ABSA lengkap pengumpulan data, preprocessing, pelabelan otomatis, klasifikasi aspek, hingga

Dipindai dengan CamScanner

perbandingan model yang dapat diadaptasi untuk penelitian analisis sentimen berbahasa Indonesia pada domain lain.

c. Referensi Ilmiah: Temuan aspek-aspek diskusi komunitas kreator YouTube Indonesia dapat dijadikan referensi bagi penelitian selanjutnya yang mengkaji ekosistem kreator konten digital di Indonesia.

2. Manfaat Praktis

a. Bagi YouTube Indonesia: Temuan profil sentimen per aspek dapat menjadi masukan berbasis data untuk penyempurnaan kebijakan YPP dan peningkatan layanan dukungan teknis bagi kreator Indonesia.

b. Bagi Kreator Konten YouTube: Pemahaman tentang aspek-aspek yang mendominasi diskusi komunitas dapat membantu kreator dalam menyusun strategi konten yang lebih efektif.

c. Bagi Akademisi dan Peneliti: Pipeline dan model yang dikembangkan dalam penelitian ini dapat dijadikan kerangka referensi untuk membangun sistem monitoring sentimen otomatis pada platform digital lain di Indonesia.

d. Bagi Regulator dan Pemangku Kebijakan: Penelitian ini memberikan evidence-based insights mengenai respons komunitas kreator terhadap kebijakan platform digital.

1.6 Sistematika Penulisan

Penulisan penelitian ini disusun dalam lima bab yang saling berkaitan. Bab I Pendahuluan menguraikan latar belakang, rumusan masalah, batasan masalah, tujuan, manfaat, dan penelitian terkait. Bab II Tinjauan Pustaka membahas teori-teori landasan: sentiment analysis, ABSA, BERT, IndoBERT, SVM, dan YouTube Partner Program. Bab III Metodologi Penelitian menjelaskan seluruh tahapan penelitian meliputi pengumpulan data, preprocessing, pelabelan, klasifikasi aspek, pengembangan model, serta metrik evaluasi. Bab IV Hasil dan Pembahasan

Dipindai dengan CamScanner

menyajikan dan menganalisis hasil tiap tahapan penelitian. Bab V Kesimpulan dan Saran merangkum temuan utama dan memberikan rekomendasi.

1.7 Penelitian Terkait

Tabel 1.1 Penelitian Terkait

Penulis dan Tahun	Judul	Hasil / Pencapaian
Wilie et al. (2020)	IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding	Menghadirkan IndoBERT dan 12 benchmark NLP bahasa Indonesia; IndoBERT unggul atas mBERT pada semua tugas yang diuji
Do et al. (2019)	Deep Learning for Aspect-Based Sentiment Analysis: A Comparative Review	BERT-based models mengungguli LSTM pada 8 dari 10 dataset benchmark ABSA; menjadi acuan utama metodologi ABSA.
Zhang et al. (2022)	A Survey on Aspect-Based Sentiment Analysis: Tasks, Methods, and Challenges	Pemetaan sistematis subtask ABSA; gap pada domain spesifik dan bahasa rendah-sumber teridentifikasi sebagai prioritas riset
Ma'aly et al. (2024)	Exploring Sentiment Analysis for the Indonesian Presidential Election Through Online Reviews Using	Deep learning mencapai akurasi tinggi pada 8 kelas emosi komentar YouTube; menggunakan sentimen

Dipindai dengan CamScanner

	Multi-Label Classification	level dokumen tanpa dekomposisi aspek.
(Bastian & Perdana, 2025)	Sentiment Analysis of Public Comments on YouTube Regarding the Inaugural Speech of the President of Indonesia Using VADER and BERT	Hybrid VADER+BERT efektif menangani teks informal Indonesia; akurasi 87,4% namun tanpa fine-tuning domain spesifik
Firdaus & Trisnawarman (2025)	Analisis Sentimen Publik Terhadap Program Tabungan Perumahan Rakyat Menggunakan Model IndoBERT Lite pada Komentar YouTube	IndoBERT Lite mencapai akurasi 78% pada 14.618 komentar YouTube berbahasa Indonesia; efektif untuk analisis opini kebijakan publik.
Sumartha (2025)	Implementasi IndoBERT untuk Analisis Sentimen Opini Publik Terhadap Kebijakan Kenaikan UKT	IndoBERT efektif mengklasifikasikan sentimen opini publik terhadap kebijakan pemerintah; mayoritas opini negatif teridentifikasi
Anugrah & Agatha (2025)	Utilizing IndoBERT and BERTopic to Explore Public Opinion on BPS Instagram Posts	IndoBERT efektif mengekstrak opini publik terhadap kebijakan pemerintah di media sosial; potensi penerapan pada

Dipindai dengan CamScanner

		domain kebijakan platform digital.
Azzouzy et al. (2024)	Transformer-based Models for Sentiment Analysis of YouTube Video Comments	BERT mencapai F1-score 88% pada komentar YouTube; model transformer secara konsisten ungguli ML tradisional termasuk SVM.
Murfi et al. (2022)	BERT-based Combination of Convolutional and Recurrent Neural Network for Indonesian Sentiment Analysis	Kombinasi BERT+CNN+RNN mencapai akurasi 91,3% pada 9 topik teks Indonesia; unggul atas model BERT tunggal.

Dipindai dengan CamScanner