
Sistem Rekomendasi Influencer Multiobjektif dan Analisis Konten serta Perilaku Audiens Berbasis NLP

Azril Fahmiardi

Universitas Gadjah Mada

azrilfahmiardi@mail.ugm.ac.id

Fahmi Haqqul Ihsan

Universitas Gadjah Mada

fahmihacquihsan@mail.ugm.ac.id

Ni Ayu Putu Gayatri Maheswari

Universitas Gadjah Mada

niayuputugayatrimaheswari@mail.ugm.ac.id

Abstract

Dalam era pemasaran digital yang semakin kompetitif, pemilihan *influencer* yang tepat menjadi krusial untuk keberhasilan kampanye sebuah brand. Penelitian ini mengusulkan sebuah sistem rekomendasi influencer berbasis data yang mampu mengotomatisasi proses seleksi dengan mempertimbangkan berbagai faktor penting, seperti kesesuaian persona, kecocokan psiko-demografi, potensi performa kampanye, dan efisiensi anggaran. Sistem yang dikembangkan mengintegrasikan pendekatan *state-of-the-art* dalam bidang Natural Language Processing (NLP) seperti *Sentence-BERT*, *fine-tuned BERT* untuk analisis caption dan komentar, serta optimasi *Mixed Integer Programming* untuk perhitungan konten optimal. Selain itu, sistem juga dirancang untuk mendukung fleksibilitas strategi kampanye melalui skema prioritas yang dapat disesuaikan brand, seperti awareness, targeted conversion, atau efisiensi biaya. Evaluasi dilakukan menggunakan *ablation analysis* terhadap 16 brand dari berbagai industri, menunjukkan kontribusi masing-masing komponen skor bervariasi tergantung pada konteks kampanye. Hasil eksperimen menunjukkan bahwa sistem mampu memberikan rekomendasi yang lebih relevan dan terukur dibandingkan baseline konvensional.

1 Pendahuluan

1.1 Latar Belakang

Tingkat kompetitif yang semakin tinggi dalam dunia bisnis membuat konsumen dihadapkan berbagai pilihan produk [1]. Tercatat bahwa jumlah brand lokal yang terdaftar dalam *e-commerce* di Indonesia pada tahun 2023 mencapai 3.816.750, dengan pertumbuhan mencapai 27.4% setiap tahunnya [2]. Banyaknya pilihan yang ada membuat konsumen mempertimbangkan 4 hingga 5 brand yang berbeda untuk kategori *blue ocean* dan lebih dari 5 brand untuk kategori *red ocean*, seperti pada produk *fashion* dan *skincare* sebelum melakukan pembelian [3]. Di tengah banyaknya pilihan yang tersedia, brand perlu memiliki strategi marketing yang mampu menarik perhatian sekaligus membangun kepercayaan konsumen secara efektif.

Dengan jumlah pengguna aktif mencapai sekitar 143 juta, Indonesia menempati posisi pertama di Asia Tenggara sebagai pasar media sosial terbesar [4]. Media sosial menjadi salah satu saluran komunikasi paling aktif di Indonesia yang berperan penting dalam membentuk preferensi konsumen. Kondisi ini membuka peluang luas bagi brand untuk mengembangkan strategi marketing yang lebih

inovatif, salah satunya melalui penerapan strategi *influencer* marketing, yaitu dengan memanfaatkan *influencer* atau *Key Opinion Leader* (KOL) sebagai representasi brand melalui bentuk endorsement. Di Indonesia, *influencer* memiliki peranan yang signifikan dalam memengaruhi keputusan pembelian, sebanyak 68% konsumen pernah membeli produk yang direkomendasikan oleh *influencer*, 63% secara aktif mengikuti akun *influencer* untuk mendapatkan tips atau ulasan sebelum membeli, dan 62% menyatakan bahwa mereka membeli suatu produk karena diulas oleh *influencer* [4]. Di antara berbagai platform media sosial, Instagram menjadi saluran yang paling banyak digunakan untuk aktivitas pemasaran. Hal ini ditunjukkan oleh tingkat penggunaannya yang mencapai 84,8%, jauh lebih tinggi dibandingkan TikTok yang hanya sebesar 63,1% [4]. Sebanyak 69,9% brand juga menjalankan kampanye *influencer* melalui Instagram. Selain itu, 63% konsumen juga mengaku lebih memilih mengikuti akun *influencer* dibandingkan akun resmi milik brand [4].

Namun, dengan semakin banyaknya jumlah *influencer* dan kompleksitas preferensi audiens, brand menghadapi tantangan dalam memilih *influencer* yang paling relevan dan efektif. 1.200 dari 4.500 brand dalam survei yang dilakukan pada tahun 2023 mengaku mengalami kesulitan dalam memilih *influencer* yang cocok dengan image yang ditampilkan brand dan memiliki audiens serupa [5]. Proses mencari dan menyeleksi *influencer* juga sangat memakan waktu, umumnya brand yang melakukan prosesnya secara manual menghabiskan 15-25 jam kerja, atau sekitar 3-5 hari untuk satu *influencer* [6]. Dalam skenario terburuk, pemilihan *influencer* yang tidak tepat (miss-match) berpotensi menimbulkan sentimen negatif dan dapat merusak citra baik brand di mata konsumen [7] [8]. Misalnya, sentimen negatif terhadap Kiky Saputri saat bekerja sama dengan brand Somethinc karena karakternya yang dianggap masyarakat menggiring opini terhadap salah satu pasangan calon yang sedang ramai isunya, akibatnya Somethinc mengalami kerugian reputasi [9].

Dalam penelitian ini, penulis mengusulkan pendekatan berbasis *Artificial Intelligence* (AI) untuk mengotomatisasi proses pencocokan brand-influencer melalui pengintegrasian filtering anggaran, pencocokan persona dan audiens, serta evaluasi performa berbasis konten, sehingga keputusan kolaborasi dapat menjadi lebih presisi, efisien, dan berbasis data.

1.2 Analisis Praktik dan Kebutuhan Brand

Wawancara dilakukan secara langsung pada tanggal 29 Juni 2025 dengan dua praktisi industri guna memperoleh pemahaman mendalam mengenai tantangan dalam strategi *influencer marketing*.

Andika Dwijatmiko, selaku *Branding Corporate Consultant*, memiliki pengalaman dalam menangani berbagai kebutuhan brand dalam pengembangan identitas citra. Berdasarkan diskusi yang dilakukan, diperoleh sejumlah wawasan mengenai tantangan yang saat ini dihadapi oleh brand:

1. Tantangan utama brand bukan lagi sekadar membangun kepercayaan di mata konsumen, tetapi juga menciptakan identitas yang berkelanjutan dan melekat dalam jangka panjang.
2. Banyak brand kesulitan menyampaikan pesan kampanye secara efektif karena kurangnya keterlibatan emosional maupun partisipatif dari konsumen, serta absennya elemen *call-to-action* dalam penyampaian pesan.
3. Kesalahan yang umum terjadi dalam strategi endorsement adalah pemilihan *influencer* yang hanya mengandalkan tingkat popularitas, tanpa mempertimbangkan kecocokan audiens antara *influencer* dan target pasar brand, sehingga menurunkan relevansi kampanye.

Ghatsa Zahira Shofa, selaku *Brand Marketing* di salah satu perusahaan FMCG (Fast Moving Consumer Goods) di Indonesia yang terlibat dalam pengelolaan lebih dari 130 kolaborasi strategis dengan *influencer*, mencakup tahapan menyeluruh dari riset pasar awal hingga perumusan *brief marketing campaign*. Berdasarkan wawancara yang dilakukan, diperoleh insight terkait pertimbangan brand dalam memilih *influencer*:

1. Kesesuaian persona *influencer* dengan citra dan nilai brand.
2. Kecocokan harga dengan alokasi anggaran *marketing campaign*.
3. Potensi konversi yang diharapkan dari kolaborasi.

1.3 Analisis Kompetitor

Sebagai upaya mempercepat proses pemilihan *influencer*, banyak brand menengah hingga besar kini mulai memanfaatkan platform seperti KOL.ID, Allstars, ICE, hingga agensi marketing. Namun, pendekatan platform-platform tersebut umumnya masih menjangkau pada metrik sederhana seperti

reach, *rate card*, dan *engagement rate*. Padahal, strategi marketing yang efektif bukan hanya soal angka, tapi tentang bagaimana pesan brand bisa match dengan karakter *influencer* dan tersampaikan secara autentik kepada audiens [10]. Celah inilah yang menjadi pijakan utama dari solusi yang kami kembangkan untuk membuat sistem pemilihan influencer yang tidak hanya berbasis data, tetapi juga memahami konteks komunikasi dan persona. Analisis lengkap terhadap lanskap kompetitor dapat dilihat lampiran [Lampiran 1].

1.4 Value Proposition

Solusi yang dikembangkan dalam paper ini bertujuan untuk menjawab kekosongan tersebut dengan pendekatan berbasis AI yang komprehensif. Model menawarkan pendekatan yang mampu menilai kecocokan *influencer* dengan brand secara lebih menyeluruh. Sistem tidak hanya memperhitungkan kesesuaian persona dan nilai brand, tetapi juga memberikan insight analisa caption, komentar audiens, serta potensi performa kampanye berdasarkan *call-to-action* dan tren sentimen publik. Analisis *value proposition* secara lengkap dapat dilihat di lampiran [lampiran 2].

2 Kajian Teori

2.1 Kajian Teoritis Bisnis

Untuk memastikan solusi yang dikembangkan tidak hanya berbasis teknis, tetapi juga memiliki fondasi konseptual yang kuat, kajian ini merujuk pada sejumlah teori dalam marketing. Teori-teori berikut memungkinkan pendekatan relevan dalam memahami dinamika perilaku konsumen serta praktik *influencer endorsement* dalam industri bisnis.

1. **Match-Up Fit Theory** teori ini menyatakan bahwa kesesuaian antara *influencer* dan produk akan menciptakan citra yang lebih positif, yang berujung pada kepercayaan yang lebih besar terhadap iklan tersebut. Teori ini mengklasifikasikan aspek penting dalam penilaian terhadap *influencer*, yaitu berdasarkan *daya tarik* (attractiveness) dan *kredibilitas* (credibility), sementara untuk brand terdiri dari *brand awareness* dan *brand association* [11].
2. **Segmenting, Targeting, Positioning (STP)** teori ini digunakan untuk memahami audiens dan kemudian mencocokkan segmen pasar yang relevan untuk produk. Dalam model, STP digunakan untuk mengidentifikasi kesamaan audiens antara *influencer* dan brand, guna menciptakan strategi yang lebih terfokus dan efektif [12].
3. **Hierarchy of Effects Model** teori ini memetakan tujuan penggunaan *influencer* dalam mencapai berbagai tahap pengaruh konsumen. Tahapan tersebut dibagi menjadi tiga, yaitu *cognitive* (meningkatkan kesadaran dan pengetahuan), *conative* (membangun hubungan emosional dan preferensi), serta *affective* (mendorong tindakan). Dalam model, teori digunakan untuk merencanakan strategi *influencer* dengan memahami kebutuhan brand [13].
4. **Marketing Funnel** teori ini memetakan proses pengambilan keputusan konsumen dari tahap pertama mengenal produk hingga menentukan pilihan produk. Dalam model, teori digunakan sebagai tahap lanjutan dalam memahami kebutuhan brand dengan mendalami perilaku konsumen dan mengaitkannya dengan proses marketing yang dilakukan brand. Proses terdiri dari menarik perhatian (*attention*), membangun ketertarikan (*interest*), membangkitkan keinginan (*desire*), hingga mendorong tindakan (*action*) [14].

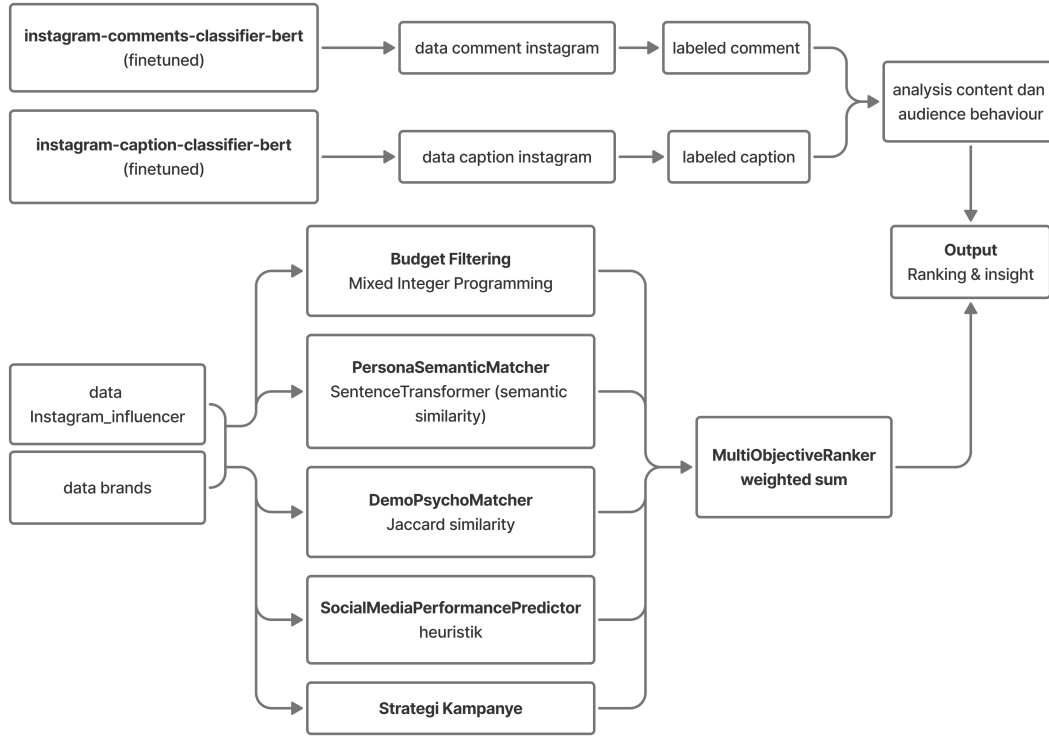
2.2 Kerangka Metodologis Model

1. **Sentence-BERT dan Cosine Similarity**, untuk mengukur kesamaan semantik antara narasi brand dan *influencer* dalam bentuk representasi vektor.
2. **BERT Fine-Tuning**, digunakan dalam klasifikasi gaya caption dan respons audiens berbasis teks.
3. **Mixed Integer Programming (MIP)**, sebagai pendekatan optimasi kombinasi konten kampanye dalam batasan anggaran.
4. **Weighted Jaccard Similarity**, digunakan untuk mencocokkan profil demografis dan psikografis audiens dengan target pasar brand.
5. **Heuristik Penilaian Performa**, untuk menghasilkan skor performa influencer berdasarkan empat dimensi utama: *engagement*, *authenticity*, *reach*, dan *brand fit*.

3 Solusi Usulan

Dalam penelitian ini diusulkan sebuah sistem rekomendasi *influencer* berbasis data yang mampu memberikan rekomendasi *influencer* sesuai kebutuhan brand secara otomatis dan terukur. Seluruh komponen utama sistem ini dirancang dengan mengadopsi pendekatan state-of-the-art (SOTA) di bidang Natural Language Processing (NLP) dan optimasi. Kami menggunakan Sentence-BERT untuk pencocokan semantik persona, BERT yang di-*fine-tune* secara khusus untuk klasifikasi caption dan komentar Instagram, serta Mixed Integer Programming untuk optimasi alokasi anggaran.

Sistem ini juga dirancang untuk mendukung fleksibilitas strategi kampanye melalui skema pembobotan dinamis (*brand priorities*) yang dapat dipilih oleh brand sesuai dengan tujuan kampanye mereka, seperti *awareness*, efisiensi biaya, atau konversi tersegmentasi. Pemilihan model-model ini didasarkan pada performa terbaik di berbagai benchmark NLP, sehingga memastikan sistem yang dikembangkan memiliki akurasi dan relevansi tinggi dalam aplikasi nyata. Sistem ini terdiri dari beberapa komponen utama yang saling terintegrasi, yaitu:



Gambar 1: Diagram alur sistem rekomendasi influencer yang diusulkan.

3.1 Budget Filtering

Budget filtering berperan dalam menyaring *influencer* berdasarkan batas anggaran dari setiap brand, kemudian menghitung kombinasi konten yang paling optimal (*story*, *feeds*, *reels*) untuk dimasukkan dalam kampanye. Proses ini menggunakan pendekatan Mixed Integer Programming (MIP) yang mempertimbangkan efektivitas biaya, dampak kampanye, dan keberagaman jenis konten agar strategi yang dihasilkan tetap realistis dan praktis secara operasional. Filter ini menggunakan tiga variabel keputusan: x , y , dan z yang masing-masing merepresentasikan jumlah *story*, *feeds*, dan *reels* yang dipilih untuk satu *influencer*. Fungsi objektif dari model dituliskan sebagai:

$$\text{Maximize: } x \cdot I_s + y \cdot I_f + z \cdot I_r + \delta \cdot (u_s + u_f + u_r) \quad (1)$$

$$\text{Subject to: } x \cdot R_s + y \cdot R_f + z \cdot R_r \leq \text{AnggaranBrand} \quad (2)$$

Di mana I_s, I_f, I_r merupakan estimasi skor dampak untuk *story, feeds*, dan *reels*, yang dihitung berdasarkan *engagement rate, tier influencer*, dan status tren. Variabel biner $u_s, u_f, u_r \in \{0, 1\}$ digunakan untuk menandai apakah jenis konten tersebut diaktifkan, sedangkan δ adalah parameter *diversity bonus* yang diberikan apabila lebih dari satu jenis konten digunakan secara bersamaan dalam kampanye. Sementara itu, R_s, R_f, R_r pada rumus 2 adalah *rate card* atau biaya masing-masing konten (*story, feeds*, dan *reels*). Pendekatan ini memungkinkan sistem untuk merekomendasikan strategi kampanye yang tidak hanya efisien secara biaya, tetapi juga bervariasi secara format, sehingga mampu meningkatkan eksposur audiens dan daya tarik kampanye secara keseluruhan.

3.2 Persona Matcher

Bertugas mengevaluasi sejauh mana narasi dan gaya komunikasi influencer sesuai dengan ekspektasi dan karakter brand. Modul ini mengadopsi pendekatan semantik menggunakan *Sentence-BERT (SBERT)* untuk mengkonversi teks menjadi representasi vektor (*embedding*) bermakna, salah satu model *state-of-the-art (SOTA)* untuk representasi kalimat, sehingga mampu mengenali kesamaan naratif yang lebih dalam dibanding metode konvensional. Sistem mengubah atribut *brand criteria* menjadi vektor representatif yang mencerminkan kepribadian merek, termasuk gaya bahasa, tone, dan afinitas emosional. Di sisi lain, setiap *influencer* memiliki vektor yang dihasilkan dari gabungan *bio* serta seluruh *caption* historis yang tersedia di dataset. Proses pencocokan dilakukan dengan menghitung cosine similarity antar kedua vektor tersebut.

$$PersonaFitScore = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} \quad (3)$$

Pada rumus 3, A adalah *embedding* representatif dari brand dan B adalah *embedding* narasi *influencer*. Skor hasil pencocokan berada dalam rentang 0–1. Semakin tinggi skor, semakin kuat kesamaan antara karakter komunikasi *influencer* dengan arah kreatif brand. Berbeda dari pencocokan berbasis kategori eksplisit seperti demografi, pendekatan semantik mampu mengenali kesamaan naratif yang lebih dalam dan tidak selalu eksplisit, seperti tone kasual, gaya bercerita, atau humor. Pendekatan ini penting terutama untuk campaign yang mengutamakan aspek *self-expression*, seperti campaign dengan target Gen Z yang sangat peka terhadap autentisitas dan nilai personal.

3.3 Psycho-Demografi Matcher

Komponen pencocokan psiko-demografi berperan penting dalam memastikan bahwa audiens dari seorang influencer memiliki kesamaan karakteristik dengan target pasar brand. Analisis dilakukan terhadap dua aspek utama, yaitu atribut demografis (usia, gender, dan pendapatan) serta atribut psikografis (gaya hidup, kebiasaan, dan nilai—baik nilai brand maupun nilai utama influencer).

Setiap brand dan influencer mendefinisikan distribusi audiensnya dalam format array list atribut. Dengan format seragam ini, sistem dapat membandingkan kedua profil secara langsung. Untuk mengukur kesesuaian, digunakan pendekatan weighted Jaccard similarity, yaitu versi berbobot dari kesamaan Jaccard yang mempertimbangkan distribusi proporsional dari setiap kategori.

$$JaccardSimilarity = \frac{|A \cap B|}{|A \cup B|} \quad (4)$$

di mana A adalah himpunan gabungan atribut demografi dan psikografi dari audiens *influencer*, dan B adalah himpunan atribut dari target audiens brand. Penyebut $|A \cup B|$ mencakup seluruh atribut unik dari kedua pihak, sedangkan pembilang $|A \cap B|$ merepresentasikan atribut yang dimiliki bersama. Nilai akhir dari versi berbobot ini mempertimbangkan frekuensi relatif kemunculan kategori, sehingga dapat menangani kasus di mana satu atribut memiliki beberapa nilai sekaligus (misalnya: ["Generasi Z", "Milenial"]).

3.4 Social Media Performance Potential

Untuk mengestimasi potensi kinerja kampanye secara kuantitatif, penelitian ini mengembangkan *Social Media Performance Predictor*, sebuah modul berbasis heuristik yang mengevaluasi setiap influencer dari beberapa dimensi. Modul ini menghasilkan skor komposit yang mencerminkan kontribusi influencer terhadap keberhasilan kampanye dan digunakan dalam sistem pemeringkatan akhir.

Komponen Penilaian Prediksi performa kampanye dihitung berdasarkan empat aspek utama: **Engagement**, **Authenticity**, **Reach**, dan **Brand Fit**, dengan bobot sebagai berikut:

$$PerformanceScore = 0.3 \times Engagement + 0.25 \times Authenticity + 0.25 \times Reach + 0.2 \times BrandFit \quad (5)$$

Bobot ditentukan berdasarkan riset primer dan sekunder yang menyoroti aspek prioritas brand. *Engagement* menilai keterlibatan audiens yang berpengaruh besar dalam distribusi konten. *Authenticity* merefleksikan kepercayaan terhadap brand, *Reach* digunakan untuk kampanye berbasis awareness, dan *Brand Fit* membantu menjaga keselarasan identitas [15] [16].

Engagement Score Skor ini mengukur interaksi audiens terhadap konten melalui rumus berikut:

$$EngagementScore = \min \left(\frac{ER}{BenchmarkER}, 2.0 \right) \div 2.0 \quad (6)$$

dengan *ER* adalah actual engagement rate dari influencer, dan *BenchmarkER* merupakan rata-rata berdasarkan tier pengikut.

Authenticity Score Skor ini menilai keaslian konten dengan mempertimbangkan rasio konten berbayar dan konsistensi perilaku, dihitung sebagai:

$$AuthenticityScore = \min ((1 - endorse_rate) \cdot 0.7 + c \cdot 0.3, 1.0) \quad (7)$$

di mana *endorse_rate* adalah proporsi konten sponsor, dan *c* bernilai 1 jika perilaku konsisten, 0 jika tidak.

Reach Potential Potensi jangkauan dihitung dari rata-rata *views* dan status tren konten:

$$ReachPotential = \min \left(\left(\min \left(\frac{avg_views}{BenchmarkViews}, 2.0 \right) \div 2.0 \right) \cdot t, 1.0 \right) \quad (8)$$

dengan *avg_views* adalah rata-rata *views*, *BenchmarkViews* adalah nilai acuan, dan *t* adalah faktor tren (1.2 jika trending, 1.0 jika tidak).

Brand Fit Kesesuaian antara influencer dan brand dihitung menggunakan representasi semantik teks dari model Sentence Transformer:

$$BrandFit = \cos(\vec{v}_{influencer}, \vec{v}_{brand}) \quad (9)$$

di mana $\vec{v}_{influencer}$ adalah vektor embedding deskripsi influencer, dan \vec{v}_{brand} adalah representasi industri brand.

3.5 Penyesuaian Strategi Kampanye melalui Skema Prioritas

Sistem ini mendukung penyesuaian rekomendasi melalui parameter *brand priorities*, yaitu skema bobot yang menentukan kontribusi masing-masing komponen skor dalam proses pemeringkatan influencer. Terdapat empat komponen utama yang dipertimbangkan: **persona_fit**, **demo_fit**, **performance_pred**, dan **budget_efficiency**. Pemilihan bobot bergantung pada jenis kampanye yang dijalankan oleh brand. Untuk itu, sistem menyediakan beberapa skema preset yang dapat dipilih sesuai kebutuhan brand. Tabel 1 merangkum lima skema yang diimplementasikan dalam sistem.

Skema Strategi	persona_fit	demo_fit	performance_pred	budget_efficiency
Balanced (default)	0.25	0.25	0.25	0.25
Targeted Campaign	0.10	0.45	0.35	0.10
Awareness Campaign	0.30	0.20	0.40	0.10
Cost Efficiency	0.10	0.20	0.30	0.40
Persona-Centric Brand	0.40	0.25	0.25	0.10

Tabel 1: Preset Skema Prioritas dalam Sistem Rekomendasi Influencer

Dengan menyediakan opsi preset, sistem memungkinkan brand untuk menyesuaikan strategi rekomendasi tanpa harus memahami kompleksitas teknis di balik algoritma pemeringkatan. Selain itu, pendekatan ini juga membuka peluang eksplorasi strategi preferensi industri yang lebih luas.

3.6 Analisis Konten dan Audiens dengan BERT Fine-Tuned

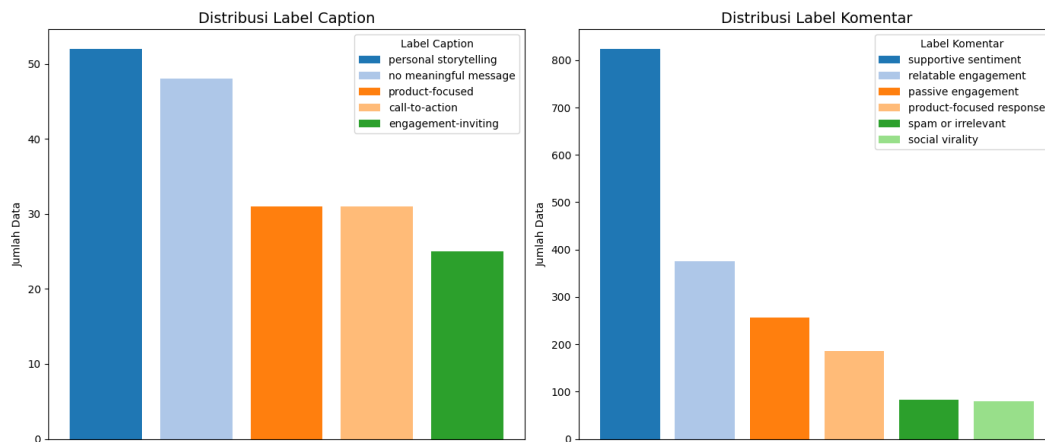
Salah satu keunggulan utama sistem kami adalah integrasi dua model BERT yang telah di-finetune secara terpisah untuk klasifikasi caption dan komentar Instagram. Model BERT dilatih menggunakan dataset hasil labeling manual, sehingga mampu mengenali berbagai gaya komunikasi influencer dan tipe interaksi audiens.

1. Model klasifikasi caption untuk mendeteksi gaya komunikasi dengan label: *call-to-action*, *engagement-inviting*, *no meaningful message*, *personal storytelling*, dan *product-focused*.
2. Model klasifikasi komentar audiens untuk mengelompokkan tipe interaksi menjadi: *passive engagement*, *product-focused response*, *relatable engagement*, *social virality*, *spam or irrelevant*, dan *supportive sentiment*.

Model dilatih menggunakan arsitektur *Bert For Sequence Classification* dari pustaka *Hugging Face Transformers*. Proses fine-tuning dilakukan pada GPU menggunakan Google Colab. Output dari model digunakan untuk menghitung:

1. Memberikan insight mendalam terkait kualitas dan relevansi interaksi audiens pada setiap influencer.
2. Mengidentifikasi gaya caption yang paling sesuai dengan tujuan kampanye brand (*awareness*, *engagement*, *conversion*).
3. Menyajikan metrik seperti *high-value comment rate*, distribusi label caption, dan contoh konkret hasil klasifikasi pada output rekomendasi.

Dataset yang digunakan untuk pelatihan masing-masing model berasal dari subset data yang telah dilabeli secara manual. Dari total 705 caption 16.973 komentar yang tersedia, hanya 87 *caption* dan 1.803 komentar yang digunakan untuk pelatihan model, setelah melewati proses anotasi manual dan pembersihan data. Distribusi label dari masing-masing dataset ditampilkan pada Gambar 2. Integrasi analisis berbasis BERT memastikan rekomendasi influencer tidak hanya berdasarkan angka (*engagement*, *reach*), tetapi juga kualitas komunikasi dan potensi konversi yang lebih kontekstual. Pendekatan analisis konten dan audiens berbasis BERT fine-tuned belum diadopsi di Indonesia, sehingga menjadi nilai tambah utama dan keunggulan kompetitif solusi kami.



Gambar 2: Distribusi label untuk data caption dan komentar yang digunakan dalam fine-tuning BERT (n = 187 caption, n = 1803 komentar)

3.7 Alur Sistem

1. Brand dan *influencer* memiliki data atribut masing-masing, termasuk persona, demografi, serta metrik performa historis.

2. Brand menentukan kriteria kampanye, alokasi anggaran, serta memilih skema prioritas strategi (*brand priorities*) sesuai objektif kampanye.
3. Sistem menyaring influencer berdasarkan budget dan menghitung kombinasi konten optimal untuk masing-masing.
4. *Influencer* yang lolos dihitung skornya berdasarkan kecocokan persona, demografi-psikografi, dan prediksi performa, lalu digabungkan menggunakan bobot skema prioritas yang dipilih.
5. *Influencer* diurutkan berdasarkan skor akhir, dan sistem menampilkan ringkasan *insight* naratif berdasarkan hasil klasifikasi *caption* dan komentar menggunakan model BERT.

3.8 Contoh Output

Berikut ringkasan output sistem untuk brand **Good Day**. Pada contoh ini, brand memilih skema prioritas *Targeted Campaign*, di mana aspek demografi-psikografi dan potensi performa kampanye mendapat bobot lebih tinggi dalam pemeringkatan akhir. Seluruh *insight* dan metrik pada tabel serta tampilan *prototype* dihasilkan secara otomatis oleh sistem berdasarkan skema prioritas terpilih. Panduan dan tampilan *output* secara lengkap dapat dilihat di lampiran [Lampiran 4].

Influencer	Overall Score	Persona Fit	Perf. Potential	Dampak Est.	Strategi
@attahalilintar	81.6%	12.8%	64.3%	3.6 pts	1 Post
@raditya_dika	79.2%	28.9%	60.6%	3.4 pts	1 Post
@rizkyfbian	79.0%	32.8%	59.8%	6.4 pts	2 Posts

Tabel 2: Ringkasan rekomendasi top 3 influencer untuk Good Day dengan detail performa.

4 Hasil Eksperimen dan Pengujian

4.1 Deskripsi Dataset

Sistem diujikan menggunakan beberapa dataset utama yang dikumpulkan dan dikurasi secara manual maupun otomatis. Data *caption*, komentar, dan *bio* Instagram diperoleh melalui proses scraping langsung dari platform Instagram sebanyak lima postingan terakhir dari setiap influencer hingga tanggal 30 Juni 2025 sebagai batas waktu pengambilan data.

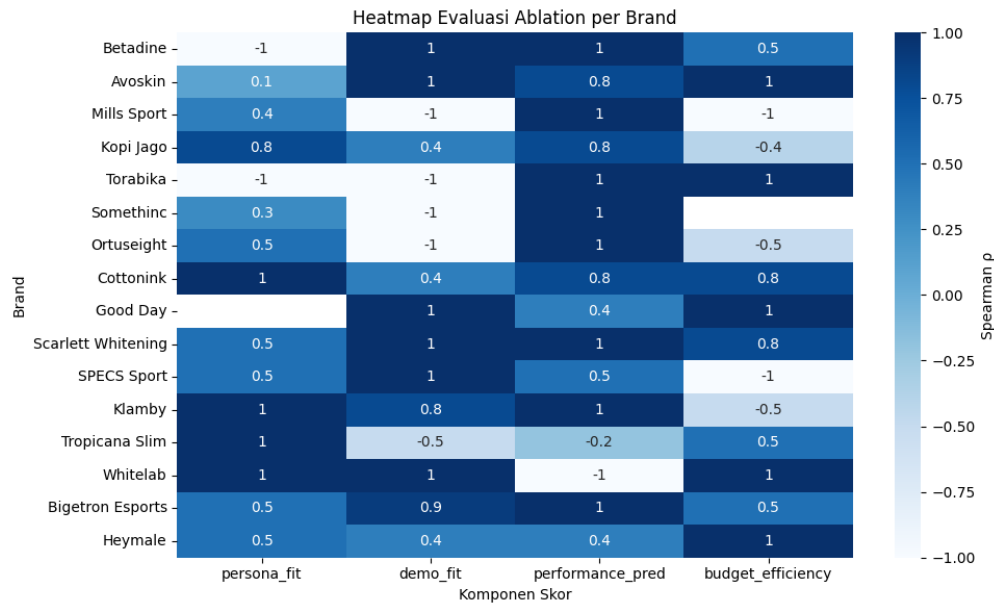
Secara keseluruhan, data yang berhasil dikumpulkan mencakup 153 **influencer Instagram**, 200 **brand**, 705 **caption**, dan 16.973 **komentar**. Untuk data brand dan influencer, nama-nama utama ditentukan secara manual dengan mempertimbangkan diversitas dunia nyata, sedangkan atribut dan data tambahan digenerasi menggunakan model GPT-4 dari OpenAI guna memungkinkan keluaran sistem memberikan *insight* terhadap berbagai skenario.

Sementara itu, untuk *caption* dan komentar, sebagian data dilabel secara manual dan digunakan sebagai data latih dan validasi untuk *fine-tuning* model BERT. Distribusi jumlah data per label dapat dilihat pada Gambar 2, dengan total 187 data *caption* dan 1.803 data komentar. Dataset hasil pelabelan tersebut kemudian dibagi menjadi data pelatihan dan validasi dengan rasio 80:20. Sisa data yang belum dilabel digunakan sebagai data uji tambahan.

4.2 Evaluasi Komponen Skoring (Ablation Analysis)

Untuk mengevaluasi kontribusi relatif tiap komponen skor dalam sistem rekomendasi, dilakukan pengujian *ablation analysis* menggunakan metode *leave-one-out*. Setiap komponen dihilangkan secara bergantian dari perhitungan skor akhir, dan dilakukan pemeringkatan ulang *influencer*. Korelasi Spearman (ρ) digunakan untuk mengukur seberapa jauh perubahan ranking dibanding model penuh.

Hasil evaluasi pada 16 brand menunjukkan bahwa **demo_fit** dan **performance_pred** adalah komponen paling berpengaruh: penghapusannya sering mengakibatkan pergeseran besar dalam ranking. Sebaliknya, **persona_fit** dan **budget_efficiency** menunjukkan pengaruh yang lebih bervariasi tergantung karakteristik brand. Gambar 3 memperlihatkan nilai korelasi Spearman antar kondisi, dengan nilai $\rho = 1$ mengindikasikan urutan identik terhadap model penuh, dan $\rho = 0$ menunjukkan peringkat berubah total secara acak.



Gambar 3: Heatmap nilai Spearman Rank Correlation (ρ) hasil ablation analysis pada 16 brand.

Secara umum, *demo_fit* mencatat rata-rata $\rho = 0.68$ dengan variasi rendah antar brand, dan *performance_pred* menunjukkan nilai $\rho = 0.84$ dengan stabilitas sedang. Sebaliknya, *persona_fit* memiliki rata-rata $\rho = 0.49$ dan variasi tinggi, sementara *budget_efficiency* mencatat $\rho = 0.54$ dan juga dipengaruhi oleh konteks brand. Temuan ini menunjukkan bahwa pentingnya komponen bersifat kontekstual, sehingga sistem memungkinkan penyesuaian *brand priorities* secara fleksibel.

5 Analisis Hasil

Conversion Potential

Analisis kualitas komentar audiens, termasuk proporsi high-value comment. Visualisasi distribusi label komentar dan high-value comment rate memberikan gambaran potensi konversi audiens influencer. Sebagai contoh, pada akun **@attahalilintar**, dari 192 komentar yang dianalisis, sistem menemukan 38.1% termasuk kategori high-value comment. Salah satu komentar bertipe *relatable engagement* adalah “*Pernah jadi A-TEAM pada masa nya , ternyata benar kata coki-muslim musuh besar A-TEAM adalah kedewasaan*”. Komentar ini menunjukkan adanya kedekatan emosional dan pengalaman bersama antara audiens dan influencer, di mana komunitas “A-TEAM” menjadi identitas bersama. Sistem mampu menangkap pola engagement yang relatable, bukan sekadar dukungan pasif. Pada label *social virality*, sistem mengidentifikasi komentar: “*Bang @attahalilintar berkehidupan di era skrg berkat para pejuang kemerdekaan ...*” yang memperlihatkan narasi viral dan potensi jangkauan pesan yang lebih luas.



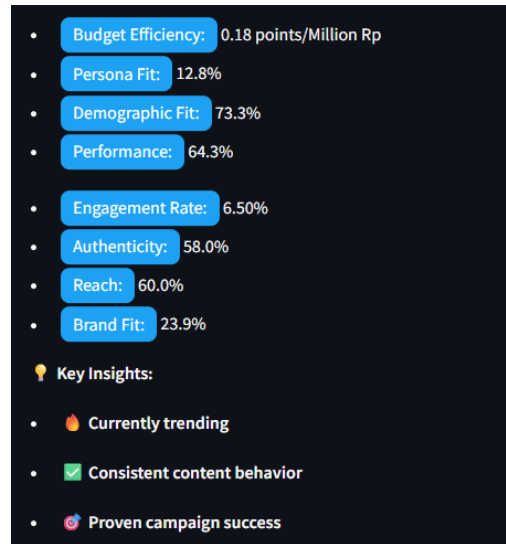
Gambar 4: Analisis komentar audiens dan high-value comment rate

Score Breakdown dan Key Insights

Rincian skor tiap aspek utama (budget efficiency, persona fit, demo fit, performance potential) sekaligus insight kualitatif seperti “currently trending”, “proven campaign success”, atau “low endorsement frequency” ditampilkan dalam satu section.

Sebagai ilustrasi, skor persona fit untuk influencer ini relatif rendah (12.8%), namun skor demographic fit (73.3%) dan performance potential (64.3%) cukup tinggi, didukung oleh engagement rate 6.5% dan authenticity score 58.0%.

Sistem juga menandai influencer ini sebagai “currently trending” dan memiliki “consistent content behavior”, insight penting bagi brand dalam memilih kolaborator yang relevan dan stabil.



Gambar 5: Rincian skor dan insight tambahan

Caption Behavior Summary

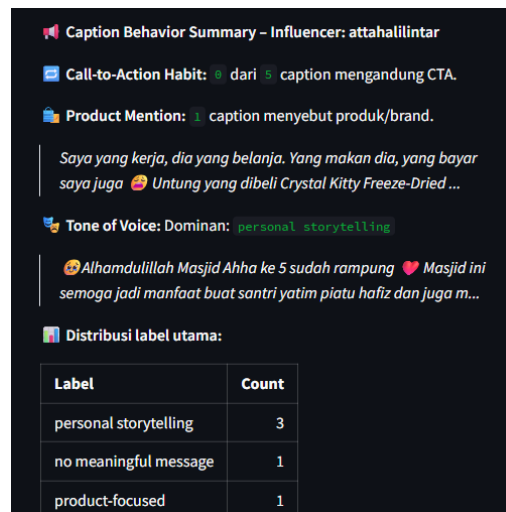
Menunjukkan gaya komunikasi influencer, seperti dominasi storytelling, call-to-action, atau product-focused.

Pada analisis caption, sistem mendeteksi bahwa dari 5 caption terakhir @attahalilintar, tidak ditemukan call-to-action (CTA), menandakan gaya komunikasi yang lebih naratif daripada ajakan eksplisit.

Tone of voice yang dominan adalah *personal storytelling*, seperti pada caption:

“Alhamdulillah Masjid Ahha ke 5 sudah rampung. Masjid ini semoga jadi manfaat buat santri yatim piatu hafiz dan juga m...”

Sistem mampu menangkap pola narasi yang personal dan inspiratif, sesuai dengan persona ceria dan ekspresif yang diharapkan brand.



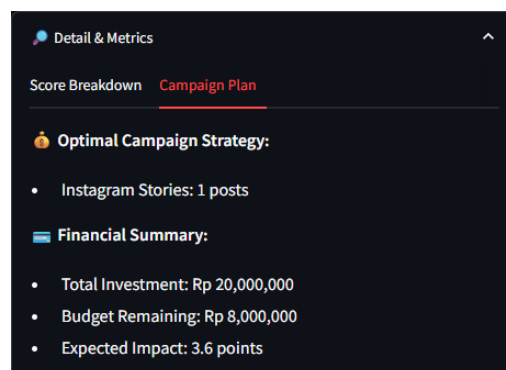
Gambar 6: Analisis gaya komunikasi caption influencer

Optimal Campaign Strategy dan Financial Summary

Rekomendasi strategi konten optimal (jumlah story, feed, reels) dan estimasi biaya divisualisasikan untuk membantu brand mengoptimalkan alokasi anggaran.

Sistem merekomendasikan, misalnya, 1 Instagram Story dengan total investasi Rp 20.000.000 dan sisa anggaran Rp 8.000.000, dengan estimasi dampak 3.6 points.

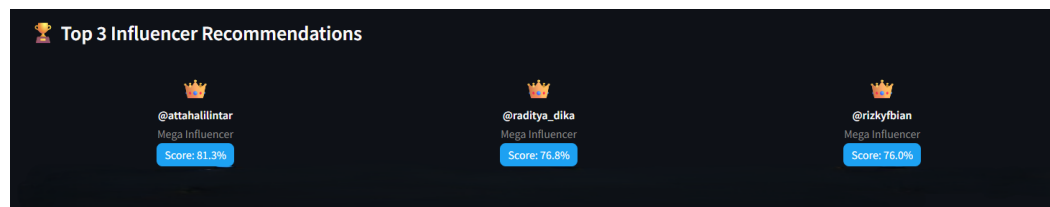
Dengan insight ini diharapkan brand dapat memaksimalkan hasil kampanye dalam batas anggaran yang tersedia.



Gambar 7: Strategi dan proyeksi anggaran kampanye

Overall Match Score

Skor akhir yang menggabungkan kecocokan persona, demografi-psikografi, performa, dan efisiensi anggaran. Visualisasi skor akhir dan ranking influencer membantu brand memilih kandidat terbaik secara objektif.



Gambar 8: Skor akhir dan ranking influencer

Dengan membedah tiap bagian output dan menampilkan visualisasi terkait, brand dapat mengambil keputusan lebih informasional dan strategis, tidak hanya berdasarkan angka, tetapi juga kualitas interaksi dan potensi konversi. Sistem ini memberikan insight mendalam yang tidak hanya kuantitatif, namun juga kualitatif, sehingga dapat meningkatkan efektivitas pemilihan influencer untuk kampanye brand.

6 Dampak Solusi

6.1 Dampak Non-finansial

Model ini memberikan dampak non-finansial bagi brand, terutama dalam menjaga citra brand, mengurangi risiko reputasi, dan mendukung prinsip Environmental, Social, and Governance (ESG). Sistem kami membantu brand untuk mengurangi risiko kesalahan pemilihan influencer yang dapat berujung pada penolakan publik atau sentimen negatif terhadap brand. Serta, menilai etika dan kepercayaan publik terhadap influencer melalui analisis gaya komunikasi dan komentar audiens, sehingga brand dapat menghindari kerja sama dengan figur yang memiliki tren negatif. Dampak ini penting, mengingat dalam ekosistem digital saat ini, kesalahan komunikasi dari satu kampanye dapat berimbas pada turunnya kepercayaan publik terhadap brand dalam jangka panjang.

6.2 Dampak Finansial

Dalam penelitian ini, pendekatan yang digunakan untuk mengetahui estimasi efisiensi biaya adalah Comparative Cost Analysis (CCA). CCA merupakan metode yang digunakan untuk membandingkan biaya dari pilihan strategi yang ada, yaitu (1) metode manual (in-house), (2) penggunaan agensi/platform, dan (3) sistem rekomendasi berbasis AI, dengan mempertimbangkan aspek finansial dan efektivitas yang relevan dalam pengambilan keputusan [17]. Setiap pendekatan dievaluasi berdasarkan estimasi biaya aktual terhadap tiga komponen utama:

1. **Biaya tenaga kerja pemasaran**, yaitu waktu dan biaya yang dibutuhkan oleh tim brand untuk melakukan proses seleksi.
2. **Mismatch cost**, yaitu kerugian akibat kesalahan pemilihan influencer yang tidak sesuai dengan citra brand.
3. **Opportunity cost**, yaitu potensi kerugian akibat keterlambatan kampanye yang seharusnya dapat dioptimalkan lebih cepat.

Perbandingan Biaya (Rp)			
Komponen Biaya	Solusi AI	Agensi/Platform	Manual
Biaya Marketing per Influencer	120.000	632.000	1.440.045
Biaya Mismatch	750.000	3.500.000	7.500.000
Opportunity Cost	250.000	1.500.000	3.500.000
Total Biaya	1.120.000	5.632.000	12.440.045

Tabel 3: Analisis Comparative Cost Analysis Pemilihan Influencer

Estimasi biaya pemasaran pada pendekatan manual dihitung dengan mengacu pada rata-rata gaji bulanan posisi Brand Marketing di Indonesia [18]. Nilai tersebut kemudian disesuaikan secara proporsional dengan estimasi durasi kerja 2–3 hari dalam proses seleksi influencer untuk satu kampanye, yang umumnya menghasilkan 1–2 profil akhir. Sementara, untuk pendekatan melalui agensi atau platform, perhitungan biaya mengacu pada rata-rata langganan bulanan yang diperoleh dari package price resmi KOL.ID [19], Allstars [20], ICE [21], dan Partipost [22].

Biaya mismatch merujuk pada rata-rata biaya content marketing [23], lalu risiko kesalahan dikategorikan menjadi tiga berdasarkan asumsi umum proses kerja [24] [25], yaitu:

1. Manual memiliki risiko tinggi karena mengandalkan subjektivitas dan proses manual tanpa verifikasi otomatis.
2. Agensi/platform memiliki risiko sedang karena cenderung berbasis data tetapi tetap tidak memverifikasi semantik narasi.
3. Sistem rekomendasi berbasis AI dikategorikan rendah karena sistem mengevaluasi kesesuaian naratif, audiens, hingga tone of voice.

Opportunity cost dalam penelitian ini dihitung dari Return on Investment (ROI), yang dalam konteks *influencer* marketing merepresentasikan rasio antara nilai hasil kampanye (seperti peningkatan penjualan, *awareness*, atau *engagement*) terhadap biaya yang dikeluarkan untuk kampanye tersebut. Nilai ROI sangat bergantung pada waktu posting, kualitas *engagement*, dan relevansi *influencer*. Oleh karena itu, keterlambatan dalam proses seleksi atau kolaborasi dengan *influencer* yang tidak tepat berpotensi menurunkan ROI aktual yang bisa dicapai brand. Dalam simulasi ini, estimasi *opportunity cost* diperoleh dengan mengasumsikan potensi ROI maksimal yang hilang karena keterlambatan eksekusi kampanye selama beberapa hari, disesuaikan dengan rata-rata estimasi nilai konversi dalam kampanye produk sejenis [26].

Hasil simulasi perbandingan antar pendekatan menunjukkan bahwa solusi berbasis AI menunjukkan potensi efisiensi biaya hingga 79% dibandingkan pendekatan melalui agensi/platform, dan hingga 91% dibandingkan pendekatan manual. Selain itu, sistem mampu mempercepat proses seleksi *influencer* dengan tetap mempertimbangkan aspek relevansi persona dan potensi konversi.

7 Kesimpulan dan Saran

Penelitian ini berhasil merancang dan mengimplementasikan sebuah sistem rekomendasi *influencer* berbasis data yang mengintegrasikan pencocokan semantik narasi, pemodelan psiko-demografi audiens, prediksi performa kampanye, serta optimasi alokasi anggaran dengan pendekatan integer programming. Sistem tidak hanya menghasilkan rekomendasi yang relevan dan terukur, tetapi juga mendukung fleksibilitas strategi melalui skema *brand priorities* yang dapat disesuaikan pengguna. Evaluasi sistem dilakukan terhadap 16 brand dari berbagai industri dengan pendekatan *ablation analysis* dan korelasi peringkat Spearman (ρ). Hasil eksperimen menunjukkan bahwa setiap komponen skor memberikan kontribusi signifikan terhadap hasil akhir, dengan variasi sensitivitas tergantung pada konteks dan tujuan kampanye. Hal ini menunjukkan bahwa sistem tidak hanya akurat, tetapi juga stabil dan dapat dijelaskan secara metrik.

Seluruh rekomendasi dapat ditelusuri dan dijelaskan secara transparan melalui metrik semantik, demografi, dan prediksi dampak, sehingga meningkatkan kepercayaan dan akuntabilitas dalam pengambilan keputusan. Pengujian pada brand dari sektor yang berbeda juga menunjukkan kemampuan sistem untuk beradaptasi terhadap berbagai strategi kampanye, mulai dari *awareness* hingga efisiensi biaya atau konversi tersegmentasi. Sistem telah dibangun dalam bentuk prototipe berbasis web yang interaktif, memungkinkan brand untuk memilih preset strategi kampanye, melihat daftar rekomendasi, serta meninjau insight dari caption dan komentar secara *real-time*. Fitur seperti visualisasi metrik, optimalisasi anggaran, dan pemeringkatan multi-objektif menjadikan sistem ini siap diintegrasikan dalam proses seleksi *influencer* di lingkungan industri.

Ke depan, pengembangan sistem dapat diarahkan pada pembelajaran preferensi brand secara otomatis melalui feedback pengguna, integrasi lintas platform seperti TikTok dan YouTube, serta validasi terhadap hasil kampanye nyata melalui uji lapangan atau A/B testing. Potensi implementasi sistem ini tidak hanya terbatas pada agensi pemasaran, tetapi juga dapat diadaptasi dalam platform rekomendasi konten dan perencanaan kampanye digital berskala besar.




References


- [1] M. Bunea and M. Dumitru, "The impact of marketing innovations on competitiveness," *ResearchGate*, 2023. <https://www.researchgate.net/publication/386199649>
- [2] Badan Pusat Statistik, "Statistik E-Commerce 2023," *bps.go.id*, Jan. 30, 2025. <https://www.bps.go.id/id/publication/2025/01/30/d52af11843aee401403ecfa6/statistik-e-commerce-2023.html>
- [3] L. G. Schiffman and J. Wisenblit, *Consumer Behavior*, 12th ed. Boston, MA: Pearson, 2019.
- [4] INS Global, "Influencer marketing in Indonesia," *insg.co*, 2023. <https://www.insg.co/en/influencer-marketing-indonesia/>
- [5] Partipost, "Influencer Marketing Report 2023," *Partipost Whitepaper*, 2023. <https://www.partipost.com/whitepaper/influencer-marketing-report-2023>.
- [6] Bonsai, "Influencer marketing statistics," *ecommercebonsai.com*, 2024. <https://ecommercebonsai.com/influencer-marketing-statistics/>
- [7] D. P. Kurniawan and A. R. Hidayat, "Cancel culture's torrent: The challenges faced by celebrity and brand," *ResearchGate*, 2024. https://www.researchgate.net/publication/389219628_Cancel_Culture's_Torrent_The_Challenges_Faced_by_Celebrity_and_Brand.
- [8] A. E. Ukozor and J. E. Nwankwo, "Challenges of celebrity endorsement and how they affect the diffusion process," *European Journal of Business and Innovation Research*, vol. 11, no. 3, pp. 39–51, 2023. <https://www.eajournals.org/wp-content/uploads/Challenges-of-Celebrity-Endorsement-and-How-They-Affect-the-Diffusion-Process.pdf>.
- [9] Z. A. R. Suhma, "Cancel culture di Indonesia: Penghakiman sepihak dan merusak etika komunikasi," *Kompasiana*, Dec. 25, 2023. <https://www.kompasiana.com/zakkyakhfashramadanisuhma6031/68486da034777c17cd72d7b2/cancel-culture-di-indonesia-penghakiman-sepihak-dan-merusak-etika-komunikasi>.
- [10] Lumen Learning, "The Role of Customers in Marketing," *Intro to Business*, SUNY Herkimer Community College. <https://courses.lumenlearning.com/suny-hccc-introbusiness/chapter/the-role-of-customers-in-marketing/>
- [11] P. Mathur and R. Jain, "The effect of celebrity endorsements on consumers' brand preference," *Journal of Creative Communications*, vol. 14, no. 3, pp. 215–232, 2019. <https://journals.sagepub.com/doi/10.1177/0973258619875604>.
- [12] A. P. Silaban and R. N. Sitompul, "Pengaruh segmenting, targeting, positioning terhadap keputusan pembelian konsumen pada produk Zhalsa.ID," *Jurnal MEA (Manajemen, Ekonomi, dan Akuntansi)*, vol. 6, no. 3, pp. 1620–1629, 2022. <https://journal.stiemb.ac.id/index.php/mea/article/view/2540>.
- [13] S. Barry, "The development of hierarchy of effects model in advertising," *ResearchGate*, 1987. https://www.researchgate.net/publication/230757483_The_Development_of_Hierarchy_of_Effects_Model_in_Advertising.
- [14] Funnel Secrets, "AIDA model explained: Examples and how to use it," *funnelsecrets.us*, 2023. <https://www.funnelsecrets.us/aida-model/>
- [15] Influencity, "Influencer Discovery Platform," *Influencity.com*, 2025, <https://influencity.com/platform/influencer-discovery/>.

- [16] LinkedIn, “Your brand is receiving conflicting advice on social media. How do you decide which influencer to follow?,” *LinkedIn Advice Community*, 2025. <https://www.linkedin.com/advice/3/your-brand-receiving-conflicting-advice-social-ge0we>.
- [17] FasterCapital, “Comparative costing: Comparative Cost Analysis – A Guide for Decision Makers,” <https://fastercapital.com/content/Comparative-costing-Comparative-Cost-Analysis--A-Guide-for-Decision-Makers.html>.
- [18] Indeed, “Brand Marketing Salaries in Indonesia,” *Indeed*, 2025. <https://id.indeed.com/career/brand-marketing/salaries>.
- [19] KOL.ID, “Paket Layanan Influencer,” 2025. <https://kol.id/dashboard/package/kol>.
- [20] Allstars, “Campaign Management,” 2025. <https://client.allstars.id/home/campaign>.
- [21] ICE, “Create Brief Page,” 2025. <https://www.ice.id/campaign/create-brief>.
- [22] Partipost, “Campaign Management Dashboard,” 2025. <https://cm.partipost.com/campaigns>.
- [23] Intero Digital, “How Much Does Content Marketing Cost? A Full Breakdown,” 2025. <https://www.interodigital.com/blog/how-much-does-content-marketing-cost-a-full-breakdown/>.
- [24] Innova AI, “AI vs Traditional Marketing Tools: ROI Comparison Guide,” 2025. <https://www.innovaai.io/blog/ai-vs-traditional-marketing-tools-roi-comparison-guide/>.
- [25] v500 Systems, “Multiple Document Comparison: Manual vs AI,” 2025. <https://www.v500.com/multiple-document-comparison-manual-vs-ai-automated/>.
- [26] Influencer Marketing Hub, “Influencer Marketing ROI: How to Calculate and Measure Campaign Success,” 2025. <https://influencermarketinghub.com/influencer-marketing-roi/>.

LAMPIRAN

Lampiran 1. Analisis Kompetitor

Nama Kompetitor	Deskripsi Layanan	Target Brand	Platform yang Ditinjau	Kelebihan	Kelemahan
 KOL.ID	Menampilkan dashboard analisis sosial media dan rate card dari <i>influencer</i> pada kategori <i>micro</i> (followers 10.000-50.000) dan kategori <i>mid-tier</i> (50.000-100.000)	Mid to Large	Instagram	Menyediakan 1.000-1.500 alternatif pilihan influencer dan memungkinkan brand untuk mengontak influencer secara langsung	Filtering masih dilakukan secara manual dan hanya menyediakan data tanpa mempertimbangkan kecocokan influencer dengan brand
 Allstars	Menyediakan dashboard campaign manajemen, memiliki 2 tools utama, yaitu Allstars Project dan Allstars Assistant. Allstars Project membantu brand terhubung dengan influencer melalui konten yang dipersiapkan sendiri, sementara pada Allstars Assistant penerbitan konten akan didesain oleh Allstars.	Large	Tiktok, YouTube, Twitter	Pilihan influencer yang disediakan telah terverifikasi dan dikurasi oleh platform keaslian akunnya	Allstars tidak menyediakan analisis sosial media seperti engagement rate maupun kualitas konten
 ICE	Platform berbentuk komunitas yang menghubungkan brand	Startup	Tiktok, Instagram (namun	Berfokus pada community driven dan	Fokus pada metode partnership dalam skala besar

	dengan influencer		cenderung didominasi influencer (Tiktok)	kolaborasi kreatif	
 PARTIPOST <i>Snap . Post . Earn</i> Partipost	Platform crowd-sourced influencer, sehingga semua orang dapat menjadi influencer dalam platform ini.	Mid to Large	Tiktok, Instagram	Menyediakan 500-1.000 alternatif pilihan influencer dan memungkinkan brand untuk mengontak influencer secara langsung	Tidak mengontrol kualitas influencer yang ada dalam platform dan tidak menampilkan insight kecocokan brand dan influencer

Lampiran 2. Analisis *Value Proposition*

Market Validation	
1. Customer Jobs (Identifikasi pekerjaan brand yang butuh dibantu)	<ul style="list-style-type: none"> ● Mempersingkat prosedur yang dilakukan ketika brand membutuhkan kerja sama dengan influencer. Brand tidak perlu melakukan pengecekan secara satu persatu lagi untuk mengidentifikasi kecocokan influencer dengan value yang dimiliki oleh brand. ● Mengurangi resiko kesalahan pemilihan influencer yang berujung pada penolakan produk oleh masyarakat, hal ini juga dapat membantu brand untuk menjaga citranya dalam jangka panjang. ● Membantu brand menilai influencer yang beretika dan memiliki kepercayaan di mata masyarakat sehingga sesuai dengan tujuan brand dalam ESG (Environmental, Social, dan Governance).
2. Pains (Identifikasi keresahan brand)	<ul style="list-style-type: none"> ● Mengambil keputusan secara subjektif yang hanya didasarkan pada kesesuaian dengan budget atau memilih influencer yang sedang viral saja, hal ini menyebabkan rawan terjadi kesalahan

	<p>dalam pemilihan yang berdampak pada citra yang dimiliki brand.</p> <ul style="list-style-type: none"> • Tidak memiliki struktur penilaian yang mencakup keseluruhan aspek influencer dan brand. • Tingginya biaya jika menggunakan platform influencer marketplace dan agensi.
3. Gains (Identifikasi fasilitas yang brand inginkan)	<ul style="list-style-type: none"> • Sistem rekomendasi influencer yang relevan untuk mengidentifikasi kecocokan dengan brand dan terukur dari sisi performancenya. • Kemudahan dalam memilih influencer yang sesuai dengan ketersediaan budget yang dimiliki oleh brand • Kemudahan dalam melihat potensi influencer yang akan digunakan untuk meningkatkan awareness konsumen terhadap produk dan penjualan. • Solusi yang bisa diperluas ke platform lainnya (seperti tiktok, youtube, dan sebagainya) sehingga mampu meningkatkan performa marketing secara optimal.
Product Validation	
1. Gain Creators (Identifikasi apa yang bisa memberikan nilai lebih pada brand)	<ul style="list-style-type: none"> • Memetakan kecocokan antara influencer dengan brand dengan mempertimbangkan aspek kesesuaian profil brand dengan influencer dan mengecek apakah influencer masuk dalam ketersediaan budget brand dan memiliki potensial meningkatkan pembelian produk. • Menyediakan prediksi valid persona yang ditampilkan influencer sehingga brand dapat menyiapkan briefing campaign secara lebih optimal dengan menekankan keunggulan influencer tersebut.
2. Pain Relievers (Identifikasi keresahan brand yang dapat	<ul style="list-style-type: none"> • Melakukan identifikasi prioritas influencer berdasarkan matrix yang komprehensif, yaitu menilai kesesuaian budget berdasarkan rate card, menilai kecocokan berdasarkan analisis brand dan influencer persona, dan menilai potential peningkatan

dikurangi)	<p>pembelian melalui analisis CTA (Call-to-Action) dan komentar audiens.</p> <ul style="list-style-type: none"> ● Menggunakan analisis sentimen dari konsumen terhadap influencer untuk menghindari penggunaan artis yang punya tren negatif. ● Menghasilkan insight dari postingan influencer di Instagram untuk membantu brand melakukan penilaian apakah influencer yang digunakan sudah sesuai dan memiliki potensi dalam meningkatkan penjualan.
3. Product (Identifikasi produk yang ditawarkan)	<ul style="list-style-type: none"> ● Rekomendasi top 3 match yang didasarkan atas kesesuaian budget brand, kecocokan dengan image brand, dan potensi influencer untuk meningkatkan pembelian produk dari brand. ● Peningkatan efisiensi pemilihan influencer dengan AI yang dapat melakukan proses pemilihan kurang dari 3-5 hari kerja. ● Mengurangi biaya yang dikeluarkan brand untuk memilih influencer yang sesuai.

Lampiran 3. Pengembangan Fitur Lanjutan

Tahapan	Timeline	Deskripsi Pengembangan
Fase 1: Stabilitas dan integrasi dasar	Bulan 1-3	Optimasi pipeline pemilihan influencer agar mendukung integrasi lintas platform (Instagram, TikTok). Penambahan opsi filtering berdasarkan lokasi, niche, dan brand safety level.
Fase 2: Analisis sentimen lanjutan	Bulan 4-6	Ke depan, sistem dapat dikembangkan lebih lanjut dengan mengintegrasikan pendekatan <i>aspect-based sentiment analysis</i> untuk membedakan opini audiens terhadap produk dan terhadap <i>influencer</i> , serta pendeteksian emosi dari komentar untuk menggali dimensi psikologis dari interaksi audiens.

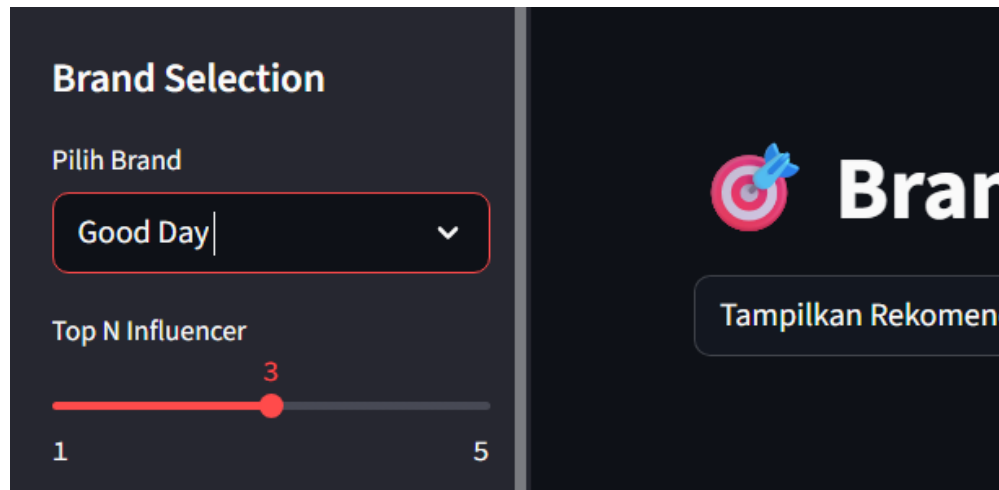
		<p>Selain itu, sistem yang saat ini menyediakan empat skema <i>preset</i> strategi kampanye dapat ditingkatkan menjadi lebih dinamis. Alih-alih hanya memilih dari template tetap, pengguna dapat memberikan deskripsi atau narasi strategi kampanye dalam bentuk teks bebas. Sistem kemudian dapat melakukan analisis semantik terhadap kalimat tersebut untuk menghasilkan pembobotan prioritas dan perancangan strategi yang lebih sesuai dan kontekstual. Pendekatan ini memungkinkan sistem beradaptasi dengan kebutuhan spesifik tiap brand secara lebih presisi, serta meningkatkan fleksibilitas dan kecerdasan sistem dalam merancang solusi rekomendasi</p>
Fase 3: Deteksi fake engagement	Bulan 7-8	Deteksi lonjakan followers tidak wajar, analisis pola komentar spam, dan identifikasi engagement manipulatif (bot-like behavior).
Fase 4: Dashboard interaktif lanjutan	Bulan 9-10	Pengembangan fitur lanjutan dashboard web untuk brand: Simulasi ROI, membuat target campaign lebih fleksibel
Fase 5: Penyesuaian dengan ESG dan value-based marketing	Bulan 11	Integrasi filter nilai dan etika influencer berbasis analisis naratif serta rekam jejak kampanye sebelumnya.
Fase 6: Scaling data	Bulan 12	Perluasan dataset caption-komentar dengan semi-supervised learning dan pretraining BERT khusus industri influencer marketing.

Lampiran 4. Panduan Penggunaan Platform Influencer-Brand Matcher

Link web : <https://influencermatcher-yujiem-rookie.streamlit.app/>

Langkah-langkah penggunaan


1. Pilih brand yang ingin dianalisis dari daftar yang tersedia. Atur jumlah influencer yang ingin ditampilkan melalui *slider* “Top N Influencer” (1–5). Centang opsi **Tampilkan detail proses** untuk melihat tahapan seleksi yang dilakukan sistem (opsional).




2. Setelah memilih brand dan jumlah influencer, pilih skema prioritas sesuai dengan tujuan kampanye yang dijalankan brand. Keterangan mengenai skema yang terpilih dan persentase pembobotan prioritas, tertera di bagian bawah setelah pemilihan skema.

Pilih Skema Prioritas Brand

Targeted Campaign ▾

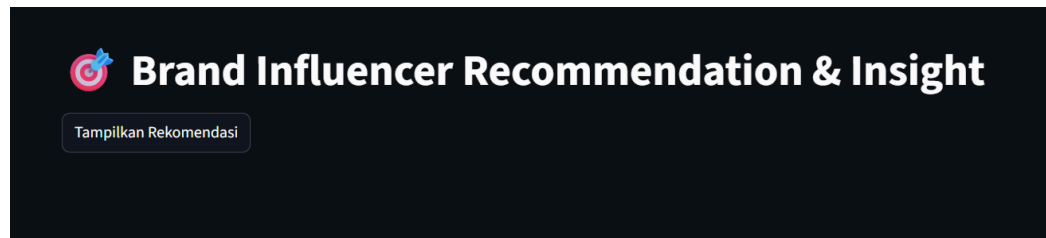
 Skema Terpilih: Targeted Campaign

 Bobot Prioritas:

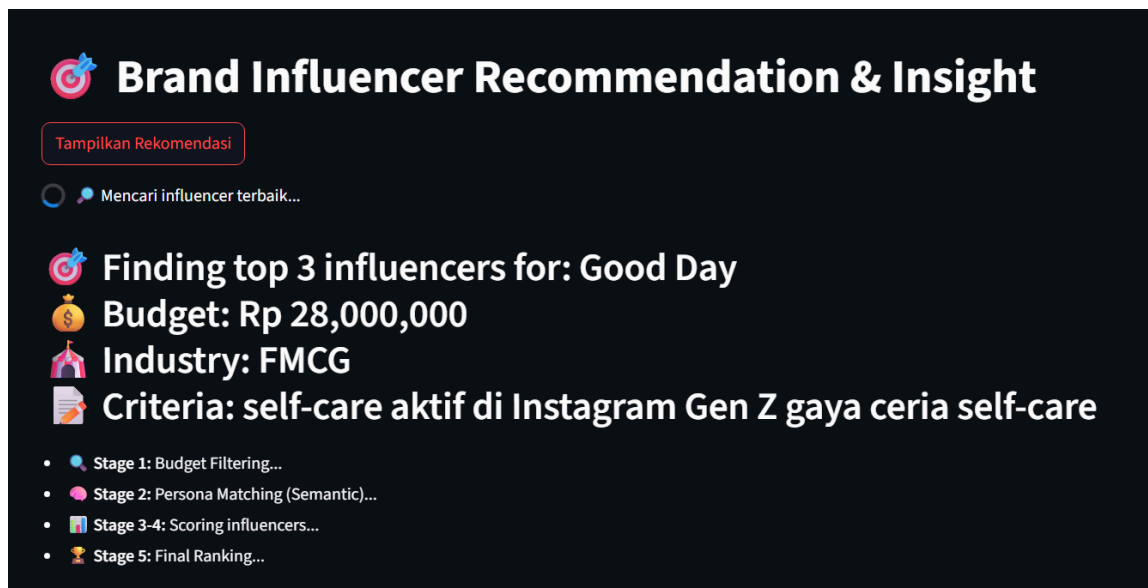
Komponen	Bobot
Persona Fit	10%
Demographic Fit	45%
Performance	35%
Budget Efficiency	10%

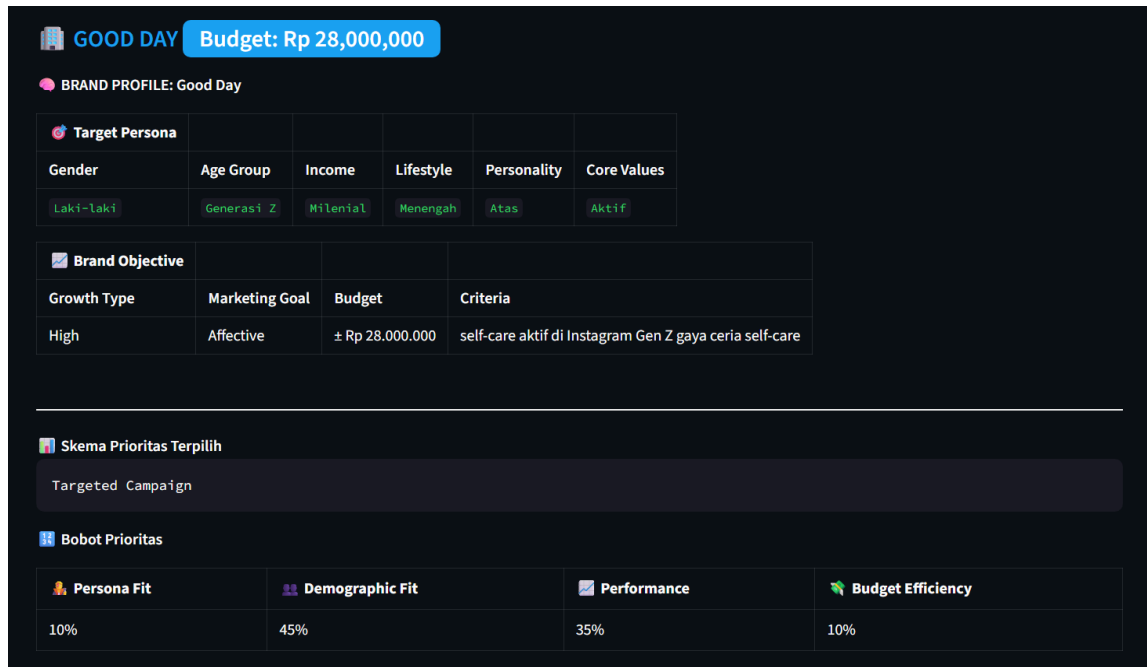
Preset skema prioritas pada sistem ini memungkinkan brand menyesuaikan strategi pemilihan influencer sesuai tujuan kampanye. Skema **Balanced** memberikan bobot yang sama pada semua aspek utama, yaitu kecocokan persona, kesesuaian demografi-psikografi, potensi performa, dan efisiensi anggaran yang membuatnya cocok untuk brand yang menginginkan hasil rekomendasi yang seimbang tanpa menonjolkan satu aspek tertentu. Skema **Targeted Campaign** memfokuskan bobot pada kesesuaian demografi-psikografi dan potensi performa, sehingga sangat sesuai untuk kampanye yang menargetkan segmen pasar spesifik dan ingin memastikan influencer benar-benar relevan dengan audiens yang dituju. Skema **Awareness Campaign** menonjolkan aspek potensi performa (terutama reach dan engagement) serta kecocokan persona, sehingga cocok digunakan saat brand ingin menjangkau audiens seluas mungkin dan membangun eksposur. Skema **Cost Efficiency** memprioritaskan efisiensi anggaran, sehingga sangat tepat bagi brand dengan budget terbatas atau yang ingin memaksimalkan dampak kampanye dengan biaya serendah mungkin. Terakhir, skema **Persona-Centric Brand** menempatkan fokus utama pada kecocokan persona antara influencer dan brand, sehingga sangat sesuai untuk brand yang sangat menjaga citra, nilai, dan storytelling, misalnya pada produk lifestyle, beauty, atau brand dengan positioning unik. Dengan fleksibilitas ini, brand dapat memilih skema yang paling sesuai dengan kebutuhan dan objektif kampanye mereka.

3. Setelah memilih skema prioritas brand, klik tombol **Tampilkan Rekomendasi**. Sistem akan memulai proses seleksi dan analisis berdasarkan kriteria brand dan performa influencer.

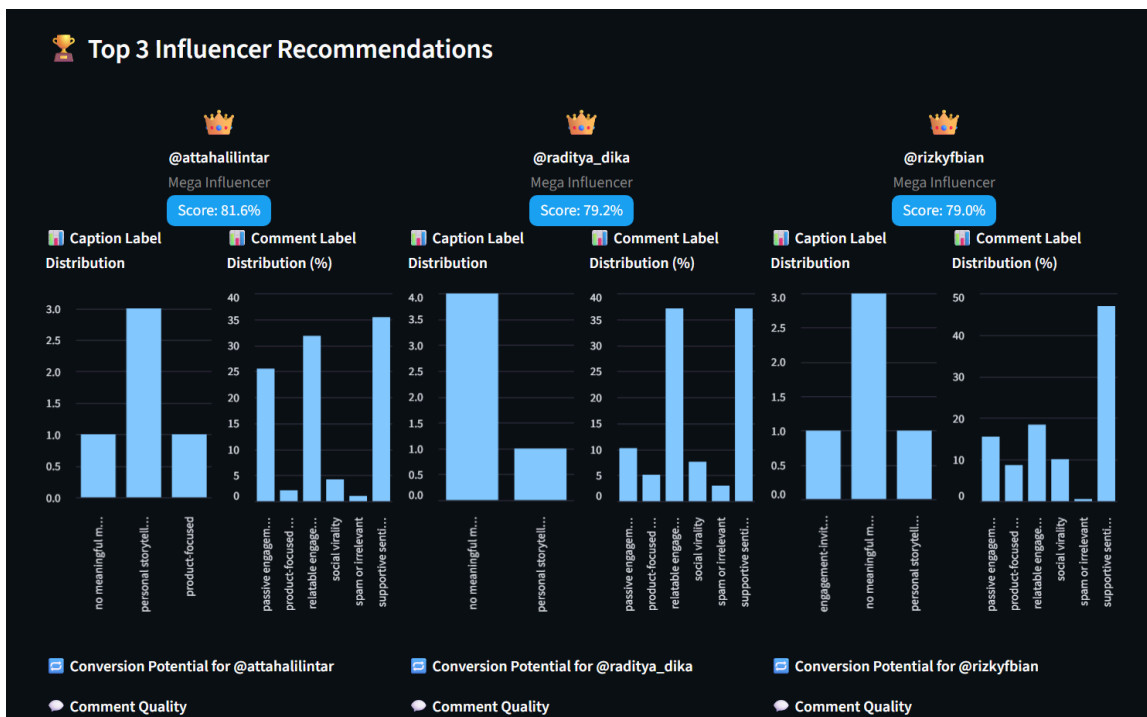


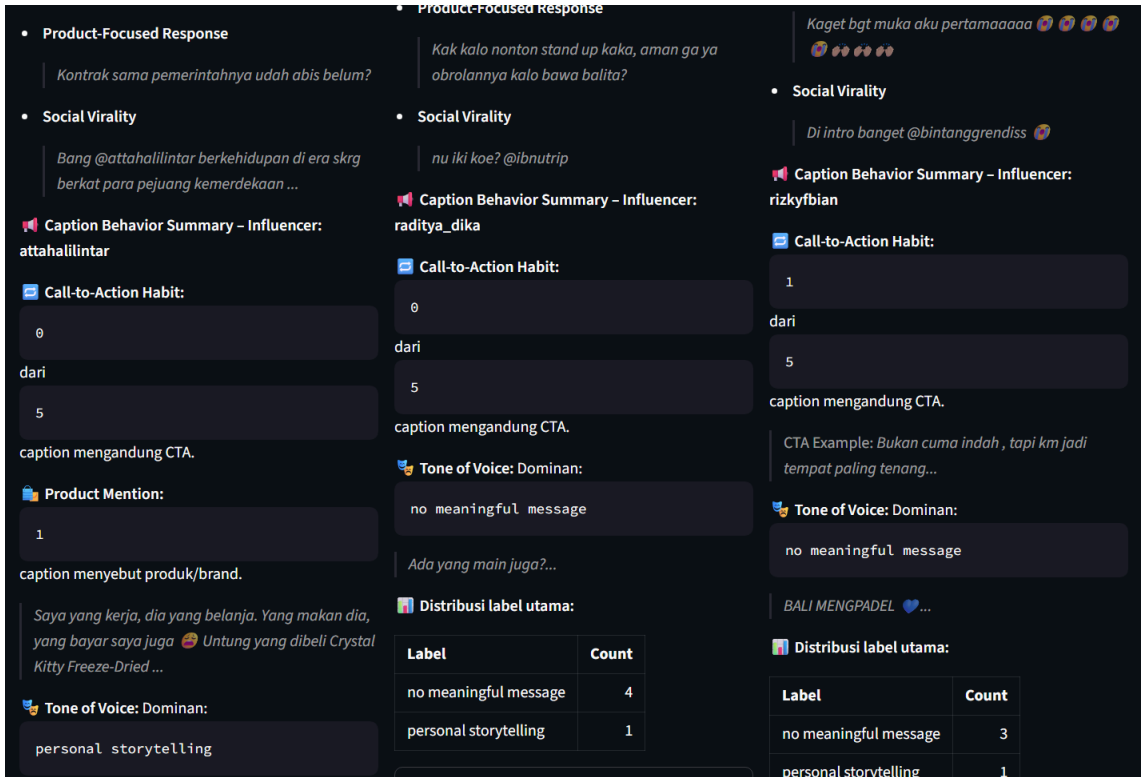
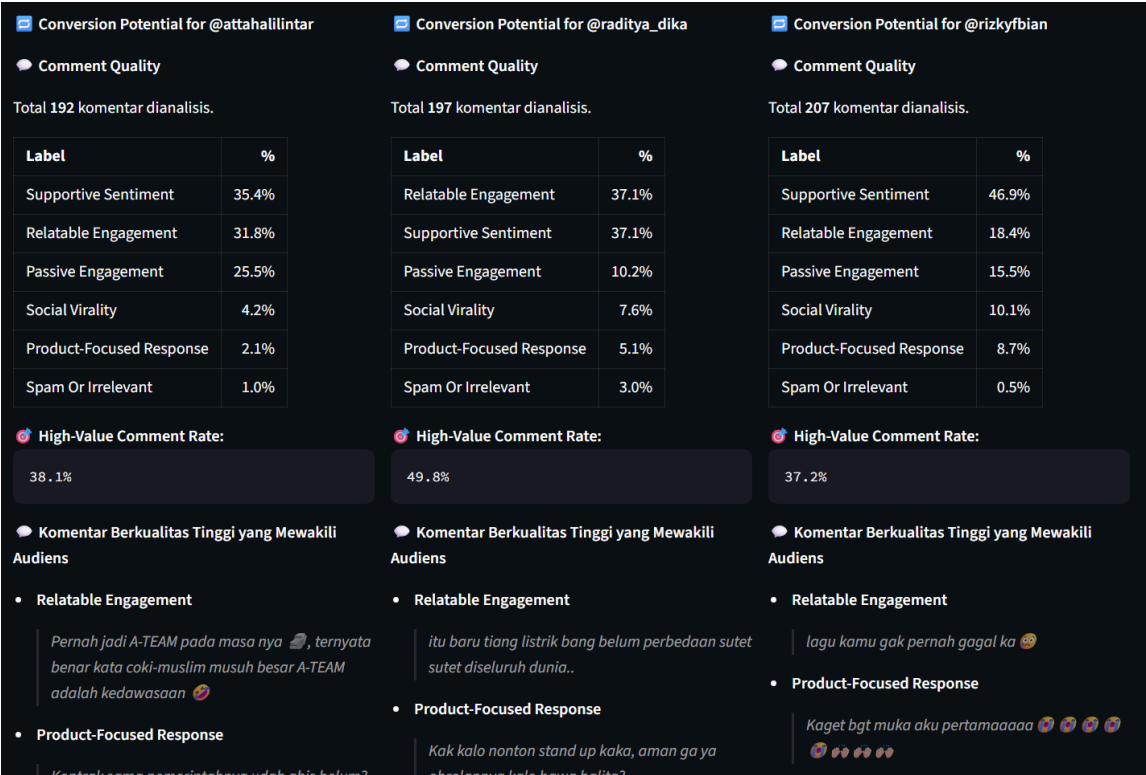
4. Setelah tombol **Tampilkan Rekomendasi** ditekan, sistem akan menampilkan ringkasan brand yang dipilih. Animasi loading menandakan sistem sedang menganalisis data influencer secara menyeluruh berdasarkan kriteria yang ditetapkan.





- Setelah proses selesai, sistem akan menampilkan **Top N Influencer** sesuai jumlah yang dipilih. Pengguna dapat scroll ke bawah untuk melihat rincian skor evaluasi dan rencana campaign rekomendasi sesuai konten optimal (feeds, story, reels).





🕌 Alhamdulillah Masjid Ahha ke 5 sudah rampung

💖 Masjid ini semoga jadi manfaat buat santri yatim piatu hafiz dan juga m...

Distribusi label utama:

Label	Count
personal storytelling	3
no meaningful message	1
product-focused	1

Detail & Metrics

Score Breakdown Campaign Plan

- Budget Efficiency: 0.21 points/Million Rp
- Persona Fit: 28.9%
- Demographic Fit: 73.3%
- Performance: 60.6%
- Engagement Rate: 6.10%
- Authenticity: 58.7%
- Reach: 44.4%
- Brand Fit: 23.9%

💡 Key Insights:

- 🔥 Currently trending
- ✅ Consistent content behavior
- 🎯 Proven campaign success

personal storytelling	1
engagement-inviting	1

Juga ditemukan upaya interaksi dua arah dengan audiens.

Detail & Metrics

Score Breakdown Campaign Plan

- Budget Efficiency: 0.27 points/Million Rp
- Persona Fit: 32.8%
- Demographic Fit: 73.3%
- Performance: 59.8%
- Engagement Rate: 5.70%
- Authenticity: 61.5%
- Reach: 33.6%
- Brand Fit: 30.2%

💡 Key Insights:

- 🔥 Currently trending
- ✅ Consistent content behavior
- 🎯 Proven campaign success

personal storytelling

🕌 Alhamdulillah Masjid Ahha ke 5 sudah rampung

💖 Masjid ini semoga jadi manfaat buat santri yatim piatu hafiz dan juga m...

Distribusi label utama:

Label	Count
personal storytelling	3
no meaningful message	1
product-focused	1

Detail & Metrics

Score Breakdown Campaign Plan

🔥 Optimal Campaign Strategy:

- Instagram Stories: 1 posts

📊 Financial Summary:

- Total Investment: Rp 16,000,000
- Budget Remaining: Rp 12,000,000
- Expected Impact: 3.4 points

personal storytelling	1
engagement-inviting	1

Juga ditemukan upaya interaksi dua arah dengan audiens.

Detail & Metrics

Score Breakdown Campaign Plan

🔥 Optimal Campaign Strategy:

- Instagram Stories: 2 posts

📊 Financial Summary:

- Total Investment: Rp 24,000,000
- Budget Remaining: Rp 4,000,000
- Expected Impact: 6.4 points