

ANALISIS KORELASI SENTIMEN YOUTUBE GADGETIN DAN KEPUASAN KONSUMEN TOKOPEDIA MENGGUNAKAN SPEARMAN RANK DENGAN PENDEKATAN NAÏVE BAYES DAN INDOBERT

Muhammad Ibnu Syna¹, Martanto², Fatihanursari Dikananda³, Mulyawan⁴.

Program Studi Teknik Informatika¹³
Program Studi Manajemen Informatika²
Program Studi Sistem Informasi⁴

STMIK IKMI Cirebon
<https://ikmi.ac.id/page/18/?lang=de>
muhammad.ibnuavcn@gmail.com

(*) Corresponding Author : muhammad.ibnuavcn@gmail.com
Published : 15 July 2026

Abstract—The rapid growth of social media has made digital reviews one of the main sources shaping public perception of products, including local technology brands such as Advan laptops. This study analyzes the relationship between public sentiment on the GadgetIn YouTube channel and the satisfaction of Advan laptop buyers on Tokopedia. The methodology used integrates the Naïve Bayes IndoBERT model for automatic labeling, validated using Cohen's Kappa coefficient, yielding values of 0.60736 (Tokopedia) and 0.71876 (YouTube) in the Substantial Agreement category. The challenge of data imbalance was addressed using the Synthetic Minority Over-sampling Technique (SMOTE) on the TF-IDF feature-based Naïve Bayes algorithm. The results show that implementing SMOTE significantly improved model performance; accuracy jumped from 71% to 97.19% on Tokopedia and from 56% to 79.52% on YouTube. The final sentiment distribution on Tokopedia was dominated by positive reviews (80.3%), while on YouTube it was more balanced (27.1% negative, 39.3% neutral, and 33.6% positive). The Spearman Rank Correlation Test yielded a coefficient of $\rho = 0,6646$ with a $-value = 0,0360$ ($\rho < 0,05$), confirming a strong and statistically significant positive relationship between digital perceptions on YouTube and actual buyer satisfaction on Tokopedia. This study demonstrates that the combination of deep learning (IndoBERT) for data preparation and data balancing techniques (SMOTE) in classical machine learning is effective in mitigating classification bias and provides empirical insights into the relationship between social media opinions and consumer experiences in Indonesia's local technology market.

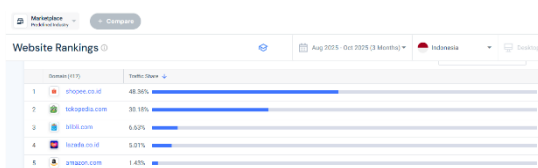
Keywords: Sentiment Analysis, IndoBERT, SMOTE (Synthetic Minority Over-sampling Technique), Naïve Bayes, Spearman Rank Correlation.

Abstrak—Perkembangan media sosial telah menjadikan ulasan digital sebagai sumber utama pembentukan persepsi publik terhadap produk, termasuk pada sektor teknologi lokal seperti laptop merek Advan. Penelitian ini menganalisis hubungan antara sentimen publik di YouTube GadgetIn dan kepuasan pembeli laptop Advan di Tokopedia. Metodologi yang digunakan mengintegrasikan model Naïve Bayes IndoBERT dan untuk pelabelan otomatis yang divalidasi menggunakan koefisien Cohen's Kappa, menghasilkan nilai 0,60736 (Tokopedia) dan 0,71876 (YouTube) dalam kategori Substantial Agreement. Tantangan ketidakseimbangan data diatasi melalui teknik Synthetic Minority Over-sampling Technique (SMOTE) pada algoritma Naïve Bayes berbasis fitur TF-IDF. Hasil penelitian menunjukkan bahwa implementasi SMOTE secara signifikan meningkatkan performa model; akurasi melonjak dari 71% menjadi 97,19% pada Tokopedia dan dari 56% menjadi 79,52% pada YouTube. Distribusi sentimen akhir pada Tokopedia didominasi ulasan positif (80,3%), sedangkan di YouTube lebih berimbang (27,1% negatif, 39,3% netral, dan 33,6% positif). Uji Korelasi Spearman Rank menghasilkan koefisien $\rho = 0,6646$ dengan $-value = 0,0360$ ($\rho < 0,05$), yang mengonfirmasi adanya hubungan positif yang kuat dan signifikan secara statistik antara persepsi digital di YouTube dengan kepuasan aktual pembeli di Tokopedia. Penelitian ini membuktikan bahwa kombinasi *deep learning* (IndoBERT) untuk persiapan data dan teknik penyeimbangan data (SMOTE) pada *machine learning* klasik efektif dalam memitigasi bias klasifikasi serta memberikan wawasan empiris mengenai keterkaitan opini media sosial dan pengalaman konsumen di pasar teknologi lokal Indonesia.

Kata Kunci: Analisis Sentimen, *IndoBERT*, SMOTE (*Synthetic Minority Over-sampling Technique*), *Naïve Bayes*, Korelasi *Spearman Rank*.

INTRODUCTION

Perkembangan teknologi informasi telah mendorong perubahan besar dalam perilaku konsumen, termasuk pergeseran dari belanja konvensional menuju ekosistem digital seperti e-commerce. Di Indonesia, tren ini didorong oleh meningkatnya penetrasi internet, penggunaan smartphone, serta gaya hidup masyarakat yang mengutamakan efisiensi dan kenyamanan (Asrul, 2024).



Gambar 1 Website Rankings

Salah satu platform yang mendominasi pasar Indonesia adalah Tokopedia. Dalam lanskap Indonesia, Tokopedia menempati posisi yang sangat strategis. Diluncurkan pada tahun 2009 yang menerapkan pendekatan berbasis data dalam pengembangan platformnya. Sistem rekomendasi, ulasan pelanggan, dan fitur rating produk menjadi fondasi penting dalam membangun kepercayaan konsumen dan memperkuat loyalitas pengguna (Setyawan et al., 2022). Sehingga, kini Tokopedia menjadi salah satu dengan pengunjung situs bulanan tertinggi di Indonesia. Hal ini didukung dengan data yang dibagikan oleh Similarweb tentang Top Marketplace Website Ranking in Indonesia. Dalam data tersebut menjelaskan bahwa pada bulan Mei 2025, pengunjung situs bulanan Tokopedia mencapai 62.5 juta pengunjung dengan 95.90% pengunjung dari Indonesia.

Pergeseran perilaku konsumen dalam mencari informasi juga semakin terlihat dengan maraknya penggunaan media sosial, di mana akses informasi tidak lagi dimonopoli oleh media tradisional atau iklan korporat. Kini, konsumen cenderung beralih ke sumber-sumber yang dianggap lebih otentik dan dapat dipercaya (Ardani, 2022). Pergeseran ini melahirkan ekosistem influencer marketing, di mana konten kreator memainkan peran signifikan dalam membentuk persepsi dan preferensi konsumen (Permana & Astuti, 2023). Konsumen cenderung memandang influencer sebagai "orang biasa" yang lebih dapat dipercaya dibandingkan iklan konvensional. (Pesa Nugraha, 2023). Kepercayaan ini menjadi

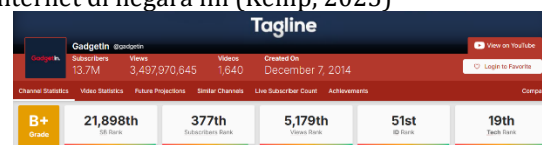
fondasi bagi influencer untuk membangun koneksi emosional dan memengaruhi keputusan pembelian audiens mereka secara signifikan (Rosyidani et al., 2024).

Pengaruh influencer menjadi sangat vital, terutama untuk produk dengan pertimbangan tinggi (high-consideration products) seperti laptop, di mana konsumen membutuhkan informasi mendalam dan tepercaya sebelum melakukan investasi finansial. Pemasaran influencer telah menjadi strategi penting dalam lanskap digital, khususnya untuk produk-produk yang banyak diminati (Lavenia & Erdiansyah, 2022). Influencer media sosial memainkan peran penting dalam membentuk keputusan pembelian konsumen dengan memberikan informasi tepercaya dan mempromosikan produk secara efektif (Fauzi et al., 2024). Influencer dikategorikan menjadi mikro, makro, dan mega-influencer, yang masing-masing mampu membuat konten menarik yang memengaruhi minat konsumen (Handayani, 2023).



Gambar 2 Penggunaan Internet di Indonesia

Di antara berbagai platform, YouTube adalah sumber informasi utama bagi konsumen sebelum membuat keputusan pembelian (Adha, 2022). Transformasi ini didorong oleh penetrasi digital yang masif di Indonesia, yang pada awal tahun 2025 telah mencapai 212 juta pengguna internet (74,6% dari total populasi) dan 143 juta pengguna aktif media sosial. Di antara berbagai platform, YouTube basis pengguna terbesar keempat di dunia dan platform yang iklannya mampu menjangkau 50,2% dari total pengguna internet di negara ini (Kemp, 2025)



Gambar 3 Statistik YouTube GadgetIn

Fenomena ini menjadi fokus utama penelitian ini, dengan studi kasus pada kanal YouTube GadgetIn. Pemilik saluran, David Brendi, atau yang

akrab dipanggil David GadgetIn, telah meluncurkan saluran ini sejak tahun 2014. Dengan keberadaannya yang sudah lama di platform YouTube, saluran GadgetIn berhasil menjadi salah satu yang paling diminati untuk ditonton. Per Juni 2025, jumlah pelanggan atau subscribers saluran ini telah melampaui angka 13,3 juta, dengan rata-rata jutaan penonton untuk setiap video yang diunggah. Pemilihan GadgetIn sebagai studi kasus didasarkan pada dua pilar utama: jangkauan dan reputasi. Dari segi jangkauan, data pada Social Blade menempatkan GadgetIn sebagai kanal teknologi ke-19 terbesar di dunia dan kanal ke-51 terbesar secara keseluruhan di Indonesia, yang menandakan potensi pengaruh pasar yang luar biasa. Berbagai penelitian telah menguji pengaruh konten GadgetIn terhadap keputusan pembelian konsumen dan menemukan bahwa kualitas konten dan informasi kanal berkontribusi hingga 80% terhadap keputusan pembelian (Natalie & Junaidi, 2024). Review dari GadgetIn dapat secara efektif memvalidasi kualitas produk atau, sebaliknya, mengungkap kekurangannya, sehingga mengurangi risiko yang dirasakan konsumen (Jumawan et al., 2024).

Fenomena menarik tidak hanya berhenti pada konten video itu sendiri, tetapi meluas hingga ke kolom komentar. Kolom komentar pada video review GadgetIn berfungsi sebagai ruang diskusi publik yang masif. Di dalamnya, ribuan calon konsumen, pengguna, dan penggemar teknologi saling bertukar pendapat, bertanya, dan berbagi ekspektasi. Sentimen yang terbentuk di ruang diskusi ini—baik itu positif, negatif, maupun netral—menciptakan sebuah gelombang ekspektasi awal (pre-purchase expectation) terhadap produk yang diulas.

Di tengah dinamika ini, pemerintah Indonesia melalui peta jalan "Making Indonesia 4.0" telah menetapkan industri elektronik sebagai salah satu dari lima sektor manufaktur prioritas yang perlu dikembangkan untuk mendorong pertumbuhan ekonomi nasional (Waluyo, 2024). Inisiatif ini membuka peluang sekaligus tantangan bagi merek-merek elektronik lokal seperti Advan, terutama laptop, untuk meningkatkan daya saing mereka di pasar domestik yang didominasi oleh pemain global. Interaksi antara influencer terkemuka, produk lokal yang strategis, dan platform dominan menciptakan sebuah ekosistem digital yang kompleks dan menarik untuk diteliti (Riyantie & Rusadi, 2024).

Di sisi lain, dari perspektif teknis, analisis sentimen menggunakan metode komputasi untuk mengklasifikasikan opini dalam data teks telah menjadi pendekatan yang mapan (Hokijulandy et al., 2023). Studi-studi ini menunjukkan bahwa Naïve Bayes merupakan pilihan yang layak untuk analisis sentimen komentar YouTube. Beberapa penelitian telah menunjukkan keefektifan algoritma Naïve Bayes untuk analisis sentimen komentar YouTube, dengan tingkat akurasi bervariasi antara 67% dan 90,36%. Saputra & Pribadi, (2023) mencapai akurasi 67% pada klasifikasi tiga label setelah menangani data yang tidak seimbang. Wahid & Saputri, (2022) melaporkan akurasi 66,72% menggunakan Naïve Bayes, mengungguli Decision Tree. Hudha et al., (2022) memperoleh akurasi 90,36% dalam mengklasifikasikan komentar menjadi sentimen positif, negatif, dan netral. Zhafira et al., (2021) mencapai akurasi tertinggi sebesar 97% menggunakan 900 sampel latih, 100 sampel uji, pembobotan TF-IDF, dan validasi silang 10 kali lipat. Langkah-langkah praproses biasanya melibatkan pengumpulan data melalui web scraping, pelabelan manual, pembersihan data, dan praproses teks. TF-IDF umumnya digunakan untuk pembobotan kata (Saputra & Pribadi, 2023; Wahid & Saputri, 2022; Zhafira et al., 2021).

Aspek kebaruan yang ditawarkan dalam riset ini terletak pada pendekatan sistematis yang menggabungkan efisiensi model transformer IndoBERT sebagai instrumen pelabelan otomatis dengan ketangguhan teknik SMOTE dalam menyeimbangkan ruang fitur. Kombinasi ini kemudian diterapkan pada algoritma Naïve Bayes Classifier untuk membedakan dinamika opini antara ekspektasi audiens di YouTube dan realitas pengalaman pengguna di Tokopedia. Pendekatan lintas platform ini memberikan perspektif baru dalam memahami bagaimana penyeimbangan data secara sintesis dapat meningkatkan reliabilitas hasil uji korelasi statistik Spearman Rank pada domain produk teknologi lokal.

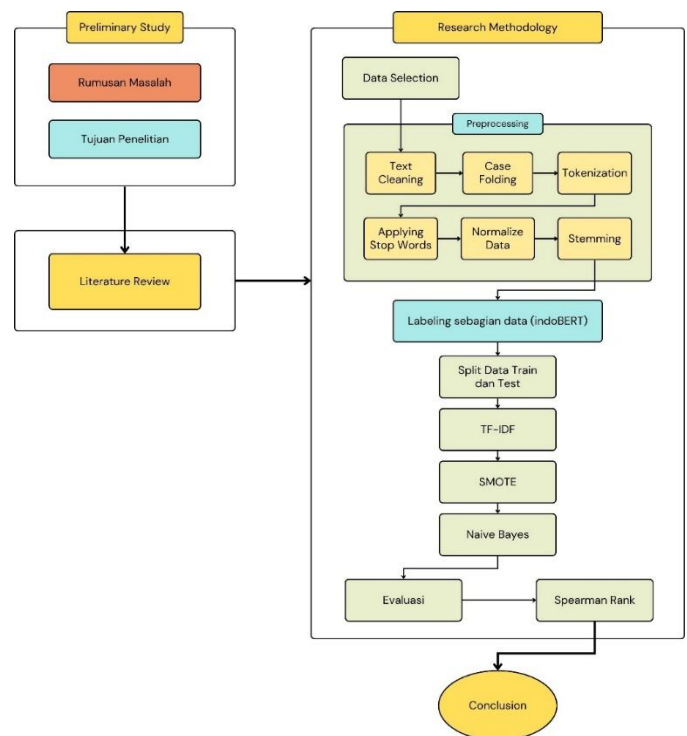
Jika penelitian ini berhasil mencapai tujuannya, yaitu menganalisis hubungan antara sentimen publik pada kanal YouTube GadgetIn dengan tingkat kepuasan pembeli laptop Advan di Tokopedia melalui kombinasi model IndoBERT sebagai alat bantu pelabelan otomatis dan teknik SMOTE dalam menyeimbangkan ruang fitur serta algoritma Naïve Bayes sebagai model klasifikasi utama yang diuji menggunakan Spearman Rank Correlation, maka hasilnya akan memberikan kontribusi penting terhadap pemahaman empiris mengenai dinamika opini publik digital dan perilaku kepuasan konsumen dalam konteks e-commerce lokal.

Temuan ini diharapkan tidak hanya memperkaya literatur terkait efektivitas influencer dan konten ulasan teknologi dalam membentuk persepsi konsumen, tetapi juga memperkuat validitas pendekatan analisis sentimen berbasis machine learning sebagai indikator awal kepuasan pelanggan. Penggunaan Naïve Bayes yang terintegrasi dengan representasi fitur TF-IDF memungkinkan penelitian ini berfungsi sebagai model referensi dalam pengembangan sistem klasifikasi opini publik yang efisien dan mudah direplikasi untuk konteks riset pemasaran digital di Indonesia.

Secara praktis, hasil penelitian ini dapat dimanfaatkan oleh pelaku industri dan brand lokal seperti Advan untuk mengevaluasi kesesuaian antara persepsi publik (pra-pembelian) dan pengalaman pengguna aktual (pascapembelian), serta menilai efektivitas strategi komunikasi melalui influencer. Dengan demikian, penelitian ini memiliki implikasi strategis terhadap peningkatan kualitas komunikasi digital, optimalisasi pengambilan keputusan berbasis data (data-driven decision making), dan pengembangan sistem rekomendasi yang lebih adaptif terhadap persepsi publik dan kepuasan pelanggan di era pemasaran berbasis komunitas daring.

MATERIALS AND METHODS

Desain penelitian ini mengikuti kerangka kerja eksperimen semu berbasis data-driven, sebagaimana disarankan oleh Macanović et al., (2024) dan Onah et al., (2023). Proses penelitian dirancang secara sistematis agar dapat direplikasi dan diukur secara kuantitatif melalui tahapan berikut:



Gambar 4 Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan desain korelasional berbasis data yang mencakup beberapa tahapan sistematis, yaitu:

Pengumpulan dan Sampling Data pengambilan komentar YouTube dari kanal GadgetIn dan ulasan pembeli di Tokopedia dilakukan menggunakan teknik *purposive sampling* untuk memastikan data relevan dengan konteks pembelian laptop lokal (Advan). Data yang digunakan dalam penelitian ini terdiri atas ribuan komentar dan ulasan yang dikumpulkan secara purposif, dengan jumlah akhir mencapai lebih dari empat ribu entri setelah proses pembersihan dan normalisasi.

Pra-pemrosesan Teks (Preprocessing) mencakup tahap *data cleaning*, *case folding*, *tokenization*, *stopword removal*, normalisasi menggunakan kamus *colloquial Indonesian lexicon*, dan *stemming* menggunakan algoritma *Sastrawi*.

Representasi Fitur (Feature Representation) setelah proses pra-pemrosesan, setiap teks diubah menjadi bentuk vektor numerik menggunakan metode *TF-IDF (Term Frequency-Inverse Document Frequency)*. Metode ini berfungsi untuk mengekstraksi bobot kata berdasarkan frekuensi kemunculannya dalam dokumen, sehingga kata yang lebih relevan terhadap konteks sentimen memperoleh nilai yang lebih tinggi. Representasi *TF-IDF* ini kemudian digunakan sebagai masukan (*input feature*) bagi algoritma *Naive Bayes* pada tahap pelatihan model.

Pelabelan dan Pelatihan Model tahap pelabelan dilakukan menggunakan *IndoBERT Sentiment*

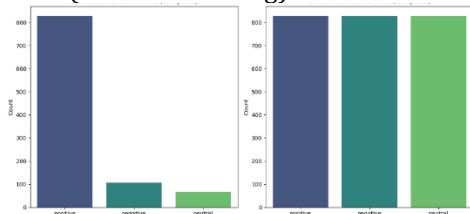
sebagai alat bantu *auto-labeling* untuk sebagian data yang belum berlabel. Hasil labeling ini kemudian digunakan sebagai *training set* untuk melatih model utama *Naïve Bayes*, yang digunakan untuk menentukan polaritas sentimen (positif, negatif, netral).

Validasi dan Evaluasi dilakukan melalui teknik *K-Fold Cross-Validation (Stratified)* dan pengukuran metrik performa seperti akurasi, presisi, *recall*, dan *F1-score* untuk memastikan keandalan model klasifikasi. Penggunaan representasi *TF-IDF* juga memungkinkan evaluasi yang lebih objektif terhadap bobot kata dan pengaruhnya terhadap hasil klasifikasi.

Analisis Korelasi Statistik pengujian hubungan antara skor sentimen hasil model *Naïve Bayes* dengan tingkat kepuasan pembeli menggunakan uji *Spearman Rank Correlation*, karena variabel bersifat ordinal dan tidak berdistribusi normal.

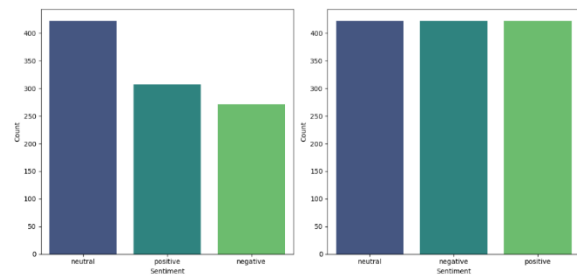
RESULTS AND DISCUSSION

Untuk mengatasi ketidakseimbangan kelas sentimen, dilakukan proses Synthetic Minority Oversampling Technique (SMOTE). SMOTE menghasilkan dataset yang lebih proporsional antar kelas sehingga model dapat belajar secara lebih adil (*balanced training*).



Gambar 5 Penerapan SMOTE (Tokopedia)

Sebelum diterapkan SMOTE, *dataset* Tokopedia menunjukkan ketimpangan distribusi yang cukup tinggi, karena sentimen positif mendominasi sebesar 80,3%, sementara negatif hanya 12,5% dan netral 7,2%. Ketidakseimbangan ini berpotensi membuat model terlalu fokus pada kelas positif dan mengabaikan pola penting pada kelas lainnya. Melalui SMOTE, jumlah data pada kelas negatif dan netral disintesis hingga mencapai jumlah yang sama dengan kelas positif, sehingga seluruh kategori sentimen menjadi seimbang setelah proses *oversampling*.



Gambar 6 Penerapan SMOTE (YouTube)

Sebelum dilakukan SMOTE, distribusi sentimen pada *dataset* YouTube tidak seimbang, yaitu 27,1% negatif, 39,3% netral, dan 33,6% positif, sehingga model cenderung lebih mempelajari pola dari kelas netral yang dominan. Gambar penerapan SMOTE menunjukkan bahwa proses *oversampling* berhasil menyeimbangkan jumlah data pada seluruh kelas dengan menjadikan masing-masing kategori memiliki jumlah sampel yang setara. Penyeimbangan ini dilakukan untuk menghindari bias model terhadap kelas tertentu dan memastikan proses pembelajaran berlangsung lebih adil.

Model *Naïve Bayes* kemudian dioptimasi menggunakan Grid Search dengan Stratified K-Fold ($k=5$) untuk menentukan parameter terbaik (α). Hasil validasi menampilkan tabel parameter dan performa uji silang sebagai berikut:

Tabel hasil Grid Search di atas menampilkan nilai parameter α (α) beserta metrik evaluasi rata-rata dari proses cross-validation sebanyak lima lipatan. Kolom *mean_test_score* menunjukkan nilai akurasi rata-rata pada setiap konfigurasi α , sedangkan *std_test_score* menggambarkan tingkat variasi performa model antar lipatan. Nilai *rank_test_score* digunakan untuk menentukan peringkat performa setiap kombinasi parameter.

Hasil proses optimasi model *Naïve Bayes* pada dataset Tokopedia menggunakan Grid Search dengan pendekatan Stratified K-Fold Cross Validation ($k = 5$) menunjukkan bahwa nilai parameter terbaik diperoleh pada α (α) = 0.1, dengan nilai akurasi rata-rata uji silang sebesar 0.9656. Dengan demikian, pemilihan nilai $\alpha = 0.1$ dianggap optimal karena memberikan keseimbangan terbaik antara kompleksitas model dan kemampuan generalisasi. Sedangkan pada dataset YouTube, hasil optimasi parameter menggunakan Grid Search menunjukkan bahwa nilai α terbaik diperoleh pada angka 0.3. Berdasarkan hasil pengujian dengan pendekatan Stratified K-Fold Cross Validation ($k = 5$), nilai $\alpha = 0.3$ memberikan akurasi rata-rata uji silang (*mean test score*) tertinggi sebesar 0.7535. Meskipun terdapat selisih yang tipis dengan nilai $\alpha = 0.1$ yang berada di peringkat kedua (0.7503), pemilihan $\alpha = 0.3$ terbukti paling efektif dalam menangani karakteristik distribusi data pada komentar YouTube yang lebih bervariasi. Hal ini

mengindikasikan bahwa tingkat smoothing yang sedikit lebih tinggi diperlukan untuk menjaga stabilitas performa klasifikasi dan mencegah overfitting pada dataset tersebut.

Evaluasi performa model dilakukan dengan mengukur nilai Precision, Recall, F1-Score, dan Support terhadap data uji yang telah diberi label otomatis menggunakan model IndoBERT (mdhugol/indonesia-bert-sentiment-classification). Pengujian dilakukan secara terpisah pada dua sumber data, yaitu ulasan Tokopedia dan komentar YouTube GadgetIn, untuk menilai konsistensi performa model terhadap karakteristik teks yang berbeda.

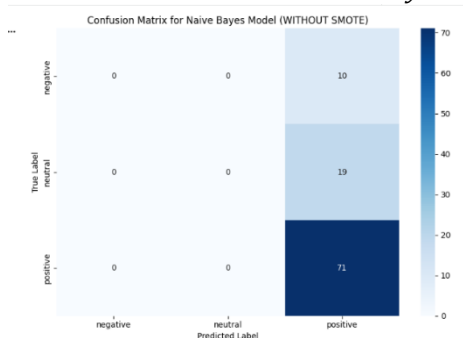
Evaluasi Model Naïve Bayes terhadap Data Tokopedia Pada tahap awal sebelum penerapan SMOTE, ulasan Tokopedia menunjukkan ketidakseimbangan kelas (*class imbalance*) yang signifikan, di mana sentimen positif jauh lebih mendominasi dibandingkan sentimen negatif dan netral. Kondisi ini menyebabkan model cenderung memiliki akurasi global yang tinggi namun memiliki nilai *recall* yang rendah pada kelas minoritas (negatif dan netral), karena model mengalami bias dan lebih sering memprediksi data sebagai kelas mayoritas.

Evaluasi model Naive Bayes TANPA SMOTE (Tokopedia) IndoBERT:

	precision	recall	f1-score	support
negative	0.00	0.00	0.00	10
neutral	0.00	0.00	0.00	19
positive	0.71	1.00	0.83	71
accuracy			0.71	100
macro avg	0.24	0.33	0.28	100
weighted avg	0.50	0.71	0.59	100

Akurasi TANPA SMOTE: 0.71

Gambar 6 Evaluasi Naïve Bayes tanpa



Gambar 7 Confusion Matrix tanpa SMOTE

Setelah dilakukan penyeimbangan data menggunakan SMOTE, hasil evaluasi menunjukkan performa klasifikasi yang sangat baik dengan tingkat akurasi mencapai 97,19%. Berdasarkan hasil *Confusion Matrix*, nilai *precision* dan *recall* tertinggi diperoleh pada kelas *negative* (0.97 dan 1.00), yang menunjukkan kemampuan model dalam mendeteksi ulasan negatif secara konsisten. Kelas *neutral* dan *positive* juga menunjukkan

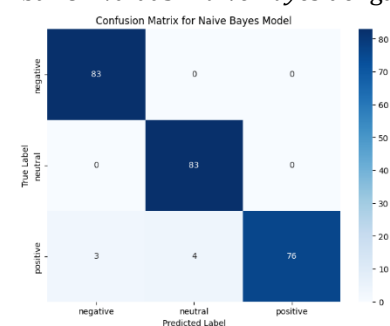
performa tinggi dengan nilai *F1-score* masing-masing sebesar 0.98 dan 0.96, menandakan model memiliki keseimbangan yang baik antara presisi dan sensitivitas setelah distribusi data diseimbangkan.

Evaluasi model Naive Bayes terhadap data (Tokopedia) IndoBERT:

	precision	recall	f1-score	support
negative	0.97	1.00	0.98	83
neutral	0.95	1.00	0.98	83
positive	1.00	0.92	0.96	83
accuracy			0.97	249
macro avg	0.97	0.97	0.97	249
weighted avg	0.97	0.97	0.97	249

Akurasi: 0.9718875502008032

Gambar 8 Evaluasi Naïve Bayes dengan SMOTE



Gambar 10 Confusion Matrix dengan SMOTE

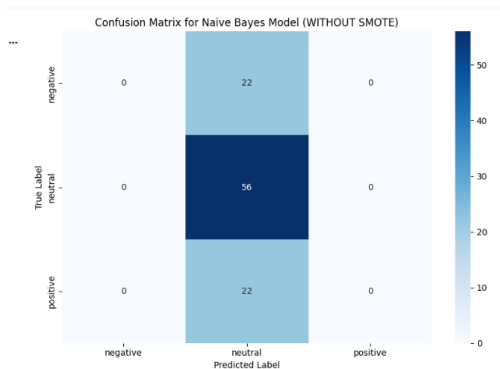
Hasil ini menunjukkan bahwa model Naïve Bayes berbasis fitur TF-IDF dan SMOTE mampu mengklasifikasikan polaritas sentimen secara konsisten dan efisien pada teks ulasan Tokopedia yang umumnya lebih terstruktur dan informatif dibandingkan komentar media sosial.

Evaluasi Model Naïve Bayes terhadap Data YouTube Serupa dengan data Tokopedia, dataset komentar YouTube sebelum SMOTE memiliki sebaran kelas yang tidak merata. Model kesulitan dalam mengidentifikasi sentimen netral dan negatif karena minimnya sampel representatif. Hal ini menyebabkan performa model pada metrik *F1-score* kelas minoritas berada di bawah standar optimal, meskipun akurasi secara keseluruhan terlihat cukup baik akibat dominasi salah satu kelas sentimen.

	precision	recall	f1-score	support
negative	0.00	0.00	0.00	22
neutral	0.56	1.00	0.72	56
positive	0.00	0.00	0.00	22
accuracy			0.56	100
macro avg	0.19	0.33	0.24	100
weighted avg	0.31	0.56	0.40	100

Akurasi TANPA SMOTE: 0.56

Gambar 11 Evaluasi Naïve Bayes tanpa SMOTE

Gambar 12 *Confusion Matrix* tanpa SMOTE

Setelah penerapan SMOTE, hasil evaluasi menunjukkan performa klasifikasi yang baik dengan akurasi sebesar 79,52%. Kelas positive memperoleh nilai recall tertinggi (0,91), menunjukkan bahwa model cukup sensitif dalam mendeteksi komentar dengan polaritas positif. Namun, performa pada kelas neutral relatif lebih rendah dengan F1-score sebesar 0,72, yang dapat disebabkan oleh karakteristik bahasa komentar YouTube yang cenderung informal, singkat, dan sering kali ambigu secara konteks.

CONCLUSION

Berdasarkan hasil penelitian mengenai hubungan antara sentimen komentar pada kanal YouTube GadgetIn dengan tingkat kepuasan pembeli laptop Advan di Tokopedia, dapat disimpulkan beberapa hal penting sebagai berikut:

Model IndoBERT terbukti sangat efektif sebagai instrumen pelabelan otomatis (*auto-labeling*) dalam mengolah dataset besar secara efisien dengan tingkat reliabilitas yang kuat. Berdasarkan pengujian statistik menggunakan koefisien *Cohen's Kappa*, diperoleh nilai sebesar 0,60736 untuk data Tokopedia dan 0,71876 untuk data YouTube. Kedua hasil tersebut masuk ke dalam kategori *Substantial Agreement*, yang mengonfirmasi bahwa hasil pelabelan otomatis memiliki konsistensi yang tinggi dan tingkat kesepakatan yang valid. Hal ini menjamin kualitas *ground truth* yang objektif sebelum memasuki tahap klasifikasi.

Implementasi algoritma *Naïve Bayes* berbasis *TF-IDF* yang diperkuat dengan teknik *Synthetic Minority Over-sampling Technique* (SMOTE) memberikan pengaruh yang sangat signifikan dalam memitigasi bias klasifikasi pada distribusi data yang tidak seimbang. Hal ini dibuktikan dengan lonjakan akurasi yang drastis, di mana pada dataset Tokopedia akurasi meningkat dari 71% menjadi 97,19%, sementara pada dataset YouTube meningkat dari 56% menjadi 79,52%.

Melalui evaluasi metrik *accuracy*, *precision*, *recall*, dan *F1-score*, model menunjukkan performa yang sangat unggul dan stabil, terutama pada dataset Tokopedia dengan nilai *precision* 0,97 dan *recall* 1,00 pada kelas negatif. Meskipun menghadapi karakteristik bahasa media sosial yang informal dan ambigu, model tetap kompetitif pada dataset YouTube dengan capaian *recall* tertinggi sebesar 0,91 pada kelas positif. Secara keseluruhan, penyeimbangan data sintetis melalui SMOTE berhasil memperkuat sensitivitas model dalam mengenali berbagai kelas sentimen secara lebih presisi dan reliabel di kedua platform tersebut.

Distribusi sentimen antar platform menunjukkan perbedaan yang signifikan dalam merepresentasikan persepsi konsumen. Ulasan pembeli di platform Tokopedia didominasi secara oleh sentimen positif (80,3%), yang merefleksikan tingkat kepuasan tinggi pengguna terhadap performa laptop Advan pasca-pembelian. Sebaliknya, komentar pada kanal YouTube GadgetIn menunjukkan persebaran yang jauh lebih berimbang antara sentimen positif (33,6%), netral (39,3%), dan negatif (27,1%). Hal ini mengindikasikan adanya dinamika diskusi yang lebih kritis, skeptis, dan variatif di kalangan audiens video ulasan dibandingkan dengan ulasan langsung di marketplace.

Hasil uji korelasi Spearman Rank antara skor sentimen dan tingkat kepuasan pembeli menghasilkan nilai koefisien $\rho = 0.6646$ dengan $p\text{-value} = 0.0360$. Nilai ini menunjukkan adanya hubungan positif yang kuat dan signifikan secara statistik antara sentimen publik di YouTube dan tingkat kepuasan pembeli di Tokopedia. Dengan demikian, dapat disimpulkan bahwa sentimen positif yang muncul pada komentar video ulasan GadgetIn dapat menjadi indikator yang valid dalam memprediksi tingkat kepuasan pembeli laptop Advan di Tokopedia.

REFERENCE

- A. H. Nasution & A. Onan. (2024). ChatGPT Label: Comparing the Quality of Human-Generated and LLM-Generated Annotations in Low-Resource Language NLP Tasks. *IEEE Access*, 12, 71876–71900. <https://doi.org/10.1109/ACCESS.2024.3402809>
- Adha, S. (2022). FAKTOR REVOLUSI PERILAKU KONSUMEN ERA DIGITAL: SEBUAH TINJAUAN LITERATUR. *JIPIS*, 31(2), 134–148. <https://doi.org/10.33592/jipis.v31i2.3286>
- Aen, K., Tyas, F. A., & Rakhmawati, H. (2025). OPTIMASI ALGORITMA SUPPORT VECTOR

- MACHINE (SVM) MENGGUNAKAN TEKNIK SMOTE UNTUK ANALISIS SENTIMEN PUBLIK TERHADAP DAMPAK PENGGUNAAN ARTIFICIAL INTELLIGENCE (AI). *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(5), 8287–8295.
<https://doi.org/10.33022/ijcs.v13i5.4384>
- Ahmad, A., Ghani, N. A., & Hamid, S. (2025). Examining the Predictors of Consumer Trust and Social Commerce Engagement: A Systematic Literature Review. *Journal of Theoretical and Applied Electronic Commerce Research*, 20(3), 247.
<https://doi.org/10.3390/jtaer20030247>
- Ahmed, C, ElKorany, A, Sayed, & El, & E. (2023). Prediction of customer's perception in social networks by integrating sentiment analysis and machine learning. *Journal of Intelligent Information Systems*.
- Al Ghazali, N. M., & Sibaroni, Y. (2024a). Sentiment classification in e-commerce using Naïve Bayes and combined lexicon-n-gram features. *JIPi (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, 10(2).
<https://doi.org/10.29100/jipi.v10i2.6157>
- Al Ghazali, N. M., & Sibaroni, Y. (2024b). *Sentiment classification in e-commerce using Naïve Bayes and combined lexicon-n-gram features*.
- Al-Kabi, M. N. (2023). Comprehensive Analysis of Naive Bayes Variants for Social Media Sentiment Classification. *Journal of King Saud University - Computer and Information Sciences*, 35(1), 120–135.
- Almansour, A., Alotaibi, R., & Alharbi, H. (2022). Text-rating review discrepancy (TRRD): An integrative review and implications for research. *Future Business Journal*, 8, 3.
<https://doi.org/10.1186/s43093-022-00114-y>
- Alshamari. (2023). Evaluating user satisfaction using deep-learning-based sentiment analysis for social media data in Saudi Arabia's telecommunication sector. *Computers*.
- Amjad, T., Timakum, T., Xie, Q., & Song, M. (2024). Editorial: Text mining-based mental health research. *Frontiers in Research Metrics and Analytics, Volume 9-2024*.
<https://doi.org/10.3389/frma.2024.1388691>
- Anderson, M., & Gómez-Rodríguez, C. (2022). The Impact of Edge Displacement Vaserstein Distance on UD Parsing Performance. *Computational Linguistics*, 48(3), 517–554.
https://doi.org/10.1162/coli_a_00440
- Ardani, W. (2022). Pengaruh Digital Marketing Terhadap Perilaku Konsumen. *Jurnal Tadbir Peradaban*, 2(1), 40–47.
<https://doi.org/10.55182/jtp.v2i1.102>
- Asrul, A. (2024). Evolusi Bisnis Digital: Dampak Teknologi terhadap Perilaku Konsumen dan Pasar Global. *INVESTASI: Inovasi Jurnal Ekonomi Dan Akuntansi*, 2(3), 130–135.
<https://doi.org/10.59696/investasi.v2i3.76>
- Bellar, O., Baina, A., & Ballafkih, M. (2024). Sentiment Analysis: Predicting Product Reviews for E-Commerce Recommendations Using Deep Learning and Transformers. *Mathematics*, 12(15), 2403.
<https://doi.org/10.3390/math12152403>