

Analisis Sentimen Komentar Terhadap Konten Tenun NTT di Youtube Menggunakan Metode SMOTE dan Logistic Regression

Erwianta Gustial Radjah^{1*}, Alfrian Carmen Talakua²

^{1,2}Program Studi Sistem Informasi, Universitas Kristen Wira Wacana Sumba, Jl. R. Suprpto, No.35, Waingapu, Sumba Timur, NTT 87113

*Korespondensi autor: erwiantaradjah@unkriswina.ac.id

ABSTRAK

Penelitian ini bertujuan untuk menganalisis sentimen komentar YouTube terkait konten tenun ikat NTT menggunakan SMOTE dan Logistic Regression. Hasil menunjukkan sentimen positif mendominasi dengan kata kunci seperti "tenun," "kain," dan "bagus," mengindikasikan apresiasi terhadap budaya NTT. Logistic Regression menghasilkan akurasi 91% dengan kinerja sempurna pada sentimen negatif, meskipun terdapat kesalahan klasifikasi pada sentimen netral dan positif akibat tumpang tindih kata kunci. Penggunaan SMOTE efektif menyeimbangkan data sehingga meningkatkan akurasi model. Penelitian ini memberikan wawasan penting tentang persepsi publik untuk mendukung pelestarian dan promosi tenun ikat NTT.

Kata kunci: Sentimen; tenun NTT; SMOTE; Logistic Regression.

ABSTRACT

This study aims to analyze the sentiment of YouTube comments related to NTT ikat weaving content using SMOTE and Logistic Regression. The results show that positive sentiment dominates with keywords such as "tenun," "kain," and "bagus," indicating appreciation for NTT's culture. Logistic Regression achieved 91% accuracy with perfect performance on negative sentiment, although there were classification errors in neutral and positive sentiments due to overlapping keywords. The use of SMOTE effectively balanced the data, improving the model's accuracy. This research provides valuable insights into public perception to support the preservation and promotion of NTT ikat weaving

Keywords: Sentiment; NTT weaving; Smote; Logistic Regression,

PENDAHULUAN

Kebudayaan merupakan pondasi dari identitas bangsa dan bernegara sehingga keberlanjutan budaya juga tidak dapat terpisah dari peran masyarakat sebagai makhluk berbudaya. Tenun Nusa Tenggara Timur (NTT) yang beraneka ragam merupakan salah satu dari warisan budaya di Indonesia. Kain tenun tidak berfungsi hanya sebagai busana harian atau busana adat, namun juga digunakan sebagai simbol penghargaan dalam berbagai jenis kegiatan (Ulfa et al., 2023). NTT memiliki potensi alam yang sangat baik dan berkualitas sehingga berbagai jenis kerajinan lokal memiliki nilai jual yang tinggi, salah satunya adalah kain tenun (Metboki et al., 2024). Tenun tidak hanya dikenal di Nusantara namun juga hingga manca negara dan Indonesia merupakan salah satu negara terbesar yang memproduksi tenun, terkhusus dalam hal kekayaan corak hiasan dari segi ragam hias, warna dan kualitas benang atau bahan yang dipakai (Edie M, 2011) (Deni, 2023). Dr. Viktor Bungtilu Laiskodat, S.H., M.Si, selaku Gubernur NTT periode 2018 – 2023, mewajibkan Aparatur Sipil Negara (ASN) di wilayah pemerintahannya untuk mengenakan busana tenun di setiap hari Rabu dan Jumat. Kebijakan ini diharapkan mendorong penenun NTT agar dapat mengembangkan usaha tenun namun tetap mempertahankan kearifan lokal dan meningkatkan kualitas tenun. Istri dari bapak Gubernur juga membawa dan memperkenalkan tenunan NTT pada kanc

Internasional yakni pada pergelaran tunggal *Couture New York Fashion Week* dan pergelaran mode bergengsi di *Paris Fashion Week* (Selan et al., 2021).

Perkembangan dan persebaran tenun ikat NTT yang semakin luas tentu memberikan pandangan, persepsi, kesan dan sentimen terhadap kain tenun ikat NTT itu sendiri. Pengalaman melihat, mendengarkan, bersentuhan atau memiliki langsung produk kain tenun dapat memberikan kesan yang beragam pada konsumen atau calon konsumen produk kain tenun NTT. Pandangan yang diberikan bisa berupa masukan, opini, saran atau kritik yang memiliki sentimen positif, netral ataupun negatif dan dapat dituangkan dalam bentuk teks (Nurian & Sari, 2023). Sentimen ini dapat dimanfaatkan sebagai bahan kajian, evaluasi ataupun perbaikan bagi produk tenun dalam rangka memberikan kepuasan pada konsumen. Bukti empiris memaparkan bahwa organisasi yang mampu meningkatkan laba dan penjualan adalah organisasi yang berfokus pada kegiatan kepuasan pemakai melalui nilai pengalaman yang kuat (Friska Mastarida, 2023).

Perkembangan teknologi saat ini memungkinkan calon konsumen untuk meningkatkan pengalaman berbelanja konsumen ataupun calon konsumen menjadi lebih baik (Suherlan, 2023). Media sosial merupakan sarana penting untuk berinteraksi dan mengemukakan pendapat mengenai berbagai hal. Sentimen publik menjadi penting karena mampu mempengaruhi keputusan, perilaku dan tindakan kelompok atau individu dalam masyarakat (Kaharudin et al., 2023). Media Sosial Youtube merupakan *platform* video daring yang berfungsi sebagai media untuk melihat, mencari dan berbagi video atau konten dari dan ke seluruh dunia melalui internet (A Hasibuan et al., 2023). Komentar pada konten Youtube dapat dimanfaatkan sebagai sumber data teks dalam bentuk opini, masukan, kritik atau saran guna menganalisis sentimen yang diberikan terhadap konten tertentu. Analisis sentimen bertujuan melakukan klasifikasi teks maupun dokumen yang didalamnya berisi opini yang mewakili maksud negatif, positif dan netral (Maulana et al., 2023).

Menurut studi sebelumnya, berjudul “Analisis sentimen komentar youtube terhadap pemindahan ibu kota negara menggunakan metode *Naïve Bayes*”, penggunaan analisis sentimen pada studi ini bertujuan untuk melihat pandangan masyarakat Indonesia melalui kolom komentar Youtube. Hasil dari studi tersebut menunjukkan bahwa sentimen masyarakat lebih banyak yang netral (Huwaida et al., 2024). Penelitian terdahulu berikutnya, berjudul “Analisis Sentimen Komentar Netizen Terhadap Pembubaran Konser NCT 127 Menggunakan Metode *Naïve Bayes*”, maksud dari penelitian ini yaitu mengetahui seberapa banyak tweets positif dan negatif yang diberikan pengguna twitter. Hasil yang didapatkan yaitu terdapat komentar negatif sebanyak 1.892 komentar dan komentar positif sebanyak 559 komentar (Rizkina & Hasan, 2023). Menurut studi terdahulu selanjutnya, dengan judul “Studi Komparasi Algoritma SVM Dan Random Forest Pada Analisis Sentimen Komentar Youtube BTS”, tujuan penelitian ini yaitu melakukan studi komparasi algoritma SVM dan RF terhadap komentar Youtube BTS. Penelitian ini menunjukkan Nilai akurasi kedua algoritma ditentukan dari hasil perbandingan algoritma yang dievaluasi. Untuk algoritma SVM akurasi data latih sebesar 96% dan akurasi data uji 85%, sedangkan untuk algoritma *Random Forest* akurasi data latih sebesar 82% dan data uji 80%. Hal ini menunjukkan bahwa algoritma SVM memberikan nilai akurasi yang lebih tinggi dibandingkan *Random Forest* dalam analisis sentimen komentar *boy grup* YouTube BTS (Syafia et al., 2023). Berdasarkan 3 studi terdahulu tersebut dibandingkan dengan penelitian saat ini, terdapat perbedaan penerapan metode pada studi terdahulu yaitu penggunaan metode *SMOTE (Synthetic Minority Over-sampling Technique)* dan Logistic Regression dalam memproses data komentar kemudian di analisis dengan analisis sentimen. *SMOTE* merupakan metode yang dapat meningkatkan akurasi klasifikasi pada dataset yang mengalami ketidakseimbangan kelas (Syukron et al., 2023). Regresi logistik merupakan algoritma klasifikasi yang memprediksi probabilitas kelas tertentu, seperti sentimen positif atau negative. Algoritma ini digunakan karena kesederhanaan dan keefektifannya, terutama pada saat

dikombinasikan dengan teknik ekstraksi fitur seperti *TF-IDF* (Tan et al., 2023). Kedua metode ini digunakan untuk menyeimbangkan data positif dan negatif yang tidak seimbang jumlahnya.

Berdasarkan penjabaran latar belakang tersebut, studi ini bertujuan untuk melakukan analisis sentimen terhadap komentar di Youtube pada konten yang berhubungan dengan tenun ikat NTT menggunakan Metode SMOTE dan Logistic Regression sehingga hasilnya nanti dapat menunjukkan sentimen seperti apa yang diberikan pengguna youtube terhadap konten tenun NTT.

METODE PENELITIAN

Secara garis besar penelitian ini menggunakan data kualitatif komentar yang terdapat pada konten tenun NTT di media sosial Youtube. Proses pengumpulan dan pemrosesan data secara keseluruhan menggunakan Bahasa pemrograman python dengan bantuan *Google Collab*. Data kualitatif teks diolah dan dianalisis diubah menjadi data kuantitatif menggunakan metode analisis sentimen. Data yang diambil merupakan data pada bulan Maret hingga bulan Agustus 2024 pada 10 konten teratas terkait tenun ikat. Studi ini secara utuh menggunakan dataset input yang diambil dari komentar pengguna Youtube pada konten – konten yang menampilkan tenun NTT. Tahap proses dimulai dengan menentukan tujuan analisis, yaitu melakukan analisis sentimen komentar-komentar yang berhubungan dengan "Tenun NTT" di platform YouTube. Proses ini menggunakan kombinasi metode *SMOTE* dan *Logistic Regression*, metode – metode ini menyediakan pendekatan yang sederhana namun kuat untuk meningkatkan model analisis sentimen pada data tidak seimbang, seperti pada kasus data komentar YouTube. Kombinasi ini bukan hanya meningkatkan akurasi dan presisi, namun juga membantu menjaga konsistensi hasil di berbagai tipe data analisis sentimen. Analisis data kualitatif diproses melalui 10 tahapan: a). Penentuan tujuan analisis, b). *Scraping Data*, c). *Preprocessing Data*, c). *Labelling Data*, d). *Text Vectorization*, d). *Train-Test Split*, e). *Model Training*, f). Evaluasi Model, g). Kesimpulan.

HASIL DAN PEMBAHASAN

Hasil

Data komentar hasil *scraping* komentar Youtube melalui *google collab* berjumlah 934 komentar. Sampel hasil data komentar dapat dilihat pada Tabel 1.

Tabel 1. Hasil Sampel *Scraping*

No.	Komentar
1.	Bisa Pesan kh
2.	Unik Sekali
3.	Epic
4.	Kak izin download buat tugas kuliah
5.	Woww luar biasa masyarakat adat
..	
929	Sarung bagus.warna keren
930	Rimbo suka berteman
931	Keren sukses selalu ya
933	Mantap
934	Keren

Hasil *preprocessing* membersihkan dan mempersiapkan data awal agar siap digunakan meliputi tokenisasi untuk memecah teks menjadi kata-kata atau token pada Tabel 2 dan *stemming*, yaitu mengubah kata-kata menjadi bentuk dasar pada Tabel 3.

Tabel 2. Hasil 10 Sampel *Tokenize*

No.	Komentar
1.	'bisa', 'pesan', 'kh'
2.	'Unik', 'Sekali'
3.	'Epic'
4.	'Kak', 'izin', 'download', 'buat', 'tugas', 'kuliah'
5.	'Woww', 'luar', 'biasa', 'masyarakat', 'adat'
..	
929	'Sarung', 'bagus', 'warna', 'keren'
930	'Rimbo', 'suka', 'berteman'
931	'Keren', 'sukses', 'selalu', 'ya'
933	'Mantap'
934	'Keren'

Tabel 3. Hasil 10 sampel *stemming*

No.	Komentar
1.	'bisa', 'pesan', 'kh'
2.	'unik', 'sekali'
3.	'epic'
4.	'kak', 'izin', 'download', 'buat', 'tugas', 'kuliah'
5.	'wow', 'luar', 'biasa', 'masyarakat', 'adat'
6.	'hallo', 'kaka', 'aji', 'finalis', 'putra', 'putri', 'kebudayaan'
7.	'keren', 'memperkenalkan', 'budaya'
8.	'ntt', 'kampung', 'mama'
9.	'tenun'
10.	'sumba'

Memberi label pada komentar secara akurat merupakan langkah penting dalam analisis sentimen. Pelabelan data komentar melibatkan penentuan apakah setiap komentar termasuk dalam kategori sentimen tertentu, seperti positif, negatif, atau netral. Proses pelabelan dilakukan secara otomatis menggunakan *lexicon sentistrength* (Musfiroh et al., 2021). *SentiStrength* adalah alat analisis sentimen berdasarkan pendekatan berbasis kosakata (Khaira et al., 2020). Alat ini memungkinkan untuk dapat menganalisis sentimen dengan cepat dengan menilai kekuatan sentimen positif dan negatif dalam teks pendek seperti tweet, postingan media sosial, dan komentar YouTube. Total data positif 592, negatif 62 dan netral 280. Sampel data komentar yang dilabeli dapat dilihat pada Tabel 4.

Tabel 4. Hasil 10 sampel *labelling*

No.	Komentar	Sentimen
1.	pesan	Positif
2.	unik	Positif
3.	epic	Positif
4.	kak izin download tugas kuliah	Negatif
5.	wow masyarakat adat	Netral
6.	bos pesan	Positif
7.	hallo kaka aji final putera puteri budaya wakil	Netral
8.	semangat	Positif
9.	bagus	Positif
10.	kampung mamaku	Netral

Kemudian Setelah preprocessing, ubah komentar menjadi representasi numerik dengan menggunakan *TF-IDF* (*Term Frequency-Inverse Document Frequency*). *TF-IDF* merupakan teknik yang dipakai dalam pemrosesan bahasa alami (NLP) dan penambangan teks untuk merepresentasikan teks dalam bentuk vektor numerik dan pembobotan dokumen (Gifari et al., 2022). Teknik ini digunakan untuk mengevaluasi pentingnya kata-kata dalam suatu dokumen dibandingkan dengan kumpulan dokumen (Fransiska et al., 2020).

Tabel 5. Hasil 10 Besar *TF-IDF*

Dokumen	Kata	Bobot TF-IDF
1.	pesan	0.6788569722398622
2.	unik	1.0
3.	epic	1.0
4.	kak izin download tugas kuliah	0.5022749138365066
5.	wow masyarakat adat	0.4808323436044952
6.	bos pesan	0.46420019398843826
7.	hallo kaka aji final putera puteri budaya wakil	0.450610738263771
8.	semangat	0.31303084630242745
9.	bagus	0.6600145168554687
10.	kampung mamaku	0.5524124793296166

Kata dengan bobot TF-IDF tertinggi yaitu “unik” dan “epic” memiliki bobot TF-IDF 1.0, menunjukkan kata ini hanya tampil dalam satu komentar tertentu dan tidak ada dalam komentar lainnya. Hal ini menunjukkan kata ini sangat spesifik untuk komentar tersebut dan kemungkinan memberikan informasi unik yang kuat tentang komentar tersebut.

Kata dengan bobot sedang yaitu kata seperti “pesan” (0.6788) dan “bagus” (0.6600) mempunyai bobot TF-IDF yang cukup besar, menandakan bahwa kata-kata ini muncul di beberapa dokumen tetapi masih cukup unik untuk membedakannya. Komentar “kak izin download tugas kuliah” (0.5022) memiliki relevansi yang tinggi untuk komentar tertentu tetapi masih memiliki distribusi yang lebih luas dibandingkan kata-kata unik lainnya. Kata dengan bobot relatif rendah yaitu kata seperti “semangat” (0.3130) dan “bos pesan” (0.4642) mempunyai bobot TF-IDF yang lebih rendah diantara kata lainnya. Meskipun masih relevan untuk komentar tertentu, kata-kata ini mungkin lebih umum dibandingkan kata lain pada hasil ini.

Pada proses pelabelan dapat dilihat komentar positif, netral dan negatif tampak tidak seimbang sehingga dibutuhkan proses penyeimbangan data tersebut. Metode *SMOTE* digunakan untuk menyeimbangkan data komentar yang ada karena ketidakseimbangan antara jumlah komentar positif dan negative serta netral. Hasil pemrosesan menggunakan *SMOTE* dilihat pada Tabel 6.

Tabel 6. Hasil 5 Besar *SMOTE*

Data ke	Kata	Label
1.	(0, 956) 0.7071067811865476 (0, 1103)0.7071067811865476	Netral
2.	(0, 166) 0.3381662247711262 (0, 1657)0.4262762557993248 (0, 873) 0.2765404878280739 (0, 1019)0.5105992034730186 (0, 184) 0.6055955499921	Negatif
3.	(0, 1903)0.1605191005371242 (0, 942) 0.3091862063377313 (0, 1643)0.20372838192780637 (0, 1578)0.2536788904701505 (0, 902) 0.3603925977711236 (0, 425)0.3991287375289788 (0, 1388)0.3603925977711236 (0, 1899)0.3603925977711236 (0, 1949)0.28652701715669204 (0, 975)0.37646954834793	Positif
4.	(0,511) 0.32066950282326345 (0,251) 0.25086997302422 (0,2007)0.3069754662909459 (0,1046)0.2683741453557043 (0,53)0.3399701632908842 (0,1111)0.32066950282326345 (0,1634)0.3399701632908842 (0,1274)0.3399701632908842 (0,1580)0.3399701632908842 (0,1059)0.3206695028232634	Positif
5.	(0,1528)0.5548709610574433 (0,1033)0.588267888200261 (0,2141)0.58826788820026	Positif

Hasil performa pembobotan TF – IDF:

1. Komentar Netral

Kata-kata dengan skor TF-IDF tinggi pada komentar netral (informasi data 1) mempunyai nilai yang hampir mirip seperti (0, 956) dan (0, 1103), masing-masing dengan skor 0.707. Ini menandakan bahwa komentar ini kemungkinan menggunakan kata-kata yang tidak mengandung makna emosional yang kuat, sehingga menghasilkan sentimen netral.

2. Komentar Negatif

Pada komentar negatif (informasi data 2), bobot tertinggi (0.605) yaitu dari (0, 184), diikuti oleh (0, 1019) dengan 0.510. Ini mengindikasikan bahwa komentar - komentar ini sangat penting dalam menentukan sentimen negatif. Kata-kata ini mungkin mengacu pada aspek yang mengekspresikan ketidakpuasan atau keluhan.

3. Komentar Positif

Komentar positif menghasilkan pola dengan beberapa komentar memiliki bobot *TF-IDF* tinggi, seperti pada data 3, kata (0, 975) memiliki skor tertinggi (0.376), menunjukkan elemen penting yang berperan pada sentimen positif. Data 4, komentar dengan skor tertinggi (0, 1274) memiliki nilai 0.339. Ini memperlihatkan komentar ini kaya dengan kata - kata yang menunjukkan sisi positif atau apresiasi. Data 5, kata (0, 1033) dan (0, 2141) mempunyai bobot tertinggi, yaitu 0.588, mengindikasikan apresiasi yang sangat kuat dalam komentar ini.

Pada tahap berikutnya yaitu penggunaan *SMOTE* untuk menangani ketidakseimbangan data dengan menambahkan data sintetis pada kelas minoritas, Sentimen netral dan negatif terlihat kurang terwakili dibandingkan dengan data sentimen positif. Penggunaan *SMOTE* menyeimbangkan distribusi kelas ini, sehingga model dapat lebih baik dalam memprediksi sentimen.

Sebelum masuk pada analisis *logistic regression* pada hasil *SMOTE*, data dibagi menjadi dua bagian yaitu data latih sebanyak 80% dan data uji sebanyak 20%. Hasil dan evaluasi performa Model *logistic regression* ditunjukkan pada table 7 dan pada tabel 8.

Tabel 7. Hasil Klasifikasi *Logistic Regression*

	precision	recall	f1-score	support
Negatif	1.00	0.99	1.00	113
Netral	0.83	0.94	0.88	124
Positif	0.93	0.80	0.86	117
accuracy			0.91	354
macro avg	0.92	0.91	0.91	354
weighted avg	0.92	0.91	0.91	354

Hasil Kinerja Model:

1. Negatif, dengan *precision* 1.00 mengindikasikan bahwa semua prediksi negatif benar (tidak ada *False Positive*). *Recall* dengan nilai 0.99 menandakan hampir semua data negatif terprediksi dengan benar, dengan hanya satu data aktual negatif yang salah diklasifikasikan. *F1-Score* dengan skor 1.00 menghasilkan kinerja sempurna pada kelas negatif.
2. Netral, dengan *precision* bernilai 0.83, menunjukkan bahwa beberapa prediksi kelas netral salah diklasifikasikan sebagai kelas lain (*False Positive* cukup signifikan). *Recall* bernilai 0.94 menunjukkan sebagian besar data aktual netral terklasifikasi dengan benar, hanya sebagian kecil yang salah. *F1-Score*: 0.88, menandakan kinerja yang baik tetapi ada kemungkinan untuk perbaikan pada pengenalan kelas netral.
3. Positif, dengan *precision*: 0.93 menunjukkan sebagian besar hasil prediksi positif akurat, tetapi masih ada beberapa *False Positive*. *Recall* dengan hasil 0.80 menunjukkan bahwa ada data aktual positif yang salah diklasifikasikan sebagai kelas lain (*False Negative* cukup besar). *F1-Score* dengan nilai 0.86, menandakan bahwa kinerja untuk kelas positif sedikit lebih rendah dibandingkan kelas lainnya.

4. Akurasi Keseluruhan bernilai 0.91 (91%), menunjukkan bahwa model cukup baik untuk klasifikasi sentimen secara umum.
5. *Macro Average* dan *Weighted Average*, mengindikasikan semua metrik berada di kisaran 0.91–0.92, menghasilkan model dengan kinerja yang stabil di seluruh kelas, tanpa terlalu bias terhadap kelas tertentu.

Tabel 8. Hasil Confusion Matrix

	Prediksi Negatif	Prediksi Netral	Prediksi Positif
Aktual Negatif	112	1	0
Aktual Netral	0	117	7
Aktual Positif	0	23	94

Kelas Negatif, dari 113 data aktual negatif, 112 terprediksi dengan baik sebagai negatif, dengan hanya 1 data yang ekliiru diklasifikasikan sebagai netral. Tidak terdapat data yang salah diklasifikasikan sebagai positif. Kemudian pada kelas netral dari 124 data aktual netral, 117 terprediksi dengan baik sebagai netral. Namun, terdapat 7 data netral keliru diklasifikasikan sebagai positif. Hal ini memperlihatkan tantangan dalam mengklasifikasi komentar netral dan positif, kemungkinan karena ada *overlap* kata-kata kunci antara kedua kelas ini. Terakhir pada kelas positif, dari 117 data aktual positif, 94 terkonfirmasi dengan benar sebagai positif, namun dengan adanya 23 data positif salah diklasifikasikan sebagai netral. Kesalahan ini menandakan bahwa beberapa komentar positif memiliki elemen yang menyerupai komentar netral, sehingga sulit diklasifikasikan oleh model. Frekuensi kata positif, netral dan negative setelah melalui proses pada logistic regression terlihat pada Tabel 9.

Tabel 9. Hasil 10 Besar Kata Positif, Negatif dan Netral.

Kata Positif	Frekuensi	Kata Negatif	Frekuensi	Kata Netral
tenun	129	kain	13	keren
kain	110	kak	8	ntt
ntt	102	sabu	8	kain
bagus	73	you	8	tenun
mantap	71	to	7	sabu
nya	61	tenun	7	nya
yang	57	tv	7	yg
cantik	56	in	7	pulau
sabu	53	its	6	sumba
sukses	53	nah	5	bangga

Kata positif yang paling sering muncul yaitu "tenun" (129) dan "kain" (110) adalah kata yang paling sering muncul dalam sentimen positif. Kata-kata ini menandakan fokus utama pada apresiasi terhadap budaya atau produk lokal (seperti kain tenun). Kata-kata seperti "ntt" (102), "bagus" (73), dan "mantap" (71) juga mencerminkan dukungan atau pujian terhadap wilayah NTT dan aspek-aspek terkait seperti karya seni atau budaya. Komentar seperti "cantik" (56) dan "sukses" (53) menandakan apresiasi pribadi atau penghargaan terhadap hasil kerja atau prestasi tertentu. Kata "nya" (61) dapat diindikasikan penggunaan kata sifat atau deskripsi yang sering dikaitkan dengan hal-hal positif.

Kata yang Negatif dominan, yaitu "kain" (13) adalah kata yang juga terlihat di sentimen negatif. Namun, kemunculannya lebih sedikit dibandingkan dengan sentimen positif, menunjukkan konteks yang berbeda, kemungkinan keluhan atau kritik terhadap kain tertentu. Kata "kak" (8), "sabu" (8), dan "you" (8) cenderung berkonotasi negatif dalam konteks ini, kemungkinan terkait dengan pengalaman yang kurang memuaskan. Makna kontekstual, kata seperti "to" (7) dan "nah" (6) menunjukkan adanya penekanan atau ekspresi frustrasi, tetapi tidak secara langsung memberikan makna sentimen negatif yang kuat.

Kata dominan seperti "keren" (8), "ntt" (8), dan "kain" (7) adalah kata-kata yang paling sering terlihat dalam sentimen netral. Kata-kata ini mengindikasikan deskripsi atau narasi tanpa emosi yang kuat. Kata seperti "pulau" (5), "sumba" (5), dan "bangga" (5) menunjukkan fokus pada informasi tentang lokasi atau budaya tanpa indikasi sentimen emosional yang berarti. Makna kontekstual dari kata-kata seperti "yg" (6) atau "nya" (5) cenderung menjadi bagian dari narasi deskriptif atau informatif.

Pembahasan

Pentingnya kata kunci pada kata-kata dengan bobot TF-IDF terbesar seperti “unik,” “epic,” dan “pesan” berkontribusi secara besar dalam membentuk topik utama dari komentar. Kata-kata ini mencerminkan elemen yang paling relevan dan khas dalam komentar. Kata-kata positif seperti “bagus” dan “semangat” menandakan jika sebagian besar komentar cenderung mengekspresikan sentimen positif. Namun, kata-kata netral seperti “pesan” tetap dominan, menunjukkan pentingnya memahami konteks kata secara lebih dalam untuk menghindari kesalahan klasifikasi.

SMOTE menunjukkan hasil data sintesis untuk kelas minoritas, yaitu kelas negatif dan netral, untuk menyeimbangkan dataset yang kemudian mempengaruhi kelas netral dan negatif mendapatkan peningkatan representasi, memungkinkan model klasifikasi untuk lebih baik memahami pola kata yang relevan untuk sentimen tersebut. Penyeimbangan ini perlu diperhatikan mengingat hasil sebelumnya menunjukkan bias terhadap kelas positif akibat distribusi data yang tidak merata. *Logistic Regression* menunjukkan kinerja yang sangat baik untuk klasifikasi sentimen, dengan akurasi keseluruhan 91%. Model ini sangat baik dalam mengenali komentar negatif, dengan *precision* dan *recall* sempurna. Kesalahan klasifikasi pada kelas netral dan positif disebabkan oleh tumpang tindih kata kunci yang sering muncul di kedua kelas tersebut. Kata-kata seperti “semangat” mungkin digunakan dalam konteks yang lebih ambigu, sehingga sulit diklasifikasikan dengan tepat. Dengan implementasi SMOTE, hasil analisis sentimen dapat lebih baik, akurat dan seimbang, terutama untuk kelas yang sebelumnya kurang terwakili. Kata-kata kunci yang relevan tetap berperan penting, dan teknik augmentasi ini memungkinkan model untuk lebih memahami konteks yang lebih luas dari setiap sentimen.

Pandangan pengguna YouTube terhadap konten terkait tenun ikat NTT sangat berpengaruh dalam mempengaruhi persepsi dan keputusan pembelian produk. Sentimen positif yang lebih banyak dalam komentar menandakan adanya apresiasi yang tinggi terhadap kualitas dan keindahan tenun ikat NTT, yang berpotensi mendukung pemasaran produk ini. Ketika audiens merespons secara positif melalui media sosial seperti YouTube, hal ini dapat meningkatkan citra merek dan menguatkan minat konsumen, baik lokal maupun internasional.

Pemasaran yang memanfaatkan analisis sentimen dapat mengidentifikasi kata kunci yang resonan dengan penikmat konten, seperti "bagus," "unik," dan "keren," yang dominan dalam komentar positif. Dengan memahami sentimen ini, produsen atau pemasar tenun ikat dapat menyusun strategi promosi yang lebih terarah, memanfaatkan kata-kata yang berkonotasi positif untuk meningkatkan minat beli dan membangun hubungan yang lebih kuat dengan calon konsumen.

Selain itu, pemahaman terhadap sentimen negatif atau netral juga penting untuk memperbaiki dari sisi produk dan layanan, karena kritik atau ketidakpuasan yang disampaikan pengguna dapat menjadi bahan evaluasi untuk mengembangkan kualitas produk atau pengalaman berbelanja. Secara keseluruhan, analisis sentimen di platform seperti YouTube memberi pandangan yang sangat berharga untuk strategi pemasaran yang lebih efektif, memanfaatkan umpan balik audiens dalam bentuk sentiment untuk meningkatkan daya tarik dan keberhasilan produk tenun ikat NTT di pasar.

KESIMPULAN

Analisis sentimen pada komentar Youtube yang berhubungan dengan tenun ikat NTT memberikan gambaran bahwa konten ini disambut dengan sangat positif oleh audiens. Metode SMOTE dan Logistic Regression terbukti efektif dalam menganalisis sentimen, dengan hasil akurasi yang tinggi dan kemampuan untuk menangkap sentimen negatif dan netral yang sebelumnya kurang terwakili. Namun, untuk meningkatkan akurasi lebih lanjut, metode berbasis konteks seperti Word2Vec atau BERT dapat digunakan untuk menangkap hubungan semantik antar kata. Analisis lebih dalam pada kata-kata ambigu seperti "kain" dan "ntt" diperlukan untuk membedakan konteks sentimen netral dan positif. Penelitian ini menunjukkan potensi besar untuk menggunakan analisis sentimen dalam memahami respon audiens terhadap budaya lokal, khususnya tenun ikat NTT, dan memberikan wawasan yang berguna untuk pelestarian serta promosi budaya.

Persepsi pengguna YouTube terhadap konten tenun ikat NTT dapat secara signifikan mempengaruhi pemasaran produk tersebut. Sentimen positif dalam komentar memperkuat gambaran produk dan menarik minat pembeli, sementara kritik atau sentimen negatif memberi peluang untuk perbaikan yang lebih terarah. Dengan memanfaatkan analisis sentimen, pemasar dapat merancang strategi promosi yang lebih efektif, menyoroti aspek yang dihargai penikmat konten dan memperbaiki kekurangan untuk meningkatkan dan memperkuat daya tarik produk di pasar.

DAFTAR PUSTAKA

- A Hasibuan, D. R., Angginami, H., Hardani Ritonga, I., Al-Rasyid Saragih, R., & Ilmu Perpustakaan, P. (2023). Pemanfaatan Media Sosial Youtube sebagai Media Edukasi di Kalangan Milenial. In *Sci-Tech Journal* (Vol. 2, Issue 2).
- Deni, G. R. (2023). Perancangan Perlengkapan Busana Berbahan Dasar Tenun Tradisional Sumbawa. *JISIP (Jurnal Ilmu Sosial Dan Pendidikan)*, 7(1), 2598–9944. <https://doi.org/10.58258/jisip.v7i1.4595>
- Edie M, T. (2011). *Tenun Ikat dan Songket*. Pelita Hati.
- Fransiska, S., Rianto, R., & Gufroni, A. I. (2020). Sentiment Analysis Provider By.U on Google Play Store Reviews with TF-IDF and Support Vector Machine (SVM) Method. *Scientific Journal of Informatics*, 7(2), 203–212. <https://journal.unnes.ac.id/nju/sji/article/view/25596>
- Friska Masterida. (2023). Hubungan Kualitas Layanan, Pengalaman Konsumen, Kepuasan Konsumen, dan Loyalitas Konsumen: Model Konseptual. *ARBITRASE: Journal of Economics and Accounting*, 3(3), 521–526. <https://doi.org/10.47065/arbitrase.v3i3.702>
- Gifari, O. I., Adha, M., Freddy, F., & Durrand, F. F. S. (2022). Film Review Sentiment Analysis Using TF-IDF and Support Vector Machine. *Journal of Information Technology*, 2(1), 36–40.
- Huwaida, S. F., Kusumawati, R., & Isnaini, B. (2024). Analisis Sentimen Komentar YouTube terhadap Pemandangan Ibu Kota Negara Menggunakan Metode Naïve Bayes. *Jambura Journal of Informatics*,

- 6(1), 26–39. <https://doi.org/10.37905/jji.v6i1.24718>
- Kaharudin, A., Agus Supriyadi, A., Baitika, H., & Derryanur, M. (2023). *OKTAL : Jurnal Ilmu Komputer dan Science Analisis Sentimen pada Media Sosial dengan Teknik Kecerdasan Buatan Naïve Bayes: Kajian Literatur Review*. 2(6). <https://harzing.com/resources/publish-or-perish>
- Khaira, U., Johanda, R., Utomo, P. E. P., & Suratno, T. (2020). Sentiment Analysis Of Cyberbullying On Twitter Using SentiStrength. *Indonesian Journal of Artificial Intelligence and Data Mining*, 3(1), 21. <https://doi.org/10.24014/ijaidm.v3i1.9145>
- Maulana, A., Inayah Khasnaputri Afifah, Asghafi Mubarrak, Kiagus Rachmat Fauzan, Ardhan Dwintara, & Zen, B. P. (2023). COMPARISON OF LOGISTIC REGRESSION, MULTINOMIALNB, SVM, AND K-NN METHODS ON SENTIMENT ANALYSIS OF GOJEK APP REVIEWS ON THE GOOGLE PLAY STORE. *Jurnal Teknik Informatika (Jutif)*, 4(6), 1487–1494. <https://doi.org/10.52436/1.jutif.2023.4.6.863>
- Metboki, G., Nababan, D., Kelen, Y. P., Studi Teknologi Informasi Fakultas Pertanian Sains Dan Kesehatan, P., & Negeri Timor, U. (2024). *E-UMKM Studi Kasus Penjualan Kain Tenun Berbasis Android dengan Metode Waterfall*. 4(1), 24–29. <https://doi.org/10.54259/satesi.v4i1.2729>
- Musfiroh, D., Khaira, U., Eko, P., Utomo, P., Suratno, T., Studi, P., Informasi, S., Sains, F., & Teknologi, D. (2021). MALCOM: Indonesian Journal of Machine Learning and Computer Science Sentiment Analysis of Online Lectures in Indonesia from Twitter Dataset Using InSet Lexicon Analisis Sentimen terhadap Perkuliahan Daring di Indonesia dari Twitter Dataset Menggunakan InS. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 1, 24–33.
- Nurian, A., & Sari, B. N. (2023). ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI GOOGLE PLAY MENGGUNAKAN NAÏVE BAYES. *Jurnal Informatika Dan Teknik Elektro Terapan*, 11(3s1), 2830–7062. <https://doi.org/10.23960/jitet.v11i3s1.3348>
- Rizkina, N. Q., & Hasan, F. N. (2023). Analisis Sentimen Komentar Netizen Terhadap Pembubaran Konser NCT 127 Menggunakan Metode Naive Bayes. *Journal of Information System Research (JOSH)*, 4(4), 1136–1144. <https://doi.org/10.47065/josh.v4i4.3803>
- Selan, R., Kale, A. K. A., #3, J. S. B., & Tarigan, B. V. (2021). *Pelatihan Manajemen Usaha Kelompok Pengrajin Tenun Ikat Khas NTT*.
- Suherlan, S. A. (2023). Analysis Of Perceived Ease Of Use And Trust To Purchase Intention Mediated By Attitude To Shop Online. In *Management Studies and Entrepreneurship Journal* (Vol. 4, Issue 6). <http://journal.yrpiaku.com/index.php/msej>
- Syafia, A. N., Hidayattullah, M. F., & Suteddy, W. (2023). *Studi Komparasi Algoritma SVM Dan Random Forest Pada Analisis Sentimen Komentar Youtube BTS*. 8(3).
- Syukron, A., Sardiarinto, S., Saputro, E., & Widodo, P. (2023). Penerapan Metode Smote Untuk Mengatasi Ketidakseimbangan Kelas Pada Prediksi Gagal Jantung. *Jurnal Teknologi Informasi Dan Terapan*, 10(1), 47–50. <https://doi.org/10.25047/jtit.v10i1.313>
- Tan, K. L., Lee, C. P., & Lim, K. M. (2023). A Survey of Sentiment Analysis: Approaches, Datasets, and Future Research. *Applied Sciences (Switzerland)*, 13(7). <https://doi.org/10.3390/app13074550>
- Ulfa, S., br Sinulingga, T. E., & Sinulingga, J. (2023). Kain Tenun Tradisional: Warisan Budaya dan Industri Kreatif. *Jurnal Pendidikan Tambusai*, 7(3), 29709–29715. <https://www.jptam.org/index.php/jptam/article/view/11780>