

Analisis Sentimen Kepuasan Pelanggan Terhadap Warunk Upnormal Berdasarkan Ulasan Google Maps Menggunakan IndoBERT Dan Topic Modeling

Razika Raehan Al Aziz¹, Herry Irawan², Yusza Reditya Murti³

¹ Manajemen Bisnis Telekomunikasi dan Informatika, Fakultas Ekonomi dan Bisnis, Universitas Telkom, Indonesia, razikaraehan@student.telkomuniversity.ac.id

² Manajemen Bisnis Telekomunikasi dan Informatika, Fakultas Ekonomi dan Bisnis, Universitas Telkom, Indonesia, herryir@telkomuniversity.ac.id

³ Manajemen Bisnis Telekomunikasi dan Informatika, Fakultas Ekonomi dan Bisnis, Universitas Telkom, Indonesia, yuszaa@telkomuniversity.ac.id

Abstract

Indonesia's culinary industry is growing rapidly, especially among the younger generation. Warunk Upnormal, as one of the popular restaurants, has shown stable performance despite Google Trends indicating a decline in search interest since late 2020. This study aims to analyze customer sentiment and identify the main topics discussed in reviews related to Warunk Upnormal. The methods used include sentiment analysis with the IndoBERT model and topic modeling using BERTopic. A total of 21,917 reviews were collected from Google Maps through web scraping and classified into positive, negative, and neutral categories. The results show that IndoBERT achieved a high accuracy of 94.42%. Of all the reviews, 75.53% were positive and 24.47% were negative. Positive sentiment reflects customer satisfaction with aspects such as convenience, food taste and variety, pricing, service quality, and social influence. In contrast, negative reviews highlight issues such as poor Wi-Fi connection, mismatch between price and quality, declining food taste, slow service, and inconsistent cleanliness. These findings indicate a gap between customer expectations and service reality, particularly at the Bandung branch. The insights from this study are expected to help management develop strategies to improve service quality and enhance customer satisfaction.

Keywords-sentiment analysis, BERTopic, Google Maps, IndoBERT, customer satisfaction

Abstrak

Industri kuliner Indonesia berkembang pesat, terutama di kalangan generasi muda. Warunk Upnormal, sebagai salah satu restoran populer, menunjukkan performa stabil meskipun data Google Trends mencatat penurunan minat pencarian sejak akhir 2020. Penelitian ini bertujuan menganalisis sentimen pelanggan dan mengidentifikasi topik utama dalam ulasan terkait Warunk Upnormal. Metode yang digunakan mencakup analisis sentimen dengan model IndoBERT dan pemodelan topik menggunakan BERTopic. Sebanyak 21.917 ulasan dari Google Maps dikumpulkan melalui web scraping dan diklasifikasikan ke dalam kategori positif, negatif, dan netral. Hasil menunjukkan bahwa IndoBERT memiliki akurasi tinggi sebesar 94,42%. Dari total ulasan, 75,53% tergolong positif dan 24,47% negatif. Sentimen positif mencerminkan kepuasan pelanggan terhadap aspek kenyamanan tempat, rasa dan variasi makanan, harga, kualitas pelayanan, dan pengaruh sosial. Sebaliknya, ulasan negatif menyoroti masalah seperti koneksi Wi-Fi yang buruk, ketidaksesuaian harga dan kualitas, penurunan cita rasa, pelayanan lambat, serta kebersihan yang tidak konsisten. Temuan ini mengindikasikan adanya kesenjangan antara harapan dan kenyataan layanan, terutama di cabang Bandung. Hasil studi ini diharapkan membantu manajemen dalam merancang strategi peningkatan layanan dan kepuasan pelanggan.

Kata Kunci-analisis sentimen, BERTopic, Google Maps, IndoBERT, kepuasan pelanggan

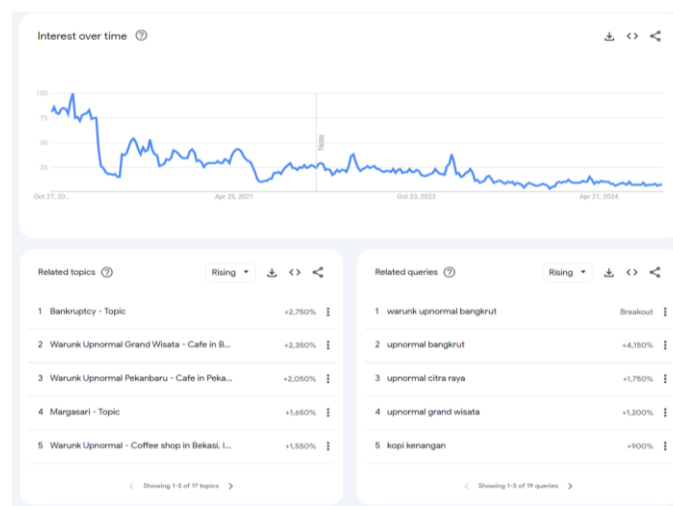
I. PENDAHULUAN

Pertumbuhan teknologi informasi dan komunikasi telah mendorong perubahan besar dalam perilaku konsumen, khususnya dalam memanfaatkan internet untuk mencari informasi dan memberikan umpan balik terhadap produk atau layanan. Berdasarkan data laporan We Are Social, penggunaan internet di Indonesia tercatat sebanyak 185,3 juta atau sekitar 66,5% dari populasi, menunjukkan betapa dominannya internet dalam kehidupan sehari-hari. Salah satu platform digital yang aktif digunakan dalam konteks konsumsi dan layanan adalah Google Maps, yang memungkinkan pengguna memberikan penilaian dan ulasan terhadap tempat yang dikunjungi, termasuk restoran (Rifa et al., 2021).

Ulasan pelanggan dalam platform digital menjadi salah satu sumber data penting dalam mengevaluasi persepsi

publik terhadap layanan bisnis. Dalam konteks industri makanan dan minuman yang terus mengalami pertumbuhan, memahami ulasan pelanggan menjadi krusial untuk mempertahankan daya saing. Warunk Upnormal merupakan salah satu pelaku bisnis kuliner yang sempat populer dengan konsep kafe modern dan menu khas Indonesia. Namun, berdasarkan data *Google Trends*, minat pencarian terhadap Warunk Upnormal mengalami penurunan signifikan sejak tahun 2021 hingga 2024. Topik pencarian bahkan mengindikasikan kemungkinan kebangkrutan, yang bertolak belakang dengan tren positif pertumbuhan industri kuliner secara umum. Menurut (Bennix, 2023), berbagai faktor memengaruhi penurunan minat pelanggan terhadap Warunk Upnormal, mulai dari harga yang tidak sebanding dengan kualitas makanan, ekspansi bisnis yang terlalu cepat, hingga pendekatan manajerial yang kurang adaptif. Kondisi ini menunjukkan adanya kesenjangan antara ekspektasi dan pengalaman pelanggan, yang perlu dianalisis lebih lanjut melalui pendekatan berbasis data.

Di era digital, data ulasan pelanggan dalam bentuk teks dapat dianalisis menggunakan teknik analisis sentimen, yakni metode untuk mengidentifikasi dan mengklasifikasikan opini publik ke dalam kategori positif, negatif, atau netral. Analisis ini menyajikan pemahaman yang lebih dalam mengenai persepsi konsumen dan dimensi-dimensi yang memengaruhi kepuasan mereka. Menurut (Mehta & Pandya, 2020), analisis sentimen memanfaatkan data dari media sosial, blog, maupun ulasan daring untuk menggambarkan reaksi emosional publik.



Gambar 1 Grafik *Google Trend* Warunk Upnormal (*trends.google.com*)

Berdasarkan data fenomena pada Gambar 1, terlihat bahwa Warunk Upnormal mengalami penurunan minat pencarian dari tahun 2020 hingga kini, dengan topik-topik pencarian yang mengarah pada kebangkrutan. Di sisi lain, data industri menunjukkan bahwa sektor makanan dan minuman secara keseluruhan mengalami pertumbuhan yang stabil. Perbedaan ini menunjukkan urgensi untuk mengevaluasi kepuasan pelanggan Warunk Upnormal melalui pendekatan analisis sentimen dan *topic modeling* terhadap ulasan pengguna di *Google Maps*. Analisis ini penting untuk memberikan wawasan yang lebih komprehensif mengenai persepsi pelanggan serta menjadi dasar evaluasi layanan dan pengambilan keputusan strategis.

Penelitian sebelumnya, yang dilakukan oleh (Alfarobby & Irawan, 2024), memiliki tujuan untuk menganalisis sentimen pengguna terhadap layanan Gojek dan Grab di *Google Play Store*. Penelitian tersebut menerapkan metode kuantitatif dengan dua pendekatan analisis, yaitu analisis sentimen menggunakan model *IndoBERT* dan analisis *topic modeling*. Hasil analisis menunjukkan bahwa sentimen positif Gojek mencapai 60,8%, sementara Grab 79,5%. Analisis topik mengungkapkan bahwa pengguna Gojek lebih banyak mengeluhkan kualitas layanan, sementara pengguna Grab fokus pada masalah harga dan ketepatan titik penjemputan. Penelitian ini memberikan wawasan bagi perusahaan dalam meningkatkan kualitas layanan dan kepuasan pelanggan. Dengan demikian, penelitian ini menghadirkan pemahaman yang mendalam dan terstruktur mengenai persepsi masyarakat terhadap kedua *platform* melalui penerapan pendekatan kuantitatif yang metodologis dan terukur.

Ulasan atau umpan balik sangat penting bagi bisnis dan juga pemerintah karena mereka dapat memperoleh informasi berharga mengenai wawasan terhadap produk dan layanan (Irawan et al., 2019). Penelitian akan dilakukan dengan menggunakan *IndoBERT* pada analisis sentimen serta *topic modeling* akan dilakukan dengan menggunakan *BerTopic* untuk mengambil informasi dari hasil analisis sentimen.

Berdasarkan latar belakang yang telah dipaparkan, penelitian akan dilakukan lebih lanjut dengan judul “Analisis Sentimen Kepuasan Pelanggan Terhadap Warunk Upnormal Berdasarkan *Google Maps Review* Menggunakan *IndoBERT* Dan *Topic Modeling*”.

II. TINJAUAN LITERATUR

A. Kepuasan Pelanggan

Kepuasan pelanggan merupakan kondisi emosional, baik berupa rasa senang maupun kecewa, yang timbul ketika seseorang menilai kesesuaian antara persepsi kinerja atau hasil produk dengan ekspektasi. (Tjiptono & Diana, 2022).



Gambar 2 Dimensi Kepuasan Pelanggan (Rezende et al., 2024)

Berdasarkan Gambar 2, (Rezende et al., 2024) mengidentifikasi empat dimensi utama kepuasan pelanggan dalam industri makanan, yakni *service quality* (kecepatan dan keandalan layanan), *price* (kesesuaian harga dengan kualitas), *overall food quality* (rasa, penyajian, dan variasi menu), serta *culture/social influence* (pengaruh nilai budaya dan sosial terhadap persepsi pelanggan).

B. Analisis Sentimen

Analisis sentimen memiliki peranan penting dalam berbagai studi yang dilakukan oleh sejumlah peneliti (Masrury et al., 2019). Analisis sentimen, atau penggalan opini, adalah proses yang bertujuan untuk mengidentifikasi, mengenali, dan mengkategorikan emosi atau opini pengguna terkait layanan tertentu, seperti film, produk, acara, atau atribut lainnya, ke dalam kategori positif, negatif, atau netral (Mehta & Pandya, 2020). Tujuan dari analisis sentimen ini adalah agar perusahaan terkait dapat memahami persepsi masyarakat atau Pelanggan atas produk atau jasa yang disediakan. (Alfarobby & Irawan, 2024).

Dalam konteks klasifikasi sentimen pada *Google Maps Review*, terdapat dua jenis sentimen utama, yaitu sentimen positif dan sentimen negatif. Analisis ini berkontribusi dalam memahami bagaimana masyarakat merespons produk maupun layanan yang ditawarkan oleh instansi bisnis terkait.

C. Topic Modeling

Topic modeling adalah metode yang digunakan untuk merepresentasikan suatu topik dan merupakan bagian dari bidang *Natural Language Processing*. Metode ini berfungsi utama untuk menganalisis teks dengan memanfaatkan algoritma tertentu. (Adjie et al., 2022). Pemodelan ini bertujuan untuk mengungkap pola-pola tersembunyi dalam kumpulan teks dengan memanfaatkan metode pembagian kata yang terdapat di dalam teks atau dokumen. Hasil dari proses *topic modeling* berupa sekumpulan topik yang masing-masing terdiri atas kelompok kata yang cenderung muncul bersamaan berdasarkan pola tertentu di dalam teks atau dokumen tersebut. *Topic modeling* adalah model statistik yang digunakan untuk mengidentifikasi pokok bahasan atau tema dalam sekumpulan dokumen. (Rusdhi & Sari, 2022).

D. IndoBERT

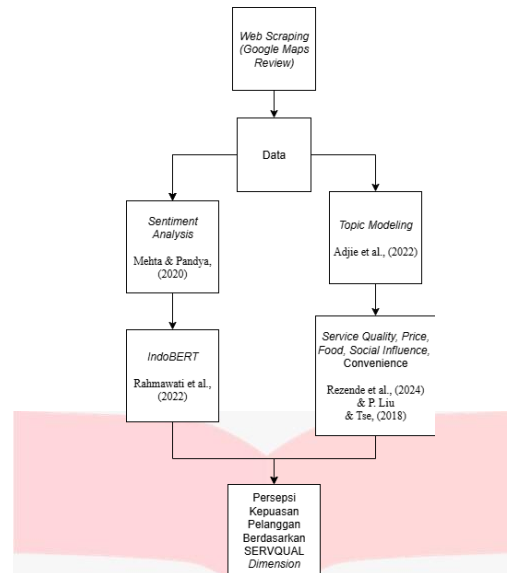
IndoBERT adalah model pra-terlatih yang menggunakan *transformer* dan dikembangkan untuk bahasa Indonesia (Rahmawati et al., 2022). Model *IndoBERT* juga Suatu versi modifikasi dari model BERT (*Bidirectional Encoder Representations from Transformers*) yang dibuat secara khusus untuk memproses teks berbahasa Indonesia (Situmorang & Purba, 2024). *IndoBERT* dikembangkan untuk memenuhi kebutuhan *Natural Language Processing* (NLP) pada bahasa Indonesia. Model ini dilatih menggunakan beragam kumpulan teks berbahasa Indonesia yang luas, mencakup sumber seperti berita, media sosial, serta berbagai dokumen lainnya.

IndoBERT adalah salah satu model BERT (*Bidirectional Encoder Representations from Transformers*) yang dikembangkan khusus untuk memproses berbagai bahasa, termasuk Bahasa Indonesia (Wilie et al., 2020). Model ini hadir dalam tiga varian, yaitu *IndoBERT-lite*, *IndoBERT-base*, dan *IndoBERT-large*. Perbedaan utama antara ketiganya terletak pada jumlah layer atau lapisan yang digunakan, yang secara langsung memengaruhi kemampuan model dalam memproses dan memahami teks dengan lebih mendalam. Model *IndoBERT* memanfaatkan arsitektur

transformer dua arah untuk mengekstraksi fitur yang digunakan dalam berbagai tugas lanjutan (Dudija et al., 2022).

E. Kerangka Pemikiran

Berdasarkan penjelasan studi literatur review, maka gambaran kerangka pemikiran penelitian ini adalah sebagai berikut:



Gambar 3 Kerangka Pemikiran

III. METODOLOGI PENELITIAN

Penelitian ini merupakan penelitian kuantitatif deskriptif yang bertujuan untuk menganalisis sentimen kepuasan pelanggan terhadap Warunk Upnormal berdasarkan ulasan *Google Maps*. Menurut (Sugiyono, 2018), penelitian kuantitatif berlandaskan positivisme, dilakukan pada kelompok populasi atau sampel spesifik, menggunakan alat ukur penelitian, lalu dianalisis secara statistik untuk menguji hipotesis. Dalam studi ini, penelitian dibagi ke dalam beberapa kategori sebagaimana ditampilkan pada Tabel 1.

Tabel 1. Karakteristik Penelitian

N o .	Karakteristik Penelitian	Jenis
1	Berdasarkan Metode	Kuantitatif
2	Berdasarkan Tujuan	Deskriptif
3	Strategi Penelitian	Studi Kasus
4	Unit Analisis	Individu
5	Keterlibatan Penelitian	Minimal
6	Latar Penelitian	<i>Noncontrived</i>
7	Waktu Pelaksanaan	<i>Longitudinal (Cross-sectional)</i>

A. Operasionalisasi Variabel

Untuk mengukur kepuasan pelanggan, penelitian ini menggunakan empat dimensi utama yang telah diidentifikasi oleh (Rezende et al., 2024) serta P. Liu & Tse, (2018). Dimensi-dimensi ini dijadikan sebagai variabel bertujuan mengevaluasi tingkat kepuasan pelanggan atas produk atau layanan yang ditawarkan, yang terdiri dari:

Tabel 3. 1 Variabel Penelitian

Variabel	Dimensi	Indikator
Kepuasan Pelanggan	Menurut Rezende et al., (2024) serta P. Liu & Tse, (2018), terdapat 5 dimensi yang dapat mempengaruhi kepuasan konsumen:	Pelanggan puas jika harga sesuai dengan kualitas, atau harga dianggap terjangkau.
	1. Price	
	2. Service Quality	Pelanggan puas jika pelayanan cepat, ramah, dan sesuai harapan.
	3. Overall Food Quality	Pelanggan puas jika makanan enak, variasi menu baik, dan presentasi menarik.
	4. Culture/Social Influence	Pengalaman bersantap didukung lingkungan sosial/budaya yang nyaman.
	5. Convenience	Pelanggan puas jika lokasi mudah diakses, waktu tunggu singkat, dan pemesanan praktis.

B. Tahapan Penelitian

Penelitian ini dimulai dengan tahap identifikasi permasalahan dan perumusan tujuan, dilanjutkan dengan kajian literatur sebagai landasan konseptual. Data dikumpulkan melalui teknik *web scraping* dari ulasan pengguna di Google Maps, kemudian diproses menggunakan model IndoBERT untuk membangun klasifikasi sentimen dengan akurasi tinggi. Selanjutnya, *topic modeling* berbasis BERTopic diterapkan secara otomatis untuk mengidentifikasi topik utama dalam setiap kategori sentimen melalui pendekatan *machine learning*.

C. Pengumpulan Data

Data untuk penelitian ini diperoleh melalui tahap *web scraping* pada *Google Maps Review*. Proses *web scraping* dilakukan secara otomatis menggunakan kata kunci tertentu yang relevan dengan penelitian. Periode waktu pengambilan data ditentukan sesuai kebutuhan penelitian, dan data yang diperoleh akan disimpan ke dalam format CSV untuk dianalisis lebih lanjut.

D. Teknik Analisis Data

1. Text Mining

Apify sebagai pengumpulan data untuk melakukan *scraping* data pada ulasan *Google Maps*. Setelah memperoleh data, data disimpan ke dalam format .csv.

2. Pre-processing

Setelah data terkumpul, dilakukan *preprocessing* sebagai langkah awal pengolahan data. Proses ini bertujuan untuk membersihkan data dan menjadikannya siap untuk dianalisis. Langkah-langkah *preprocessing* meliputi pembersihan data (*cleaning*), tokenisasi, normalisasi, penghapusan kata-kata tidak penting (*stopword removal*), dan *stemming*. Tahapan ini penting untuk mengurangi *noise* dalam data dan meningkatkan akurasi analisis.

a. Data Cleaning

Tahap pertama yakni *data cleaning* atau pembersihan data yang dilakukan untuk tujuan menghilangkan elemen-elemen yang tidak relevan atau tidak diperlukan dalam analisis, seperti karakter khusus, angka, atau spasi ganda yang dapat mengganggu proses selanjutnya. Pada tahap ini, karakter yang biasanya dihapus meliputi *URL*, nama pengguna (@username), tagar (#), tanda titik (.), koma (,), dan berbagai tanda baca lainnya (Syahril Dwi Prasetyo et al., 2023).

b. Case Folding

Case folding merupakan langkah pada proses *pre-processing* yang berfungsi untuk mengonversi semua huruf

kapital (*upper case*) dalam data menjadi huruf kecil (*lower case*), guna menyamakan format teks dan menghindari duplikasi kata yang hanya berbeda dalam penulisan huruf besar dan kecil (Nurrachmat Hidayat, 2023).

c. *Tokenization*

Tokenization adalah proses memisahkan kalimat menjadi kata-kata individual dengan menyortir dan mengenali struktur dari masing-masing kata. Proses ini juga dikenal sebagai penguraian data teks menjadi token atau potongan-potongan kecil, seperti kata atau frasa, yang dapat dianalisis lebih lanjut (Ayu Nirwana & Fathoni, 2024).

d. *Normalization*

Normalization merupakan tahap untuk memperbaiki kata-kata tidak baku maupun bentuk singkatan tertentu, dengan tujuan menyeragamkan penulisan kata. Tahapan ini dilakukan dengan mencocokkan setiap kata dalam dokumen dengan kata-kata yang diperoleh dalam kamus standar (Noviana & Rasal B A, 2023).

e. *Stopwords Removal*

Stopwords removal merupakan tahap menghapus kata-kata yang sering muncul namun tidak mempunyai makna penting dalam analisis, karena hanya berfungsi sebagai penghubung antar kata dalam kalimat. Contoh kata-kata tersebut antara lain "yang", "dan", "di", serta "ke", yang umumnya terdapat dalam daftar *stopwords* (Julianto et al., 2022).

f. *Stemming*

Stemming merupakan tahap untuk mengembalikan kata berimbuhan ke bentuk dasarnya. Proses ini dilakukan dengan menghapus awalan, sisipan, atau akhiran dari suatu kata guna memperoleh bentuk dasar yang lebih umum. Dalam bahasa Indonesia, proses *stemming* sering dibantu dengan menggunakan *library Sastrawi*, yaitu pustaka pemrosesan teks berbasis *PHP* yang dirancang untuk mengubah kata berimbuhan menjadi kata dasar. Sebagai contoh, kata “memakan” akan diubah menjadi “makan” (Darwis et al., 2020).

3. Analisis Sentimen

Tahap ini melibatkan pembangunan model klasifikasi sentimen menggunakan model *IndoBERT*. Model ini diterapkan pada data latih untuk mempelajari pola sentimen dalam ulasan pelanggan, dan hasilnya diuji menggunakan data uji. Dataset tersebut mengandung dua kemungkinan kategori sentimen, yaitu negatif dan positif. Data yang telah melewati tahap *pre-processing* selanjutnya diberi label dan dikategorikan sebagai data *training* berdasarkan sentimen konsumen terhadap kepuasan pelanggan Warunk Upnormal.

4. Evaluasi Model

Evaluasi model dilakukan menggunakan *confusion matrix*. Metode ini digunakan untuk mengukur kinerja model berdasarkan metrik seperti akurasi, presisi, *recall*, dan *f1-score* (Erfina & Wardani, 2022). Tabel *confusion matrix* merepresentasikan hasil prediksi model dibandingkan dengan kondisi aktual data, sehingga menyajikan gambaran jelas tentang kinerja model dalam mengklasifikasikan data dengan benar atau salah (Chicco et al., 2021). Hasil evaluasi akan menunjukkan sejauh mana model dapat mengklasifikasikan data dengan benar. Formula untuk metrik akurasi, presisi, *recall*, dan *f1-score* adalah sebagai berikut:

a. *Accuracy*: Akurasi digunakan untuk menilai proporsi prediksi benar yang dihasilkan oleh model.

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

b. *Precision*: Presisi menilai proporsi prediksi positif yang sesuai dengan kondisi sebenarnya.

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

c. *Recall*: *Recall* mengukur seberapa banyak prediksi positif yang benar-benar berhasil ditemukan oleh model.

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

d. *F1-Score*: *F1-score* merupakan metrik yang mengintegrasikan presisi dan *recall* guna memberikan penilaian menyeluruh terhadap kinerja model.

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

Keterangan:

- a. TP (*True Positive*): Data dengan prediksi positif yang benar.
- b. TN (*True Negative*): Data dengan prediksi negatif yang benar.
- c. FP (*False Positive*): Klasifikasi data negatif yang keliru sebagai positif.
- d. FN (*False Negative*): Klasifikasi data positif yang keliru sebagai negatif.

5. Topic Modeling

Pada tahap ini, *topic modeling* diterapkan untuk mengidentifikasi tema utama yang sering muncul dalam ulasan pelanggan. Algoritma yang digunakan, *BERTopic*, membantu dalam mengelompokkan kata-kata ke dalam topik tertentu berdasarkan kemunculannya dalam ulasan. Menurut (Grootendorst, 2022) *BERTopic* adalah sebuah model untuk pemodelan topik yang memanfaatkan teknik *transformer* dan *Sentence-BERT* untuk menghasilkan representasi teks yang lebih bermakna dan kaya akan konteks. Hasil dari *topic modeling* ini memberikan wawasan mengenai aspek-aspek yang menjadi perhatian pelanggan, seperti kualitas layanan, harga, atau suasana restoran.

BERTopic lebih efisien dibandingkan dengan metode-metode pemodelan topik tradisional, seperti *Latent Dirichlet Allocation (LDA)*, karena ia mempertimbangkan makna semantik dalam teks dan tidak hanya fokus pada kata-kata individual.

IV. HASIL DAN PEMBAHASAN

A. Karakteristik Data

Data penelitian ini diperoleh dari ulasan pengguna di Google Maps pada seluruh cabang Warunk Upnormal di Bandung. Pengumpulan dilakukan melalui scraping dengan kata kunci relevan dan penyaringan bahasa Indonesia. Sebanyak 21.917 data tidak terstruktur berhasil dikumpulkan, mencakup berbagai opini seperti pujian, keluhan, dan saran, yang selanjutnya diproses menjadi data bersih untuk dianalisis.

B. Hasil Penelitian

1. Hasil Klasifikasi Sentimen Analisis Warunk Upnormal

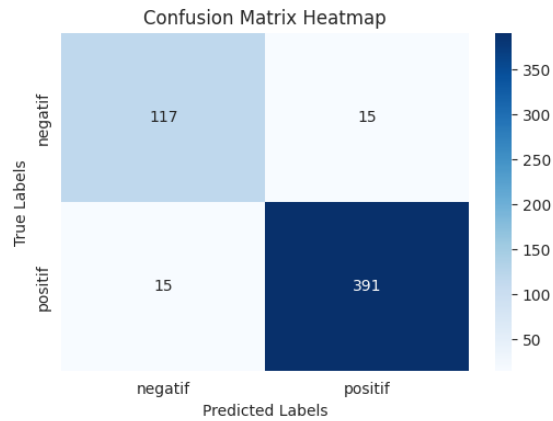
Klasifikasi sentimen dilakukan pada ulasan pelanggan Warunk Upnormal dari Google Maps menggunakan model IndoBERT. Data dibagi menjadi 80% data latih, 10% data uji, dan 10% data validasi.

Classification Report:				
	precision	recall	f1-score	support
0	0.89	0.89	0.89	132
1	0.96	0.96	0.96	406
accuracy			0.94	538
macro avg	0.92	0.92	0.92	538
weighted avg	0.94	0.94	0.94	538
Accuracy: 0.9442				

Gambar 4 *Classification Report Analysis Sentiment*

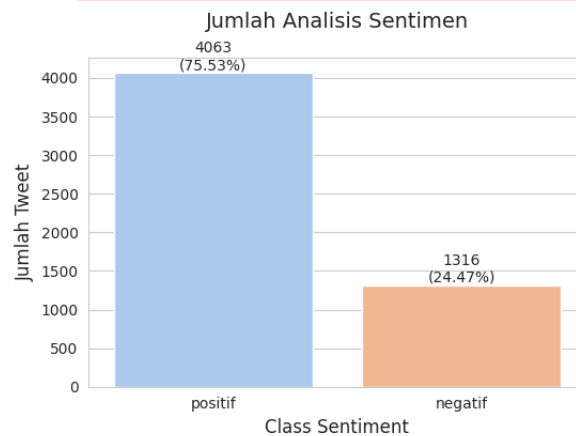
Sumber: Data Olahan Penulis (2025)

Gambar 4.1 menunjukkan hasil evaluasi model IndoBERT dengan akurasi 94,42%. Model menunjukkan kinerja seimbang antar kelas, dengan F1-score tertinggi pada kelas positif (0,96) dan terendah pada kelas negatif (0,89). Nilai macro average F1-score sebesar 0,92 dan weighted average 0,94 menunjukkan performa klasifikasi yang baik secara keseluruhan.



Gambar 5 *Confusion Matrix Heatmap*
Sumber: Data Olahan Penulis (2025)

Berdasarkan informasi yang diperoleh dari confusion matrix, model *IndoBERT* memberikan performa yang baik dalam melakukan pelabelan. Distribusi sentimen secara keseluruhan, baik positif maupun negatif, dapat dilihat pada Gambar 4.3 di bawah ini.

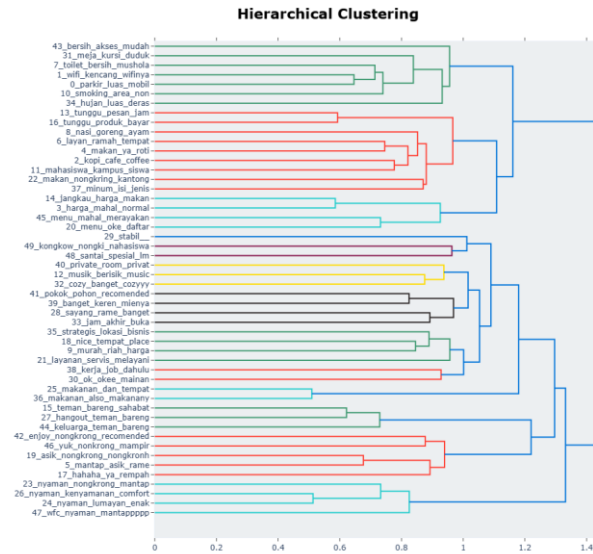


Gambar 6 Jumlah Analisis Sentimen
Sumber: Data Olahan Penulis (2025)

Gambar 4.3 memperlihatkan distribusi sentimen, dengan 75,53% (4.063 ulasan) tergolong positif dan 24,47% (1.316 ulasan) negatif. Hal ini menunjukkan mayoritas pelanggan memberikan opini positif terhadap Warunk Upnormal cabang Bandung, meskipun masih terdapat ulasan negatif yang mencerminkan ketidakpuasan dan perlu dianalisis lebih lanjut.

2. Analisis *Topic Modeling* Positif

Berdasarkan Hasil analisis sentimen positif melibatkan 4.063 data. Topik-topik di dalamnya terbentuk secara otomatis menjadi beberapa kelompok berdasarkan kesamaan makna kata-katanya.



Gambar 4. 1 Hierarchical Clustering Positif
Sumber: Olahan Data Penulis (2025)

Pendekatan *Hierarchical Clustering* digunakan untuk mengelompokkan informasi dari data yang kompleks. Pada sentimen positif, diperoleh 48 topik yang dikelompokkan ke dalam 4 *cluster*. Penjelasan masing-masing *cluster* disajikan pada tabel berikut.

Tabel 4. 1 Hasil Analisis *Topic Modeling* Positif

Cluster	Nomor Topik	Total Count	Kata Kunci Topik	Hasil Analisis
1	0, 4, 8, 14, 22, 37	1016	enak_makan_enak, makan_ya_roti, nasi_goreng_ayam, jangkau_harga_makan, makanan_kongkrong_santai, minum_isi_jenis	Fokus <i>Cluster</i> ini adalah <i>Food Quality</i> . secara jelas menunjukkan kepuasan terhadap kualitas, variasi, dan rasa makanan serta minuman. Topik lain yang tidak berkaitan langsung dengan makanan/minuman telah diabaikan
2	6, 21, 38	139	layanan_ramah_tempat, layanan_servis_melayani, kerja_job_dahulu	Fokus <i>Cluster</i> ini adalah pada <i>Service Quality</i> . indikator langsung dari kepuasan terhadap kualitas pelayanan dan interaksi dengan staf. juga bisa mengindikasikan pelayanan cepat/efisien.

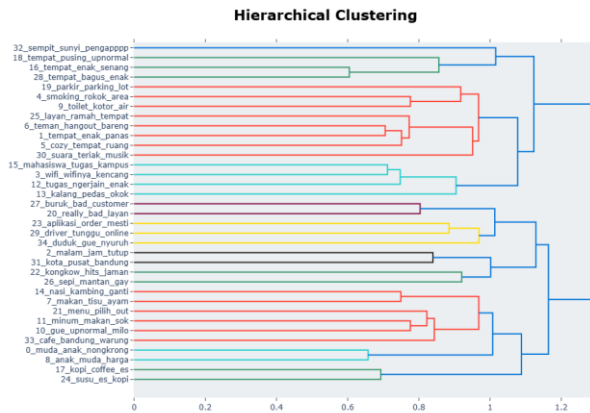
3	5, 15, 19, 27, 28, 30, 32, 33, 35, 36, 39, 40, 41, 42, 44, 46, 48, 49	497	mantap_asik_ra me, teman_bareng_s ahabat, asik_nongkrong _hangat, hangout_teman_ bareng, sayang_rame_ra mee, musik_berisik_ music, cozy_banget_co zyy, jam_akhir_buka, strategis_lokasi_ bisnis, makanan_also_r amaaee, banget_keren_m ienya, private_room_pr ivat, pokok_pohon_re comended, enjoy_nongkron g_recomended, keluarga_teman _bareng, yuk_nongkrong _mampir, santai_spesial_l m, kongkow_nongk i_mahasiswa	Fokus <i>Cluster</i> ini adalah pada <i>Social Influence</i> . semuanya sangat relevan dengan suasana, interaksi sosial, dan daya tarik tempat untuk berkumpul.
4	23, 24, 26, 47	107	nyaman_nongkr ong_mantap, nyaman_lumaya n_enak, nyaman_kenya manan_comfort, wfc_nyaman_m antappppp	Fokus <i>Cluster</i> ini, adalah pada <i>Convenience</i> . Semua topik ("nyaman_nongkro ng_mantap", "nyaman_lumayan _enak", "nyaman_kenyama nan_comfort", "wfc_nyaman_man tappppp") berfokus pada kenyamanan fisik dan lingkungan.

--	--	--	--	--

Langkah selanjutnya setelah analisis tiap *cluster* adalah mengelompokkan *cluster* tersebut ke dalam dimensi kepuasan pelanggan. Terdapat empat dimensi yang diidentifikasi, yaitu *Food Quality*, *Service Quality*, *Social Influence*, dan *Convenience*. *Cluster 1* dikaitkan dengan dimensi *Food Quality*, mencerminkan kepuasan terhadap rasa, variasi, dan kualitas makanan serta minuman. *Cluster 2* menggambarkan dimensi *Service Quality*, berfokus pada kepuasan pelanggan terhadap pelayanan dan interaksi dengan staf. *Cluster 3* mewakili dimensi *Social Influence*, di mana ulasan menyoroti suasana yang mendukung aktivitas sosial dan kebersamaan. Sementara itu, *Cluster 4* berhubungan dengan dimensi *Convenience*, menunjukkan kenyamanan fisik dan lingkungan tempat yang dirasakan pelanggan.

3. Analisis Topic Modeling Negatif

Berdasarkan Hasil analisis sentimen negatif melibatkan 1.316 data. Topik-topik di dalamnya terbentuk secara otomatis menjadi beberapa kelompok berdasarkan kesamaan makna kata-katanya.



Gambar 4. 2 Hierarchical Clustering Negatif
Sumber: Data Olahan Penulis (2025)

Pendekatan *Hierarchical Clustering* digunakan untuk mengelompokkan informasi dari data yang kompleks. Pada sentimen positif, diperoleh 35 topik yang dikelompokkan ke dalam 3 *cluster*. Penjelasan masing-masing *cluster* disajikan pada tabel berikut.

Tabel 4. 2 Hasil Analisis Topic Modeling Negatif

Cluster	Nomor Topik	Total Count	Kata Kunci Topik	Hasil Analisis
1	1, 3, 4, 5, 6, 9, 12, 13, 15, 16, 18, 19, 25, 28, 30, 32	534	tempat_enak_panas, wifi_wifinya_kencang, smoking_rokok_area, cozy_tempat_ruangan, teman_hangout_bareng, toilet_kotor_air, tugas_ngeriain_ena, kalang_pedas_oke, mahasiswa_tugas_kampus,	<i>Cluster</i> ini mencakup keluhan terkait fasilitas fisik, kenyamanan, dan suasana umum tempat. Oleh karena itu, dimensi yang paling dominan adalah <i>Convenience</i> .

			tempat_enak_senang, tempat_pusing_up normal, parkir_parking_lot, layanan_ramah_tempat, tempat_bagus_enak, suara_teriak_musik, sempit_sunyi_pengap	
2	2, 20, 23, 27, 29, 31, 34	154	malam_jam_tutup, really_bad_layanan, aplikasi_order_mesti, buruk_bad_customer, driver_tunggu_online, kota_pusat_bandung, duduk_gue_nyuruh	<i>Cluster</i> ini sangat fokus pada keluhan mengenai kualitas dan ketanggapan layanan. Ini adalah cerminan langsung dari ketidakpuasan terhadap dimensi <i>Service Quality</i> .
3	7, 11, 14, 17, 21, 24	128	makanan_tisu_ayam_sayang, minum_makan_sork_keluarga, nasi_kambing_ganti_ayam, kopi_coffee_es_manis, menu_pilih_out_solid, susu_es_kopi_enakkk	<i>Cluster</i> ini dominan dengan <i>Food Quality</i> . dengan jelas mengindikasikan masalah pada kualitas dan ketersediaan produk makanan/minuman.

Langkah selanjutnya setelah melakukan analisis pada masing-masing *cluster* adalah mengelompokkan *cluster* tersebut ke dalam dimensi kepuasan pelanggan. Terdapat tiga dimensi yang diidentifikasi, yaitu *Service Quality*, *Food*, dan *Convenience*. Dimensi *Convenience* terkait erat dengan *cluster* 1, yang berisi keluhan mengenai kondisi fisik tempat seperti ruangan sempit, toilet kotor, area merokok, dan parkir, yang mengganggu kenyamanan pelanggan. Dimensi *Service Quality* mencakup *cluster* 2, yang menyoroti keluhan terhadap pelayanan, mulai dari ketidaksiapan staf hingga proses pemesanan yang kurang efisien. Sementara itu, *cluster* 3 berkaitan dengan dimensi *Food Quality*, yang menunjukkan ketidakpuasan terhadap kualitas makanan dan minuman, termasuk masalah pada rasa, menu, dan ketersediaan produk.

C. Pembahasan Hasil Penelitian

1. Hasil Analisis Sentimen

Analisis sentimen terhadap ulasan Google Maps Warunk Upnormal cabang Bandung dilakukan menggunakan model IndoBERT yang telah melalui proses pelatihan. Model ini mencapai akurasi sebesar 94,42%, menunjukkan bahwa sebagian besar prediksi sesuai dengan label sentimen sebenarnya. F1-score tertinggi sebesar 0,96 pada kelas positif mencerminkan keseimbangan yang baik antara precision dan recall, serta menunjukkan kemampuan model dalam mengklasifikasikan sentimen secara cukup seimbang.

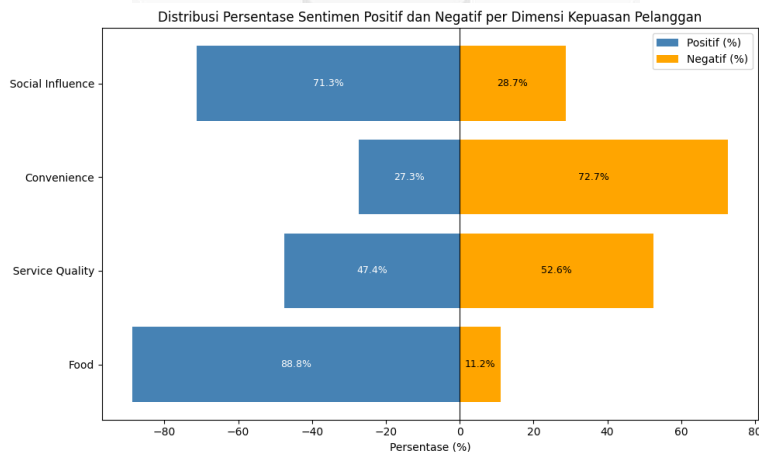
2. Hasil Analisis *Topic Modeling*

Hasil *topic modeling* dilakukan setelah klasifikasi sentimen terhadap ulasan pelanggan Warunk Upnormal. Setiap kelompok sentimen, baik positif maupun negatif, dianalisis lebih lanjut berdasarkan topik dan dikategorikan ke dalam dimensi kepuasan pelanggan. Proses ini bertujuan untuk mengidentifikasi faktor utama yang memengaruhi persepsi pelanggan terhadap layanan yang diberikan.

Tabel 4. 3 Hasil Analisis *Topic Modeling* Pada Dimensi Kepuasan Pelanggan

Dimensi Kepuasan	Sentimen	Total Count Topik
Food	Positif	1016
	Negatif	128
Service Quality	Positif	139
	Negatif	154
Convenience	Positif	107
	Negatif	258
Social Influence	Positif	497
	Negatif	200

Berdasarkan hasil *topic modeling*, setiap dimensi kepuasan pelanggan memiliki sejumlah topik relevan dari sentimen positif dan negatif. Pada dimensi *Food*, sentimen positif mencatat 1.016 *count*, menunjukkan kepuasan terhadap kualitas dan variasi makanan, sedangkan sentimen negatif sebesar 128 *count* mencerminkan kritik terhadap rasa dan ketersediaan menu. Dimensi *Service Quality* mencatat 139 *count* positif dan 154 *count* negatif, menunjukkan apresiasi sekaligus keluhan terhadap layanan staf. Pada dimensi *Convenience*, terdapat 107 *count* positif dan 258 *count* negatif, mengindikasikan kepuasan atas kenyamanan tempat namun juga banyak keluhan soal fasilitas. Sementara itu, dimensi *Social Influence* memperoleh 497 *count* positif dan 200 *count* negatif, mencerminkan peran tempat sebagai ruang interaksi sosial, meskipun ada kekhawatiran terkait suasana dan kebisingan.



Gambar 7 Presentase Sentimen Positif dan Negatif Dimensi Kepuasan Pelanggan
Sumber: Data Olahan Penulis (2025)

Berdasarkan Gambar 4.8, hasil *topic modeling* dan distribusi persentase sentimen mengungkap beberapa temuan penting terkait persepsi pelanggan terhadap Warunk Upnormal. Meskipun sentimen positif mendominasi sebagian besar dimensi kepuasan, sentimen negatif tetap muncul di tiap dimensi, menunjukkan area yang perlu perbaikan. Pada dimensi *Food*, sentimen positif mencapai 88,8%, mencerminkan kepuasan tinggi terhadap cita rasa dan variasi menu, seperti nasi goreng dan ayam. Sebaliknya, dimensi *Convenience* didominasi sentimen negatif sebesar 72,7%,

menunjukkan ketidakpuasan terhadap fasilitas seperti Wi-Fi, toilet, dan parkir. Pada dimensi *Service Quality*, sentimen negatif sedikit lebih tinggi (52,6%) dibanding positif (47,4%), menandakan adanya masalah dalam pelayanan dan responsivitas staf. Sementara itu, dimensi *Social Influence* menunjukkan sentimen positif sebesar 71,3%, mencerminkan kenyamanan dalam bersosialisasi, meskipun masih ada keluhan terkait suasana yang terlalu ramai atau kurang menarik.

D. Analisis Manajerial

Berdasarkan analisis sentimen menggunakan model IndoBERT terhadap ulasan Google Maps Warunk Upnormal cabang Bandung, diperoleh akurasi sebesar 94,42%. Meskipun secara umum sentimen cenderung positif, ditemukan ketidakpuasan pelanggan pada beberapa aspek layanan. Sentimen negatif yang mencapai 24,47% terutama muncul dalam dimensi *Service Quality*, *Food*, *Social Influence*, dan *Convenience*, yang menunjukkan perlunya peningkatan pada area tersebut. Pada dimensi *Food*, sentimen positif mendominasi (88,8%), menunjukkan kekuatan utama Warunk Upnormal pada kualitas dan variasi makanan, meskipun konsistensi rasa dan ketersediaan menu tetap perlu dijaga. Dimensi *Service Quality* menunjukkan lebih banyak sentimen negatif (52,6%), mencerminkan masalah dalam kecepatan layanan, interaksi staf, dan sistem pemesanan. Diperlukan pelatihan staf dan evaluasi proses operasional untuk meningkatkan kualitas pelayanan. Pada dimensi *Convenience*, dominasi sentimen negatif (72,7%) mengindikasikan ketidakpuasan terhadap fasilitas fisik seperti Wi-Fi, toilet, area merokok, dan parkir, sehingga perbaikan infrastruktur menjadi prioritas. Sementara itu, dimensi *Social Influence* menunjukkan sentimen positif sebesar 71,3%, menandakan kenyamanan dalam aspek sosial, meski masih ada keluhan terkait kebisingan atau suasana yang kurang mendukung. Secara manajerial, temuan ini menjadi dasar penting bagi Warunk Upnormal untuk menyusun strategi perbaikan yang lebih spesifik, mulai dari peningkatan layanan, pembenahan fasilitas, hingga optimalisasi pengalaman pelanggan secara menyeluruh agar kepuasan dan loyalitas dapat terus ditingkatkan.

V. KESIMPULAN DAN SARAN

Berdasarkan analisis ulasan pelanggan Warunk Upnormal cabang Bandung melalui Google Maps Review, penelitian ini menunjukkan bahwa model IndoBERT mampu mengklasifikasikan sentimen dengan akurasi tinggi sebesar 94,42%. Sentimen positif mendominasi dengan 75,53%, namun masih terdapat 24,47% sentimen negatif yang mencerminkan ketidakpuasan pelanggan pada aspek tertentu. Hasil *topic modeling* mengungkap bahwa dimensi *Food*, *Social Influence*, dan *Service Quality* menjadi kekuatan utama berdasarkan sentimen positif, sedangkan dimensi *Convenience* dan *Service Quality* paling banyak dikritik, terutama terkait fasilitas fisik, kualitas layanan, dan responsivitas staf. Berdasarkan temuan ini, Warunk Upnormal disarankan untuk menjaga konsistensi rasa dan variasi menu makanan, meningkatkan pelayanan melalui pelatihan staf dan evaluasi sistem pemesanan, serta memperbaiki fasilitas fisik seperti Wi-Fi, toilet, area merokok, dan kenyamanan ruang. Langkah-langkah ini penting untuk menjawab ekspektasi pelanggan dan meningkatkan kepuasan serta loyalitas secara berkelanjutan.

REFERENSI

- Adjie, K. R., Santoso, P., Husna, A., Putri, N. W., & Rakhmawati, A. (2022). Analisis Topik Tagar Covidindonesia pada Instagram Menggunakan Latent Dirichlet Allocation. In *Jurnal Informatika Sunan Kalijaga* (Vol. 7, Issue 1).
- Alfarobby, A. N., & Irawan, H. (2024). *Analisis Sentimen Kepuasan Konsumen Pengguna Transportasi Online Pada Ulasan Google Playstore Menggunakan Indobert Dan Topic Modeling (Studi kasus: Gojek dan Grab)* (Vol. 11, Issue 1).
- Ayu Nirwana, Z. F., & Fathoni, M. Y. (2024). Sentiment Analysis of JNT Express, JNE and Sicepat Customer Satisfaction on Twitter Using Multinomial Naïve Bayes Method. *International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, 716–721. <https://doi.org/10.1109/EECSI63442.2024.10776297>
- Bennix. (2023, January 31). *PRAHARA DIBALIK HANCURNYA WARUNK UPNORMAL 2023! INVESTOR SAHAM WAJIB TAHU!* <https://www.youtube.com/watch?v=7NkEIue4jBg&t=758s>.
- Chicco, D., Tötsch, N., & Jurman, G. (2021). The matthews correlation coefficient (Mcc) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation. *BioData Mining*, 14, 1–22. <https://doi.org/10.1186/s13040-021-00244-z>
- Darwis, D., Shintya Pratiwi, E., Ferico, A., & Pasaribu, O. (2020). PENERAPAN ALGORITMA SVM UNTUK ANALISIS SENTIMEN PADA DATA TWITTER KOMISI PEMBERANTASAN KORUPSI REPUBLIK INDONESIA. In *Jurnal Ilmiah Edutic* (Vol. 7, Issue 1).
- Dudija, N., Natalia, L., Alamsyah, A., & Romadhony, A. (2022). Identification of Extraversion and Neuroticism Personality Dimensions Using IndoBERT's Deep Learning Model. *Proceedings of the 2022 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology, IAICT 2022*, 155–159. <https://doi.org/10.1109/IAICT55358.2022.9887476>

- Erfina, A., & Wardani, N. (2022). ANALISIS SENTIMEN PERGURUAN TINGGI TERMEWAH DI INDONESIA MENURUT ULASAN GOOGLE MAPS MENGGUNAKAN SUPPORT VECTOR MACHINE (SVM). *Jurnal Manajemen Informatika & Sistem Informasi*, 5(1), 77–85.
- Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. <http://arxiv.org/abs/2203.05794>
- Irawan, H., Akmalia, G., & Masrury, R. A. (2019). Mining tourist's perception toward Indonesia tourism destination using sentiment analysis and topic modelling. *ACM International Conference Proceeding Series*, 7–12. <https://doi.org/10.1145/3361821.3361829>
- Julianto, Y., Setiabudi, D. H., & Rostianingsih, S. (2022). Analisis Sentimen Ulasan Restoran Menggunakan Metode Support Vector Machine. *Jurnal Infra*, 10(1).
- Masrury, R. A., Fannisa, & Alamsyah, A. (2019, July 1). Analyzing tourism mobile applications perceived quality using sentiment analysis and topic modeling. *2019 7th International Conference on Information and Communication Technology, ICoICT 2019*. <https://doi.org/10.1109/ICoICT.2019.8835255>
- Mehta, P., & Pandya, Dr. S. (2020). A Review On Sentiment Analysis Methodologies, Practices And Applications. *INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH*, 9(2), 601–609.
- Noviana, R., & Rasal B A. (2023). PENERAPAN ALGORITMA NAIVE BAYES DAN SVM UNTUK ANALISIS SENTIMEN BOY BAND BTS PADA MEDIA SOSIAL TWITTER. *JTS*, 2(2), 51–60.
- Nurrachmat Hidayat, F. (2023). Analisis Sentimen Masyarakat Terhadap Perekrutan PPPK Pada Twitter Dengan Metode Naive Bayes ... (Fajar Nurrachmat Hidayat, Sugiyono) Analisis Sentimen Masyarakat Terhadap Perekrutan PPPK Pada Twitter Dengan Metode Naive Bayes Dan Support Vector Machine. *Jurnal Sains Dan Teknologi*, 5(2), 665–672. <https://doi.org/10.55338/saintek.v5i1.1359>
- Rahmawati, A., Alamsyah, A., & Romadhony, A. (2022). Hoax News Detection Analysis using IndoBERT Deep Learning Methodology. *2022 10th International Conference on Information and Communication Technology, ICoICT 2022*, 368–373. <https://doi.org/10.1109/ICoICT55009.2022.9914902>
- Rezende, G. A., Mariano, A. M., Santos, M. R., & Coelho Constatin, A. C. (2024). Modeling Customer Satisfaction in the Food Industry: Insights from a Structural Equation Approach. *Procedia Computer Science*, 242, 130–137. <https://doi.org/10.1016/j.procs.2024.08.250>
- Rifa, A., Sujaini, H., Prawira, D., & Hadari Nawawi, J. H. (2021). Sentiment Analysis Objek Wisata Kalimantan Barat Pada Google Maps Menggunakan Metode Naive Bayes. *JEPIN (Jurnal Edukasi Dan Penelitian Informatika)*, 7(3), 400–407.
- Rusdhi, V. F., & Sari, I. (2022). IDENTIFIKASI TOPIK ARTIKEL BERITA MENGGUNAKAN TOPIC MODELLING DENGAN LATENT DIRICHLET ALLOCATION. *Jurnal Ilmiah Informatika Komputer*, 27(2), 169–176. <https://doi.org/10.35760/ik.2022.v27i2.6829>
- Situmorang, G. F., & Purba, R. (2024). Deteksi Potensi Depresi dari Unggahan Media Sosial X Menggunakan Teknik NLP dan Model IndoBERT. *Technology and Science (BITS)*, 6(2). <https://doi.org/10.47065/bits.v6i2.5496>
- Sugiyono. (2018). *Metode Penelitian Kuantitatif*.
- Syahril Dwi Prasetyo, Shofa Shofiah Hilabi, & Fitri Nurapriani. (2023). Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes dan KNN. *Jurnal KomtekInfo*, 1–7. <https://doi.org/10.35134/komtekinfo.v10i1.330>
- Tjiptono, F., & Diana, A. (2022). *Manajemen dan Strategi Kepuasan Pelanggan*.
- Wilie, B., Vincentio, K., Winata, G. I., Cahyawijaya, S., Li, X., Lim, Z. Y., Soleman, S., Mahendra, R., Fung, P., Bahar, S., & Purwarianti, A. (2020). *IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding*. <http://arxiv.org/abs/2009.05387>