



Analisis Topik Dan Sentimen Berbasis *Algoritma Latent Dirichlet Allocation (LDA)* Dan *Bidirectional Encoder Representations From Transformers (BERT)*: Studi Kasus Ulasan Pelanggan Pada *E-Commerce ruparupa.com*

Surya Permana¹, Sajarwo Anggai², Taswanda Taryo³

^{1,2,3} Teknik Informatika S-2, Program Pascasarjana, Universitas Pamulang, Kota Tangerang Selatan, Banten

¹suryapermana.095018@gmail.com, ²sajarwo@mail.com, ³dosen02234@unpam.ac.id

ABSTRACT

The growth of e-commerce in Indonesia has led to an increasing volume of customer reviews containing vital information. These reviews are generally in the form of unstructured text, necessitating text analysis methods to extract meaningful insights. This study aims to analyze topics and sentiments in customer reviews of the e-commerce platform ruparupa.com by utilizing Latent Dirichlet Allocation (LDA) and Bidirectional Encoder Representations from Transformers (BERT) algorithms. The LDA algorithm is used to identify the main topics frequently discussed by customers, while BERT is employed to classify review sentiments into positive, negative, and neutral categories. By using Lexicon-Based and VADER as an automatic labeling mechanism (auto-labeling), the preprocessing stage includes cleaning, case folding, and stemming using the Sastrawi library to ensure the quality of the input data. The LDA algorithm is implemented to extract latent topic structures, which are then mapped into five main categories: Price, Application, Service, Product Quality, and Delivery. Furthermore, the DistilBERT model is trained through a fine-tuning process using the AdamW optimizer for 3 epochs. The sentiment analysis results indicate that the model demonstrates very strong performance, as reflected by high accuracy and consistently optimal precision, recall, and F1-score across all sentiment classes. This customer sentiment distribution reflects the level of user satisfaction with the services of ruparupa.com. The combination of LDA and BERT methods is proven effective in providing an overview of key issues and customer perceptions.

Keywords: Sentiment Analysis, Topic Modeling, LDA, BERT, ruparupa.com

ABSTRAK

Perkembangan e-commerce di Indonesia mendorong meningkatnya jumlah ulasan pelanggan yang mengandung informasi yang sangat penting. Ulasan pelanggan tersebut umumnya berbentuk teks yang tidak terstruktur, sehingga diperlukan metode analisis teks untuk mengekstraksi informasi yang bermakna. Penelitian ini bertujuan untuk menganalisis topik dan sentimen pada ulasan pelanggan e-commerce ruparupa.com dengan memanfaatkan algoritma Latent Dirichlet Allocation (LDA) dan Bidirectional Encoder Representations from Transformers (BERT). Algoritma LDA digunakan untuk mengidentifikasi topik-topik utama yang sering dibahas oleh pelanggan, sedangkan BERT digunakan untuk melakukan klasifikasi sentimen ulasan ke dalam kategori positif, negatif, dan netral. Dengan Lexicon-Based dan VADER sebagai mekanisme pelabelan otomatis (*auto-labeling*), tahap *preprocessing* meliputi *cleaning*, *case folding*, dan *stemming* menggunakan pustaka Sastrawi untuk menjamin kualitas input data. Algoritma LDA diimplementasikan untuk mengekstraksi struktur topik laten yang kemudian dipetakan menjadi lima kategori utama: Harga, Aplikasi, Pelayanan, Kualitas Produk, dan Pengiriman. Selanjutnya, model DistilBERT dilatih melalui proses *fine-tuning* menggunakan optimizer AdamW selama 3 *epoch*. Hasil analisis sentimen menunjukkan bahwa model memiliki kinerja yang sangat baik, yang ditunjukkan oleh tingkat akurasi yang tinggi serta nilai *precision*, *recall*, dan *F1-score* yang optimal pada setiap kelas sentimen. Distribusi sentimen pelanggan tersebut mencerminkan tingkat kepuasan pengguna terhadap layanan ruparupa.com. Kombinasi metode LDA dan BERT terbukti efektif dalam memberikan gambaran mengenai isu utama dan persepsi pelanggan.

Kata Kunci : Analisis Sentimen, Pemodelan Topik, LDA, BERT, ruparupa.com

1. PENDAHULUAN

Perkembangan pesat industri *e-commerce* di Indonesia telah menyebabkan lonjakan volume ulasan pelanggan yang tersebar di berbagai *platform marketplace* [1]. Salah satunya adalah ruparupa.com, sebuah *e-commerce* yang menyajikan ribuan produk [1]kebutuhan rumah tangga. Ulasan pelanggan dalam *platform* ini tidak hanya menjadi cerminan pengalaman pembelian, tetapi juga mengandung informasi yang sangat berharga untuk perusahaan dalam merancang strategi pemasaran yang lebih *responsif*.

Menurut survei yang dilakukan oleh [2] Ulasan online semakin kritis dalam memengaruhi keputusan konsumen, dengan 87% membaca ulasan sebelum memilih bisnis lokal—naik dari 81% di 2022—dan 53% menganggapnya "sangat penting," terutama untuk layanan kesehatan dan restoran. Sebanyak 76% konsumen menghindari bisnis berrating di bawah 3 bintang, sementara 95% curiga jika bisnis tidak menanggapi ulasan negatif. Kredibilitas ulasan bergantung pada detail spesifik, dan 73% konsumen bisa mendeteksi ulasan palsu. *Google My Business* menjadi platform tepercaya utama (63%), diikuti Yelp (45%). Bisnis harus merespons ulasan—khususnya negatif—dalam 1 minggu atau kurang, karena respons sopan bisa tingkatkan kepercayaan 45% konsumen. Dengan 93% konsumen mengandalkan *internet* untuk mencari bisnis lokal dan 74% lebih memprioritaskan ulasan tertulis daripada rating bintang, manajemen reputasi digital yang proaktif menjadi kunci kesuksesan bisnis.

Penelitian mengenai analisis ulasan pelanggan dalam platform *e-commerce* telah berkembang pesat seiring kemajuan teknologi pemrosesan bahasa alami (NLP), terutama dengan kehadiran model *BERT* dan algoritma topik seperti *LDA*. Salah satu studi relevan dilakukan oleh [3] yang menganalisis perbandingan layanan dan kemudahan penggunaan aplikasi *e-commerce* menggunakan model *BERT*. Penelitian ini menunjukkan akurasi tinggi (95–99%) dan mengungkapkan bahwa sentimen positif mendominasi ulasan Lazada (55,1%), sentimen netral lebih banyak pada Shopee (47,5%), sementara Tokopedia mendapat porsi terbesar untuk sentimen negatif (16,1%). Studi ini menegaskan kemampuan *BERT* dalam memahami sentimen berbasis topik pada ulasan pelanggan.

Selanjutnya, [4] mengusulkan integrasi algoritma *LDA* dengan representasi *embedding* dari *BERT* untuk menganalisis keluhan pelanggan dalam *e-commerce*

Indonesia. Penelitian ini menghasilkan skor *coherence* tertinggi sebesar 0,388 dengan model *BERT-Base Multilingual Cased*, dan berhasil mengidentifikasi topik-topik utama keluhan seperti domain, produk, dan promo. Hal ini menunjukkan bahwa penggabungan kekuatan semantik *BERT* dengan struktur probabilistik *LDA* dapat meningkatkan kualitas identifikasi topik dari data teks berbahasa Indonesia.

Penelitian lain yang juga relevan dilakukan oleh [5] yang mengembangkan analisis sentimen berbasis aspek terhadap ulasan tempat wisata di Kalimantan Tengah. Dengan memanfaatkan kombinasi *LDA* dan *BERT*, penelitian ini menghasilkan akurasi sentimen sebesar 98,02% dan ketepatan identifikasi aspek seperti lingkungan (98,21%) dan pengalaman (99%) dari total 19.315 ulasan. Temuan ini memperkuat efektivitas model *BERT* dalam mendeteksi topik dan aspek yang terkandung dalam ulasan pelanggan di sektor pariwisata.

Selama ini, banyak penelitian menganalisis ulasan pelanggan dengan menggunakan analisis sentimen. Metode ini mengklasifikasikan opini menjadi positif, negatif, atau netral [6]. Meskipun *informatif*, pendekatan ini tidak cukup untuk menggali dimensi topikal dalam opini pengguna, seperti jenis keluhan, pujian, atau saran spesifik yang tersembunyi dalam teks

Penelitian ini mengusulkan penerapan kombinasi *LDA* dan *BERT* secara praktis untuk menganalisis ulasan pelanggan di *e-commerce* ruparupa.com. Proses dimulai dengan *web scraping* menggunakan Python untuk mengumpulkan data ulasan. Data yang diperoleh kemudian melalui tahapan pra-pemrosesan teks, seperti *Cleaning tokenizing, case folding, stopword removal, dan stemming*, untuk membersihkan teks dari elemen yang tidak relevan. Selanjutnya setelah data bersih dilakukan pemodelan topik menggunakan *LDA* guna mengelompokkan ulasan ke dalam sejumlah topik dominan. Terakhir, hasil dari pemrosesan tersebut diimplementasikan dalam sebuah aplikasi antarmuka grafis (GUI) berbasis Flask.

2. METODE

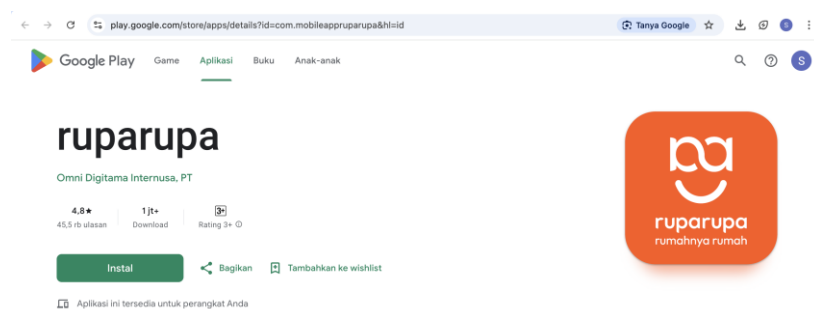
Objek penelitian ini adalah salah satu *e-commerce* terbesar di Indonesia yaitu PT.Omni digitama internusa atau yang biasa disebut ruparupa.com. adapun subjek Penelitian ini adalah Opini pelanggan yang berbelanja di *e-commerce* ruparupa.com melalui *review* pada database website google playstore.

2.1. Jenis dan Sumber data

Jenis data yang digunakan pada penelitian ini adalah data sekunder, Adapun sumber data yang digunakan dalam penelitian ini adalah *Dataset reviews* terhadap kepuasan Pelanggan terhadap pelayanan kepuasan berbelanja yang diberikan di *website Google Play*, yang mana peneliti mengambil secara online dengan link berikut ini:

<https://play.google.com/store/apps/details?id=com.mobileappruparupa&hl=id>

Setelah link diklik akan muncul Tampilan berikut



Gambar 2.1 Aplikasi ruparupa.com di *playstore*

Pengambilan data secara online ini dinamakan *web scraping* atau *scraping data*. Adapun *scraping data* pada penelitian ini dilakukan dengan menggunakan *Google Colab* dan *Python*.

2.2. Metode Pemilihan sampel

Sampel dalam penelitian ini adalah semua data *review* atau ulasan pengguna PT. Omni Digitama Internusa (ruparupa.com) dari database *website google play* dengan aplikasinya yang bernama ruparupa.com dengan total 30.000 *review*. sedangkan sampel diambil dari suatu populasi yang mana pada penelitian ini digunakan sampel ulasan ruparupa.com dari tanggal 01 Desember 2023 Sampai dengan 30 Desember 2025. dengan jumlah sampel 3.000 *review*.

2.3. Teknik Pengumpulan Data

Metode yang digunakan untuk proses pengumpulan data dari penelitian ini adalah sebagai berikut :

1. Web Scraping

Data pada penelitian ini adalah data sekunder yang diambil secara online dengan teknik *scraping data* menggunakan ekstensi Scraper yang ada di *Google Colab* dengan

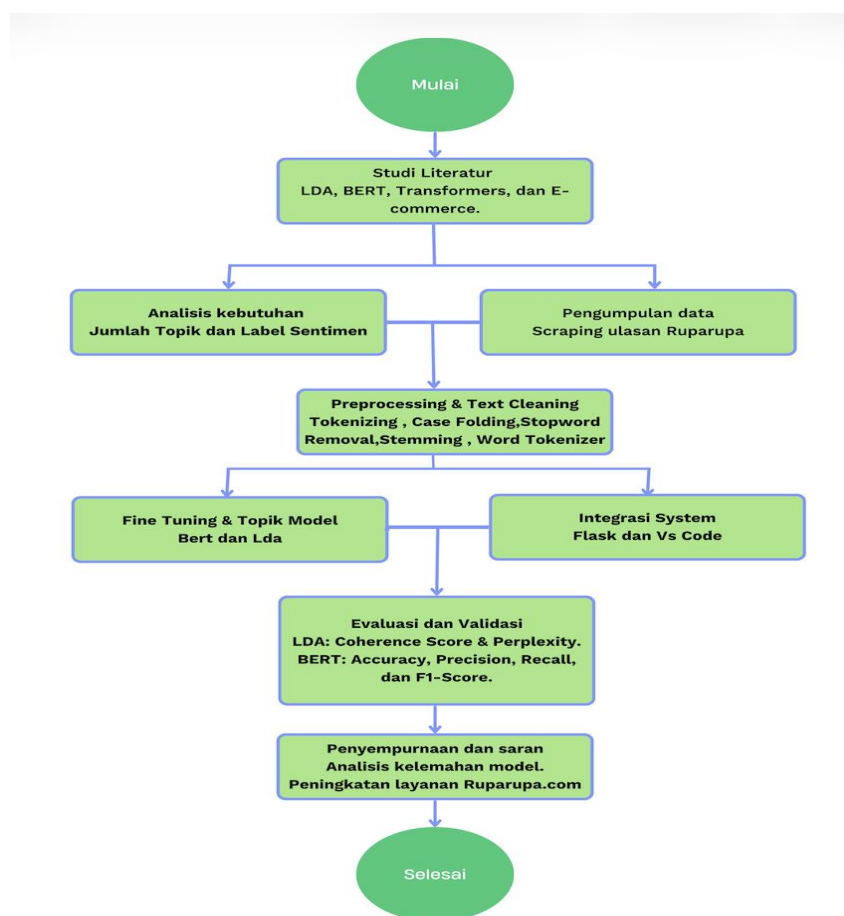
menggunakan *Python*. Melalui teknik *scraper* ini didapatkan informasi dari *website* secara otomatis tanpa perlu menyalin secara manual.

2. Studi Pustaka

Studi pustaka dilakukan dengan menggunakan beberapa kajian literatur, buku, maupun referensi jurnal yang sekiranya berkaitan dengan tujuan, rumusan, batasan, dan metode penelitian.

2.4. Rancangan Penelitian

Berdasarkan diagram alir penelitian, langkah-langkah pada penelitian ini terdiri dari sebagai berikut:



Gambar 2.2. Rancangan Penelitian

Berdasarkan Rancangan penelitian di atas, tahapan proses dimulai dari studi literatur dan pengumpulan data ulasan Ruparupa.com melalui *web scraping*, yang dilanjutkan dengan tahap *preprocessing* teks. Selanjutnya, dilakukan integrasi sistem menggunakan *framework* Flask serta penerapan *fine-tuning* Algoritma *BERT* untuk klasifikasi sentimen dan algoritma *LDA* untuk pemodelan topik. Seluruh proses ini

diakhiri dengan evaluasi performa model menggunakan metrik evaluasi guna memberikan rekomendasi perbaikan layanan yang akurat bagi perusahaan.

2.5. Teknik Analisis Data

Teknik yang digunakan dalam penelitian ini Adalah sebagai berikut

2.5.1 *Scraping Data*

Dalam penelitian ini, pengumpulan data dilakukan melalui proses *scraping* pada aplikasi ruparupa.com yang tersedia di *Google Play*. Data yang diambil berupa ulasan pengguna yang diberikan antara tanggal 01 Desember 2023 hingga 30 Desember 2025 sebanyak 3000 ulasan

2.5.2 *Text Processing*

Proses yang dilakukan untuk membersihkan dan mempersiapkan data agar siap digunakan dalam analisis topik dan sentimen. seperti Case Folding, stopwords Removal, Tokenisasi, Stemming dan Normalisasi teks Hasil dari *Preprocessing* ini adalah *Dataset* yang lebih bersih, terstruktur, dan siap untuk dianalisis.

2.5.3 *Evaluasi Model*

1. *Laten Dirichlet Allocation (LDA)*

$$\text{Coherence Score} = \frac{C_{UMass}}{|T|} \sum_{i < j} \log \frac{D(w_i, w_j) + \epsilon}{D(w_j)} \quad (1)$$

Penjelasan :

1. Mengukur ko-occurrence kata dalam dokumen
2. Menggunakan data internal Corpus
3. Nilai umumnya negative
4. Semakin mendekati nol, topik semakin koheren

$$\text{Perplexity} \quad D = EXP - \frac{1}{M} \sum_{d=1} \log p(wd) \quad (2)$$

Penjelasan :

1. D = Kumpulan Dokumen
2. Wd = kata-kata dalam dokumen ke-d
3. Nd = Jumlah kata dalam dokumen ke-d
4. P(Wd) = Probabilitas kemunculan kata dalam model
2. *Bidirectional Encoder Representation From Transformer* (BERT)

Berikut Matrix evaluasi Untuk Model BERT

1. Accuracy
$$Accuracy = \frac{\text{Jumlah prediksi yang benar}}{\text{Total jumlah prediksi}} \quad (3)$$

2. Precission
$$Precision = \frac{\text{True positives}}{\text{True positives} + \text{False positives}} \quad (4)$$

3. F1-Scoree
$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (5)$$

4. Recall

$$Recall = \frac{\text{True positives}}{\text{True positives} + \text{False negatives}} \quad (6)$$

Penjelasan :

Accuracy : Akurasi mengukur proporsi dari prediksi yang benar dibandingkan dengan total jumlah prediksi yang dibuat

Precision : proporsi dari prediksi positif yang benar dibandingkan dengan seluruh prediksi positif yang dibuat

Recall : mengukur seberapa baik model dapat menangkap semua kasus positif yang benar-benar positif dalam *Dataset*

F1-Score : ukuran keseimbangan antara *Precision* dan *Recall*, memberikan rata-rata harmonik dari keduanya

3. HASIL DAN PEMBAHASAN

3.1. *Scraping Data*

Data yang digunakan dalam penelitian ini berupa ulasan pelanggan ruparupa.com yang diperoleh dari platform Google *playstore*, hasil dari proses scraping data disimpan dalam local database. Terlampir hasil dalam bentuk gambar dibawah ini :

```
import pandas as pd

df = pd.read_csv('Ulasan.csv')
df.head()
```

	reviewId	userName	userImage	content	score	thumbsU
0	13658fab-6f92-4764-babf-d2c236ae48fc	Dwi Susanto	lh.googleusercontent.com/9-ALV-U...	pelayanan azko royal plaza sangat buruk, bingu...	1	
1	efc464e7-cc76-4243-bc26-56ce1077082f	angga Kristian	lh.googleusercontent.com/ia/ACg8oc...	kalian yang mau menggunakan aplikasi saya kasi...	1	
	533b0e84-f401-4a77-	Ricky	https://play-	aplikasi masih eror koneksi		

Gambar 3.1 Hasil *Scraping* Data

3.2. Text Preprocessing

Setelah dilakukan scraping, data dilakukan pembersihan melalui tahapan Cleaning ,case folding, tokenizing, filtering dan stemming , pengguna dapat melanjutkan ke tahap *preprocessing* dengan menekan tombol *preprocessing*. Berikut kami lampirkan hasilnya dalam bentuk gambar berikut ini :

Data Diupload (2960 baris)		Preprocessing
Tampilkan 10 s	data per halaman	Cari:
ulasan	stemmed	
pelayanan azko royal plaza sangat buruk, bingung mau komplain kemana. karena telpun toko tidak bisa. call center azko juga tidak di angkat. oder dr aplikasi dan sangat percuma ada biaya pengiriman cepat, kalau packing lelet ya tetap lelet. padahal estimasi pengiriman 1-2 jam an. tolong koreksinya.	layan azko royal plaza sangat buruk bingung mau komplain mana karena tel toko tidak bisa call center azko juga tidak di angkat oder dr aplikasi dan sangat percuma ada biaya kirim cepat kalau packing lelet ya tetap lelet padahal estimasi kirim jam an tolong koreksi	
kalian yang mau menggunakan aplikasi.saya kasih tau pengalaman saya hari ini: untuk pembatalan pesanan sangat sangat sangat sangat sulit sekali dan banyak menyita waktu.harus melewati perdebatan yang panjang,dan kalaupun disetujui kembalinya akan jadi voucher untuk dibelanjakan diaplikasi lagi....banyak aplikasi lain yang melakukan pembatalan pesanan mudah dan gampang sangat menyita pikiran dan waktu hanya untuk pembatalan	kalian yang mau guna aplikasi saya kasih tau alam saya hari ini untuk batal pesan sangat sangat sangat sangat sangat sulit sekali dan banyak sita waktu harus lewat debat yang panjang dan kalaupun tuju kembali akan jadi voucher untuk belanja aplikasi lagi banyak aplikasi lain yang laku batal pesan mudah dan gampang sangat sita pikir dan waktu hanya untuk batal.	

Gambar 3.2 Proses Text *Preprocessing*

3.2.5 Laten Dirichlet Allocation (LDA)

Berdasarkan hasil pemodelan topik menggunakan algoritma *Latent Dirichlet Allocation* (LDA) yang divisualisasikan melalui *pyLDavis*, lima topik utama yang merepresentasikan isu-isu dominan dalam ulasan pelanggan Rupa-Rupa.com, berikut tampilan hasil visualisasi nya :



3.2.6 Analysis Sentimen BERT

a) Fine tuning bert

1. Epoch 1

2.Epoch 2

3. Epoch 3

Model mencapai tingkat pembelajaran yang lebih optimal. Pola sentimen telah dipahami secara lebih menyeluruh dan performa model menjadi lebih konsisten. Penggunaan

beberapa *epoch* bertujuan untuk meningkatkan kemampuan generalisasi model dalam mengklasifikasikan sentimen tanpa menyebabkan overfitting terhadap data pelatihan

b) Hasil Sentimen dengan BERT dan Topik dengan LDA

Berikut hasil proses *Fine-Tuning* menggunakan *distilBERT* selama 3 epoch dengan optimizer AdamW, model *DistilBERT* mampu melakukan klasifikasi sentimen pada ulasan pelanggan ruparupa.com secara otomatis ke dalam kategori positif, Negatif dan Netral.

Stemmed	Topik	Sentimen
garagara petugas informa salah input nomor hp promo jadi tidak bisa di pakai percuma saat butuh untuk di edit nomor telepon di aplikasi tidak ada menunya untuk edit nomor telepon mau masuk di website karena di daftarkan oleh petugas id dan ps nya gak tau minta tolong ke petugasnya gak dibantu harus lewat cs atau website perasaan akun mobile banking jg gak seribet ini paraaah	Aplikasi	Negatif
beli mesin cuci merk kels di informa mahal mahal juta yang tabung muat kg malah udah rusak ngasih penilaian produk di aplikasi malah penilaian saya ga ditampilkan lawak emg aplikasi di refund ga bisa di tuker ga bisa nyesel banget beli alat elektronik di informa	Harga	Negatif
jadi lebih gampang belanja produk informa dan poin yg kita dapet dari belanja tuh lumayan banget bisa kepeke untuk belanja lagi atau bisa juga pake jasa cleaningnya mereka sejauh ini aplikasinya pun ga ada bug di aku aplikasinya juga mudah digunakan dan informatif juga	Aplikasi	Positif
lebih baik kalo mau ada aplikasinya harusnya harga lebih murah bukan makin tinggi	Harga	Positif
kapok belanja di informa di informa supermall karawaci pengiriman lambat seenaknya merubah jadwal pengiriman tidak sesuai kesepakatan gak bakal belanja disini lagi	Pengiriman	Negatif

Gambar 3.4 Hasil *Analysis* sentiment BERT

c) *Matrix* evaluasi

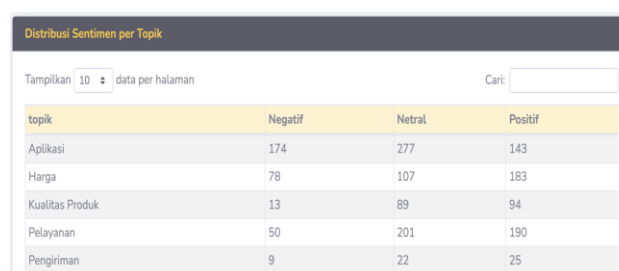
Berdasarkan hasil pengujian pada 592 data uji, model BERT menunjukkan performa yang sangat kuat dengan akurasi total sebesar 91%.terlampir dalam bentuk gambar 3.5

		precision	recall	f1-score	support
	Negatif	0.84	0.77	0.81	83
	Netral	0.90	0.90	0.90	233
	Positif	0.94	0.96	0.95	276
	accuracy			0.91	592
	macro avg	0.89	0.88	0.89	592
	weighted avg	0.91	0.91	0.91	592

Gambar 3.5 *Matrik* evaluasi Bert

3.3. Kesimpulan

Berdasarkan hasil implementasi dan evaluasi *System Analisis topik dan sentiment* berbasis *Algoritma LDA dan BERT*, berdasarkan hasil gambar dibawah ini beberapa kesimpulan kunci dapat ditarik



topik	Negatif	Netral	Positif
Aplikasi	174	277	143
Harga	78	107	183
Kualitas Produk	13	89	94
Pelayanan	50	201	190
Pengiriman	9	22	25

Gambar 3.6 Distribusi *sentiment* pertopik

1. Identifikasi Topik (LDA): Pelanggan Ruparupa.com sering membahas lima topik utama,yaitu :Aplikasi ,Harga ,Kualitas Produk ,Pelayanan dan Pengiriman
2. Analisis Sentimen (BERT): Distribusi sentimen pelanggan secara keseluruhan didominasi oleh sentimen Netral. Penggunaan algoritma BERT terbukti sangat efektif dengan tingkat akurasi mencapai 91%, di mana model sangat handal dalam mengenali kepuasan pelanggan (F1-score 0.91).

Kombinasi Metode: Integrasi LDA dan BERT berhasil memberikan gambaran bahwa isu utama pelanggan terletak pada aspek Aplikasi, yang memiliki angka sentimen negatif tertinggi. Sebaliknya, aspek Harga dan Kualitas Produk merupakan kekuatan utama perusahaan dengan persepsi positif yang sangat kuat.

4. DAFTAR PUSTAKA

- [1] A. S. Pramudita, "How reviews and ratings influence consumer purchase intentions in Indonesian e-commerce," *J. Logistik Bisnis*, vol. 14, p. 2, 2024.
- [2] B. BrightLocal, "Local Consumer Review Survey 2023," BrightLocal, 2023. [Online]. Available: <https://www.brightlocal.com/research/local-consumer-review-survey-2023/>. [Accessed 26 01 2026].
- [3] A. G. Yuda, R. Novita, Mustakim, dan M. Afdal, "Comparison of service and ease of e-commerce user applications using BERT information system," *Jurnal Sistem Cerdas*, vol. 7, no. No.2, pp. 98-107, 2024.

- [4] M. A. Khadija dan W. Nurharjadmo,, "Enhancing Indonesian customer complaint analysis: LDA topic modelling with BERT embeddings," *SINERGI (Journal)*, vol. Vol.28, no. No.1, pp. 153-162, 2024.
- [5] B. F. R. Putra, V. H. Pranatawijaya, dan P. B. A. A. Putra, "Analisis Sentimen Berbasis Aspek pada Tempat Wisata di Kalimantan Tengah dengan Memanfaatkan Model Deep Learning," *JOINTECOMS (Journal of Information Technology and Computer Science)*, vol. Vol. 4, no. No. 3, pp. 200-211, 2024.
- [6] W. Medhat, A. Hassan, dan H. Korashy, "Sentiment analysis algorithms and applications: A survey," *Ain Shams Engineering Journal*, vol. 5, no. 4, p. 1093–1113, 2014.
- [7]. Azmina, F. N., Liebenlito, M., & Zulkifli, D. U. (2025). Community Detection, Topic, and Sentiment Analysis of the Palestine-Israel Issue on Social Media X. *Jurnal Matematika, Statistika dan Komputasi*, 21(2), 502–519. <https://doi.org/10.20956/j.v21i2.41308> [2]
- [8] Febriany, A. K., Dyar Wahyuni, E., & Permatasari, R. (t.t.). Analisis Sentimen Berbasis Aspek Pada Ulasan Aplikasi Indrive Menggunakan Bidirectional Encoder Representations From Transformers (Bert). *Jurnal Ilmiah Wahana Pendidikan*, 2024(20), 105–115. <https://doi.org/10.5281/zenodo.14263982>
- [9] Hanifa, A., Debora, C., Hasani, M. F., & Wicaksono, P. (2024). Analyzing Views on Presidential Candidates for Election 2024 Based on the Instagram and X Platforms with Text Clustering. *Procedia Computer Science*, 245(C), 730–739. <https://doi.org/10.1016/j.procs.2024.10.299>
- [10] Hutama, L. B., & Suhartono, D. (2022). Indonesian Hoax News Classification with Multilingual Transformer Model and BERTopic. *Informatica (Slovenia)*, 46(8), 81–90. <https://doi.org/10.31449/inf.v46i8.4336>
- [11] Lihardo Girsang, D., Sidiq, A., & Salsabila Elenaputri, T. (2023). Analisis Sentimen Masyarakat terhadap Layanan BPJS Kesehatan dan Faktor-Faktor Pendukung Opini dengan Pemodelan Natural Language Processing (NLP). *Emerging Statistics and Data Science Journal*, 1(2).

- [12] Sriyanti, Z. A., Kartika, D. S. Y., & Najaf, A. R. E. (2024). Implementasi Model Bert Pada Analisis Sentimen Pengguna Twitter Terhadap Aksi Boikot Produk Israel. *Jurnal Informatika Dan Teknik Elektro Terapan*, 12(3), 2335–2342. <https://doi.org/10.23960/Jitet.V12i3.4743>