

Analisis Sentimen Masyarakat terhadap *Brand* Lokal di *E-Commerce* dengan Pendekatan *Deep Learning* Menggunakan Metode XLNet

I Gusti Made Diva Widia Wiartha^{a1}, I Made Agus Dwi Suarjaya^{a2}, Ni Kadek Dwi Rusjyanthi^{a3}

^aProgram Studi Teknologi Informasi, Fakultas Teknik, Universitas Udayana
Bukit Jimbaran, Bali, Indonesia, telp. (0361) 704170

e-mail: 1divawidia@student.unud.ac.id, 2agussuarjaya@it.unud.ac.id,
3dwi.rusjyanthi@unud.ac.id

Abstrak

Pertumbuhan perekonomian Indonesia dapat didongkrak dengan meningkatkan minat masyarakat untuk membeli produk brand lokal. Brand awareness, brand image, dan kualitas produk yang baik dapat memengaruhi minat masyarakat dalam membeli produk lokal. Teknologi Big Data dapat dimanfaatkan untuk menganalisis sentimen masyarakat melalui data review produk brand lokal untuk mengetahui citra dan kualitas produk brand lokal. Penelitian yang dilakukan oleh penulis bertujuan untuk mengetahui sentimen masyarakat terhadap brand lokal berdasarkan data review produknya di situs e-commerce Tokopedia, Shopee, Blibli, Bukalapak, Lazada, dan Zalora. Metode yang digunakan pada penelitian yaitu XLNet. Penelitian yang dilakukan oleh penulis memperoleh data review sebanyak 16.663.045 review dengan jumlah data yang telah diproses sebanyak 4.110.012 review serta dataset training sebanyak 11.606 review yang terdiri dari 5803 data review positif dan 5803 data review negatif. Hasil analisis sentimen terhadap produk brand lokal di situs e-commerce menunjukkan bahwa data review positif lebih dominan (87,82%) dibandingkan data review negatif (12,18%) sehingga dapat disimpulkan bahwa masyarakat Indonesia cenderung puas dan menyukai produk brand lokal yang tersedia di e-commerce. Hasil performa metode XLNet menunjukkan nilai accuracy sebesar 95,6%; precision sebesar 95,06%; recall sebesar 96,21%; dan f1-score sebesar 95,63%.

Kata kunci: analisis sentimen, e-commerce, klasifikasi, review produk, XLNet

Abstract

Indonesia's economic growth can be boosted by increasing people's interest in buying local brands products. Brand awareness, brand image, and good product quality can influence people's interest in buying local brands products. Big Data technology can be used to analyze public sentiment through product reviews to determine the brand image and product quality of local brands. The research conducted by the author aims to determine public sentiment towards local brands based on product reviews in e-commerce such as Tokopedia, Shopee, Blibli, Bukalapak, Lazada, and Zalora. The research conducted by the authors obtained 16,663,045 reviews with a total of 4,110,012 processed reviews and 11,606 training datasets consisting of 5803 positive reviews and 5803 negative reviews. The results of sentiment analysis for local brand products on e-commerce show that positive reviews are more dominant (87.82%) than negative reviews (12.18%), it can be concluded that Indonesian people tend to be satisfied and like the available local brand products in e-commerce. The performance results of the XLNet method show an accuracy value of 95.6%; precision of 95.06%; recall of 96.21%; and f1-score of 95.63%.

Keywords: sentiment analysis, e-commerce, classification, product reviews, XLNet

1. Pendahuluan

Kementerian perindustrian terus gencar mengajak seluruh masyarakat Indonesia untuk semakin mencintai, menggunakan, dan mempromosikan produk industri dalam negeri demi meningkatkan pertumbuhan industri nasional. Penelitian widiyono pada tahun 2019 menyatakan bahwa salah satu upaya untuk meningkatkan pertumbuhan industri nasional adalah

menumbuhkan rasa nasionalisme di kalangan generasi muda. Rasa nasionalisme dapat ditumbuhkan dengan selalu menggunakan produk dalam negeri dan bangga dalam menggunakan produk dalam negeri [1].

Agus Gumiwang Kartasasmita, selaku Menteri Perindustrian Indonesia mengungkapkan bahwa dampak pembelian produk dalam negeri senilai Rp400 triliun dapat meningkatkan pertumbuhan ekonomi nasional sebesar 1,67 hingga 1,71% berdasarkan hasil simulasi yang dilakukan oleh Badan Pusat Statistik. Pertumbuhan ekonomi pada tahun 2021 sebesar 3,69%, penggunaan produk dalam negeri yang maksimal diperkirakan dapat mendorong peningkatan ekonomi Indonesia sampai 5,36 hingga 5,4% [2]. Meningkatkan minat masyarakat dalam membeli dan menggunakan produk dalam negeri sangat penting agar merek dan produk dalam negeri dapat berkembang dan meningkatkan ekonomi Indonesia.

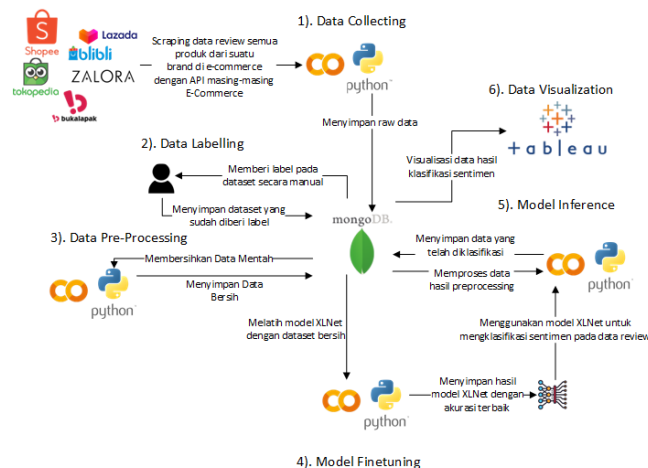
Menurut hasil survei *We Are Social* per April 2021 menyatakan bahwa 88,1% pengguna internet di Indonesia sudah memakai layanan *e-commerce* untuk membeli barang dengan total nilai transaksi mencapai Rp. 266 triliun [3]. Layanan *e-commerce* muncul untuk memudahkan konsumen dan penjual atau pihak *brand*. Data *review* dari pembeli dapat dimanfaatkan oleh pihak *brand* untuk mengetahui umpan balik dari pembeli.

Teknologi digunakan untuk mendukung berbagai aktivitas mulai dari aktivitas ringan hingga berat di berbagai bidang. Teknologi merupakan sarana utama penyampaian informasi yang dibutuhkan berupa komputer dan jaringan internet [4]. Perkembangan teknologi informasi menghasilkan data dalam jumlah besar setiap harinya dengan kumpulan data yang masif dan memiliki struktur yang besar dan kompleks sehingga diperlukan suatu metode untuk mengolah data tersebut yang dikenal dengan istilah *Big Data* [5]. Teknologi *Big Data* berkembang dalam pengelolaan data dengan jumlah yang sangat besar dan kompleks sehingga dapat membantu dalam mendapatkan berbagai wawasan, informasi, solusi; serta meningkatkan keunggulan kompetitif yang signifikan di sebagian besar organisasi, perusahaan, institusi, dan lain-lain [6]. Hasil penelitian sebelumnya terkait dengan pemanfaatan Big Data dalam meningkatkan kualitas maskapai penerbangan menunjukkan bahwa analisis sentimen pada ulasan penumpang maskapai dapat digunakan untuk mendapatkan informasi mengenai kekuatan dan kelemahan pelayanan maskapai kepada pelanggan sehingga maskapai tersebut dapat mengembangkan strategi baru dan meningkatkan pangsa pasar mereka [7]. Penelitian sebelumnya yang terkait dengan analisis sentimen berhasil menggunakan metode *Multinomial Naïve Bayes* untuk mengklasifikasi sentimen pada data *review* produk di situs Lazada dengan hasil nilai *accuracy*, *precision*, *recall*, dan *f1-score* sebesar 90% [8]. Penelitian yang serupa berhasil menganalisis sentimen pada data *review* di situs Shopee dengan metode KNN dan TF-IDF yang dapat diterapkan untuk sistem analisis pemasaran dalam meningkatkan produk beserta pelayanannya [9]. Penelitian lain yang serupa berhasil memanfaatkan metode XLNet dalam mengimplementasikan analisis sentimen pada *dataset movie review*; metode XLNet tersebut dapat mengungguli performa model *machine learning* tradisional *Naïve Bayes*, *Random Forest*, dan SVM dalam mengklasifikasi sentimen [10].

Penelitian yang dilakukan oleh penulis bertujuan untuk menganalisis sentimen masyarakat terhadap *brand* lokal berdasarkan data *review* produk *brand* lokal di *e-commerce* dengan metode XLNet sehingga informasi mengenai sentimen masyarakat terhadap suatu *brand* lokal dapat digali lebih dalam. Klasifikasi sentimen pada penelitian ini dibagi menjadi dua, yaitu: positif dan negatif. Hasil dari penelitian yang dilakukan oleh penulis dapat digunakan untuk mengevaluasi kepuasan konsumen terhadap produk dari *brand* lokal serta membantu *brand* lokal dalam mengembangkan produk, layanan, dan strategi *marketing* menjadi lebih baik sehingga *brand* lokal tersebut dapat menarik minat calon konsumen yang lebih banyak sekaligus meningkatkan pertumbuhan ekonomi Indonesia dengan menggencarkan masyarakat untuk membeli produk lokal.

2. Metode Penelitian

Metode penelitian yang digunakan untuk menganalisis sentimen opini masyarakat terhadap *brand* lokal berdasarkan data *review* produk *brand* lokal di *e-commerce* dengan XLNet dapat dipaparkan ke dalam gambaran umum penelitian yang terdiri dari enam tahapan utama, yaitu: *data collecting*, *data labelling*, *data pre-processing*, *finetuning* model XLNet, *model inference*, dan *data visualization*. Gambaran umum penelitian dapat dilihat pada Gambar 1.



Gambar 1. Gambaran Umum Metode Penelitian

2.1 Data Collecting

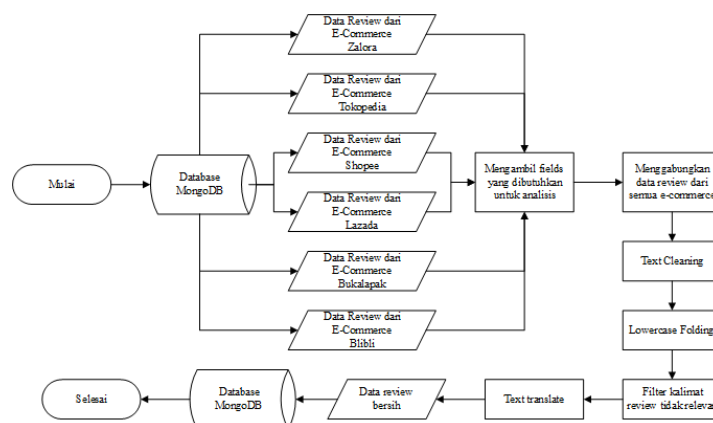
Data yang dikumpulkan adalah data *review* semua jenis produk dari setiap *brand* lokal di situs *e-commerce* Tokopedia, Shopee, Lazada, Bukalapak, Bilibli, dan Zalora yang diambil dengan teknik *scraping* pada API masing-masing situs *e-commerce*. Penelitian ini menggunakan 46 *brand* lokal yang dipilih dari penghargaan *Top Brand Award 2022* berdasarkan *Top Brand Index* tertinggi di masing-masing kategorinya [11]. Data *review* hasil *scraping* kemudian disimpan dalam format JSON ke dalam *database* MongoDB.

2.2 Data Labelling

Data *review* dari salah satu *brand* di setiap kategorinya diambil kembali dari *database* MongoDB untuk diberi label secara manual oleh penulis dan satu orang *volunteer*. Label sentimen pada setiap data *review* di penelitian ini berupa sentimen positif (1) atau negatif (0). *Dataset* yang sudah diberi label kemudian disimpan ke dalam *database* MongoDB.

2.3 Data Pre-Processing

Tahap data *pre-processing* merupakan proses mengubah data teks menjadi bentuk yang lebih mudah dipahami dan terstruktur agar data siap digunakan dalam proses analisis [12]. *Flowchart* tahap *pre-processing* data pada penelitian dapat dilihat pada Gambar 2.



Gambar 2. Flowchart Tahap Data Pre-Processing

Tahap pertama dalam proses *pre-processing* data *review* adalah mengambil data *review* salah satu *brand* dari *e-commerce* Bilibli, Bukalapak, Lazada, Shopee, Tokopedia, dan Zalora

yang telah tersimpan di *database* MongoDB. *Fields* data yang digunakan pada penelitian dapat dilihat pada Tabel 1.

Tabel 1. *Fields* Data *Review* yang Digunakan untuk Analisis

E-Commerce	Fields
Blibli	id, content, createDate, product.name, product.brand
Bukalapak	id, text_review, created_at, product.name
Tokopedia	id, message, reviewCreateTime, product.name
Lazada	reviewRateId, reviewContent, boughtDate, itemTitle
Shopee	cmtid, comment, ctime, product_items.name
Zalora	id, text, SubmittedAt, product.Name, product.Brand

Terdapat *fields* yang tidak berisi *value* nama *brand* dari beberapa produk di *e-commerce* Bukalapak, Tokopedia, Lazada, dan Shopee sehingga perlu penambahan *fields* *product_brand* yang berisikan nama *brand* dari data *review* pada tahap *pre-processing*. Tahap selanjutnya adalah menggabungkan data *review* salah satu *brand* dari setiap *e-commerce* yang telah diambil dari *database* MongoDB. Tahap *rename* dilakukan pada setiap *fields* di salah satu *brand* dari semua *e-commerce* sebelum penggabungan data *review* yang bertujuan untuk membuat data seragam dan dapat digabungkan. Hasil *field* data *review* setelah di-*rename* dapat dilihat pada Tabel 2.

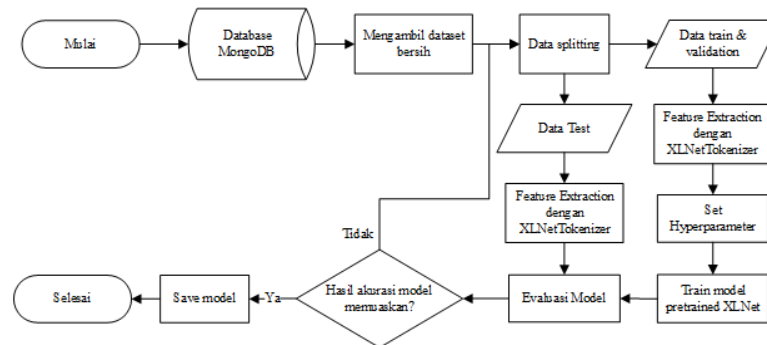
Tabel 2. *Fields* Data *Review* Setelah di-*Rename*

E-Commerce	Fields	Fields Rename
Blibli	id, content, createDate, product.name, product.brand	id, review, review_date, product_name, product_brand
Bukalapak	id, text_review, created_at, product.name	id, review, review_date, product_name
Tokopedia	id, message, reviewCreateTime, product.name	id, review, review_date, product_name
Lazada	reviewRateId, reviewContent, boughtDate, itemTitle	id, review, review_date, product_name
Shopee	cmtid, comment, ctime, product_items.name	id, review, review_date, product_name
Zalora	id, text, SubmittedAt, product.Name, product.Brand	id, review, review_date, product_name, product_brand

Fields yang telah di-*rename* sudah menjadi seragam dan dapat digabungkan menjadi satu. Tahap Proses *text cleaning* terdiri dari menghilangkan *noise* pada data *review* seperti karakter numerik, karakter spesial/symbol, spasi berlebihan, URL, emoji; mengubah huruf pada kata yang memiliki huruf yang sama lebih dari dua menjadi satu; menghilangkan data teks yang kosong/*null*; dan menghapus data *review* yang duplikat berdasarkan id; *lowercase folding* (proses mengubah kata huruf kapital menjadi huruf kecil) yang bertujuan untuk mengurangi redundansi data dan menyesuaikan kata dengan *vocabulary* pada model *pre-trained* XLNet; filter kalimat *review* yang tidak relevan dengan produk; serta *text translate* untuk menerjemahkan teks *review* bahasa Indonesia ke bahasa Inggris karena model XLNet hanya mengenali kata dalam bahasa Inggris.

2.4 Model Finetuning

Model finetuning merupakan tahapan untuk melatih model *pre-trained* XLNet dengan *dataset training* yang sudah melalui tahapan *pre-processing* dan diberi label secara manual, serta melatih model tersebut dengan beberapa *hyperparameter* untuk memperoleh model dengan performa yang mumpuni dalam mengklasifikasi sentimen data *review* produk.



Gambar 3. Flowchart Finetuning Model XLNet

Gambar 3 merupakan *flowchart finetuning* model XLNet pada penelitian. *Dataset training* yang sudah melalui tahapan *pre-processing* dan sudah diberi label digunakan untuk *training* model XLNet secara *supervised learning*. *Dataset training* dibagi menjadi tiga bagian, yaitu: data *train* dan data *validation* untuk *training* model, serta data *test* untuk proses evaluasi model XLNet. Model *pre-trained* yang digunakan untuk *finetuning* adalah 'xlnet-base-cased' yang tersedia di situs huggingface.co dengan *library* Tensorflow. *Dataset training* yang sudah dibagi perlu menjalani proses *feature extraction* untuk dijadikan input model XLNet dengan fungsi XLNetTokenizer. Nilai *hyperparameter* yang digunakan pada tahap *finetuning* model *pre-trained* XLNet meliputi *learning rate* (1e-5, 2e-5, dan 3e-5), *epoch training*, *optimizer* (Adam dan AdamW), *dropout* (0.1, 0.2, 0.3, 0.4, dan 0.5), *max sequential length* (128), *batch size* (32), dan *data split ratio* (80:10:10, dan 60:20:20). Parameter *callbacks* yang digunakan pada proses *finetuning*, yakni: *EarlyStopping* untuk menghentikan proses *training* pada *epoch* apabila *metric loss* yang dimonitor tidak mengalami penurunan sama sekali dan *ModelCheckpoint* untuk menyimpan hasil *training* model pada hasil *metrics* terbaik yang diperoleh pada setiap *epoch training*.

2.5 Model Inference

Inference model merupakan tahapan penggunaan model XLNet yang telah di-*training* untuk mengklasifikasi sentimen berupa positif dan negatif pada data *review* yang telah melalui tahap *pre-processing*. Data *review* kemudian dilakukan proses prediksi oleh model XLNet dengan hasil akhir berupa logits. Angka logits kemudian dikalkulasikan dengan fungsi aktivasi *Softmax* untuk memperoleh nilai probabilitas di setiap kelasnya; *text review* diklasifikasikan menjadi negatif jika nilai probabilitas pada *vector index* ke-0 lebih besar, *text review* diklasifikasikan menjadi positif jika nilai probabilitas pada *vector index* ke-1 lebih besar. Data yang sudah diberi label sentimen kemudian disimpan ke dalam *database* MongoDB

2.6 Data Visualization

Data visualization merupakan tahapan dalam menampilkan hasil analisis data dalam bentuk yang mudah dipahami. *Data visualization* dilakukan dengan menggunakan Tableau. Tableau adalah perangkat lunak yang dirancang untuk membuat visualisasi data, laporan, dan dasbor secara instan [13]. Data yang telah diproses dan diklasifikasi kemudian ditransformasi dalam bentuk persentase, bagan, maupun diagram seperti diagram *bar*, *line*, *pie*, dan *word cloud*.

3. Kajian Putaka

Kajian pustaka berisikan materi yang berkaitan dengan penelitian meliputi *text mining*, analisis sentimen, XLNet, dan evolusi kinerja model (*confusion matrix*).

3.1. Analisis Sentimen

Analisis sentimen adalah pola pikir atau tanggapan berupa pendapat dan komentar seseorang tentang sesuatu, seperti objek atau proses, dengan cara menganalisis dan mengklasifikasikannya, sehingga dapat dibedakan konteks yang bermuatan positif atau negatif. Analisis sentimen juga bermanfaat bagi para praktisi dan peneliti, seperti sosiologi, pemasaran, periklanan, psikologi, ekonomi, politik, sains, karena membutuhkan banyak data tentang interaksi manusia dan komputer [14]. Analisis sentimen dapat dikategorikan menjadi dua bagian, yaitu:

kategori pertama berisi tugas inti atau utama (tugas analisis sentimen dasar) dan kategori kedua berisi subtugas (subkategori utama). Analisis sentimen inti meliputi klasifikasi sentimen tingkat dokumen (*document-level sentiment classification*), klasifikasi tingkat kalimat (*sentence-level sentiment classification*), dan klasifikasi tingkat aspek (*aspect-level sentiment classification*). Kategori Sub-tugas meliputi klasifikasi sentimen multidomain (*multi-domain sentiment classification*) dan klasifikasi sentimen multimodal (*multimodal sentiment classification*) [15].

3.2. XLNet

XLNet merupakan model *deep learning* yang berbasis pada tujuan pemodelan bahasa permutasi umum. XLNet dikembangkan oleh *Carnegie Mellon University* dan peneliti Google pada tahun 2019 [16]. Model XLNet dirancang untuk mencari solusi kepada kelemahan metode *autoencoding* yang digunakan pada model BERT dan model populer lainnya. Pemodelan bahasa permutasi bertujuan untuk mengintegrasikan manfaat pemodelan bahasa *autoregressive* (AR) dan *autoencoding* (AE) [16].

Pemodelan bahasa autoregresif umum (XLNet) mempertahankan manfaat model AR dan memungkinkan model untuk menangkap konteks dua arah (*bidirectional*) yang memperbaiki kesenjangan antara AR dan *masked language model*. XLNet adalah model bahasa permutasi yang memaksimalkan kemungkinan log urutan yang diharapkan [17]. Persamaan pemodelan bahasa permutasi ditulis dengan Persamaan (1).

$$\max_{\theta} \mathbb{E}_{z \sim Z_T} \left[\sum_{t=1}^T \log p_{\theta}(x_{z_t} | \mathbf{x}_{z < t}) \right] \quad (1)$$

Z_T merupakan himpunan semua kemungkinan permutasi dari urutan panjang index T [1, 2, . . . T]. z_t merupakan elemen ke- t , $z < t$ merupakan elemen pertama $t - 1$ dari permutasi $z \in Z_T$. Fungsi persamaan tujuan untuk pemodelan bahasa permutasi mengambil token $t - 1$ sebagai konteks dan untuk memprediksi token ke- t sehingga dapat dikatakan bahwa XLNet dapat menangkap pengetahuan ketergantungan konteks yang rumit [17]. Urutan teks x diambil sampel urutan faktorisasi z pada satu waktu dan menguraikan kemungkinan $p_{\theta}(x)$ menurut urutan faktorisasi. Parameter model yang sama θ dibagikan di semua urutan faktorisasi selama pelatihan, dengan harapan x_t telah melihat setiap elemen yang mungkin $x_i \neq x_t$ dalam urutan sehingga dapat menangkap konteks dua arah [16].

3.3. Confusion Matrix

Confusion matrix dapat digunakan untuk mengevaluasi kinerja model deep learning pada tugas klasifikasi citra. Dari hasil Confusion matrix dapat dianalisis dan diidentifikasi keefektifan tugas klasifikasi pada setiap tahap pelatihan dan pengujian model [18]. Rumus yang dilakukan memiliki empat keluaran, yaitu: *Recall*, *Precision*, *Accuracy*, dan *F1 Score* [19].

Tabel 3. *Confusion Matrix*

	Positif Aktual	Negatif Aktual
Prediksi Positif	TP (True Positive)	FP (False Positive)
Prediksi Negatif	FN (False Negative)	TN (True Negative)

Tabel 3 merupakan tabel *Confusion Matrix* yang berisikan *True Positive* (TP), *False Positive* (FP), *True Negative* (TN) dan *False Negative* (FN) untuk mengetahui tingkat kedekatan antara nilai prediksi dengan nilai aktual. *True Positive* yaitu prediksi dan nilai aktualnya positif. *False Positive* yaitu prediksinya positif dan nilai aktualnya negatif. *False Negative* yaitu prediksinya negatif dan nilai aktualnya positif. *True Negative* yaitu prediksi dan nilai aktualnya negatif.

Nilai akurasi adalah tingkat keberhasilan klasifikasi. Tingkat akurasi bisa dihitung dengan Persamaan (2).

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (2)$$

Nilai presisi merupakan data TP (*True Positive*) dari data yang diperkirakan benar dan relevan dengan kebutuhan informasi pengguna. Nilai presisi dihitung dengan Persamaan (3).

$$Precision = \frac{TP}{FP + TP} \tag{3}$$

Nilai *recall* merupakan pengukur keberhasilan suatu klasifikasi yang diprediksi benar. Nilai *recall* digunakan untuk TP (*True Positive*) dengan tupel positif. Nilai *recall* dihitung dengan Persamaan (4).

$$Recall = \frac{TP}{FN + TP} \tag{4}$$

Nilai *F1 Score* merupakan nilai rata-rata harmonis dari *precision* dan *recall*. Nilai *F1 Score* dihitung dengan menggunakan Persamaan (5).

$$F1 - Score = \frac{2 \times (Precision \times Recall)}{Precision + Recall} \tag{5}$$

4. Hasil dan Pembahasan

Hasil dari penelitian yang dilakukan penulis berdasarkan proses gambaran umum penelitian yang dipaparkan sebagai berikut.

4.1 Data Collecting

Data *review brand* lokal diperoleh dengan teknik *scraping* pada API dari masing-masing *e-commerce*, yaitu: Tokopedia, Shopee, Lazada, Bukalapak, Blibli, dan Zalora. Total jumlah data *review* yang diperoleh dari keenam situs *e-commerce* pada periode Februari 2013 hingga Juni 2023 sebanyak 16.663.045 data. Data *review* produk yang diperoleh dari hasil *scraping* kemudian disimpan ke dalam *database* MongoDB.

4.2 Data Labelling

Dataset yang diperoleh dari hasil *scraping* data *review* dari berbagai *e-commerce* diambil sebanyak 23.000 data. Data *review* kemudian diberikan label secara manual. Hasil dari *data labelling* secara manual dapat dilihat pada Gambar 4.

review	sentimen
Suara dan mic puas. Hanya saya pesan yg nyala dan sudah sy pastikan lg klik yg nyala tp yg datang ga nyala	0
Good. Mudah-mudahan awet. Makasih	1
kondisi dus robek2, tolong di perhatikan lagi	0
Kualitas produk baik. Harga terjangkau dan pengirimannya cepet banget. Udh 3x order disini oke lah tp sayangnya klo buat kado gabisa ntip pake kartu ucapan atau dibungkus pake kertas kado :(0
Overall bagus, tombol berfungsi semuanya, tapi packaging kurang karena sudah penyot dan robek sedikit	0
Produk ini original, Dan kualitasnya sangat baik. kecepatan pengiriman sanagt cepat, recommended seller.	1
kualitas sangat jelek. bocor	0
terimakasih banyak ka ya, barang sudah diterima dengan baik. sukses terus ka	1
Body nya lebih berat dari cm 1579, spiral nya lebih tipis dibanding cm 1579, semoga awet, beli CM 1659 penasaran Aja, yang cm 1579 udah mau rusak mesin nya, Harga lebih murah 4k dari cm 1579 Fast delivery, courier Ramah sangat, ty	0
pesanan di terima tanggal 27 January 2021 tetapi baru bisa saya buka pagi ini tanggal 29 January 2021 pukul 06.00 dan ternyata tab SKETSA ADVAN tidak bisa di nyalakan sama sekali dan di cas juga posisi Tab tidak bisa menyala	0
sesuai gambar	1
Perlu ditingkatkan packaginya..masak iya pake clear wrap aja..paling nggak kan dimasukin ke kantong plastik dulu lah baru di wrap..so far tidak ada keluhan untuk produknya..Harga bersaing 🙏🏻	0
Barang bagus, hanya sepertinya tidak di cek dulu ya. Ada baut yg hilang 1 \nJadinya harus cari ke bengkel Pengiriman	0
barang berfungsi dengan baik. kardus nya agak rusak dan penyok penyok akibat pengiriman mungkin padahal untuk	0
bagus banget!!! Suka 🤩🙏🏻	1
Pengiriman sangat cepat cuma 2 hari . barang berpungsi sangat baik . cuma packingnya aja yg kurang aman . untungnya semua barangnya gak ada yg cacat sedikitpun 🙏🏻	0

Gambar 4. Hasil *Dataset* yang Sudah Diberi Label Secara Manual

Jumlah total dataset training yang diberi label secara manual terdiri dari 500 data review dari 46 brand lokal dengan total sebanyak 23.000 data (17.197 data positif dan 5.803 data negatif). Jumlah dataset yang digunakan untuk melatih model XLNet hanya sebanyak 11.606 data (5803 data positif dan 5803 data negatif) agar setiap label dapat mempertahankan proporsi dan distribusi label sentimen yang sama dan seimbang.

4.3 Data Pre-Processing

Proses *data pre-processing* yakni mengubah data *review* yang masih mentah/kotor menjadi data yang bersih dan terstruktur untuk dilakukan proses analisis. Contoh hasil tahap *pre-processing* data dapat dilihat pada Tabel 4.

Tabel 4. Hasil *Data Pre-Processing*

Proses	Teks Review Sebelum Diproses	Teks Review Setelah Diproses
Text Cleaning	Proses cepatttt, pengiriman cepattt, packaging aman 👍👍, produk sdh pembelian kedua... Belum ada perubahan yg signifikan di kulit wajah saya tapi tetap mau mencoba utk botol kedua inih	Proses cepat, pengiriman cepat, packaging aman, produk sdh pembelian kedua. Belum ada perubahan yg signifikan di kulit wajah saya tapi tetap mau mencoba utk botol kedua inih
Lowercase Folding	Proses cepat, pengiriman cepat, packaging aman, produk sdh pembelian kedua. Belum ada perubahan yg signifikan di kulit wajah saya tapi tetap mau mencoba utk botol kedua inih	proses cepat, pengiriman cepat, packaging aman, produk sdh pembelian kedua. belum ada perubahan yg signifikan di kulit wajah saya tapi tetap mau mencoba utk botol kedua inih
Filter Kalimat Review Tidak Relevan	proses cepat, pengiriman cepat, packaging aman, produk sdh pembelian kedua. belum ada perubahan yg signifikan di kulit wajah saya tapi tetap mau mencoba utk botol kedua inih	belum ada perubahan yg signifikan di kulit wajah saya tapi tetap mau mencoba utk botol kedua inih
Text Translate	belum ada perubahan yg signifikan di kulit wajah saya tapi tetap mau mencoba utk botol kedua inih	There hasn't been a significant change in my facial skin but I still want to try this second bottle

Teks review berawal dari data teks yang masih memiliki noise, kalimat yang tidak relevan dengan produk, emoji, dll; setelah melalui tahap *text cleaning*, *lowercase folding*, filter kalimat, dan *text translate* maka data teks review menjadi data yang bersih dan siap digunakan untuk tahap *finetuning* model dan analisis.

4.4 Finetuning Model

Beberapa *hyperparameter* seperti *learning rate*, *epoch training*, *optimizer*, *dropout*, *max sequential length*, *batch size*, dan *data split ratio* dilakukan eksperimen untuk menghasilkan model dengan performa terbaik dan optimal.

Tabel 5. Hasil Evaluasi *Finetuning* Model XLNet

Data Split Ratio	Optimizer	Learning Rate	Dropout	Epoch	Accuracy	
80:10:10	Adam	1e-5	0.1	2	94,91%	
			0.2	4	93,62%	
			0.3	11	93,97%	
			0.4	1	50,47%	
			0.5	1	53,31%	
			0.1	1	94,91%	
		2e-5	0.2	3	94,65%	
			0.3	11	94,14%	
			3e-5	0.1	1	94,65%
				0.2	2	95,60%
				0.3	5	94,57%

Tabel 5 menunjukkan hasil akurasi yang diperoleh dari setiap eksperimen *hyperparameter finetuning*. *Hyperparameter* yang memperoleh hasil akurasi pengujian terbaik sebesar 95,60% adalah *hyperparameter* yang menggunakan *data split ratio* sebesar 80:10:10, *optimizer* Adam, nilai *dropout* sebesar 0.2, nilai *learning rate* sebesar $3e-5$, nilai *batch size* sebesar 32, dan nilai *max sequence length* sebesar 128. Model XLNet dilatih dengan dua *epoch* yang dihentikan secara otomatis dengan fungsi *EarlyStopping*.

4.5 Hasil Inference Model

Hasil *inference* model membahas tentang hasil dari mengklasifikasikan label sentimen pada data *review* dengan model XLNet yang telah dilatih pada proses *finetuning*. Hasil dari proses *inference* model XLNet pada data *review* yang sebelumnya tidak memiliki label sentimen dapat dilihat pada gambar 5.

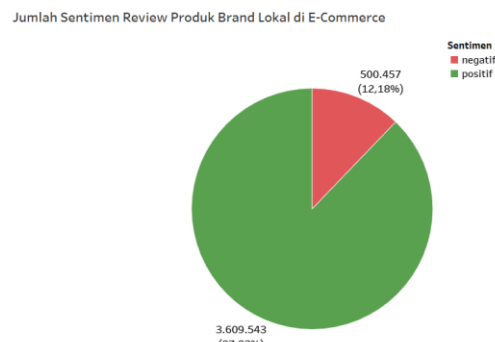
id	review	review_date	product_name	product_brand	ecommerce	review_bersih	translate	sentimen
e0e076c5-ffac-49a5-9b68-b534ed34a80d	good, nyaman untuk cuci piring dan bersih-bers...	2022-10-10T00:00:00.000Z	KENMASTER Sarung Tangan Karet - Orange	Kenmaster	Blibli	good, nyaman untuk cuci piring dan bersih bers...	good, convenient for washing dishes and other ...	positif
04c686b2-21ae-4047-af22-891306e5e2ba	oke. sesuai desk. good seller. good job!	2022-03-14T00:00:00.000Z	KENMASTER Sarung Tangan Karet - Orange	Kenmaster	Blibli	oke. sesuai desk. good job!	Okay, according to the description good job!	positif
16d9381a-3686-4de4-8a86-c474e39cb977	oke	2022-03-01T00:00:00.000Z	KENMASTER Sarung Tangan Karet - Orange	Kenmaster	Blibli	oke	Okay	positif
1059b197-32a3-4e47-838f-05df13aea2a5	Product cocox. tks	2021-10-01T00:00:00.000Z	KENMASTER Sarung Tangan Karet - Orange	Kenmaster	Blibli	product cocox. tks	product cocox. thx	positif
a7caa86b-3f2-4759-8c12-df0278e19fcc	enak dipake, lembut gak tebal banget jd gak kaku	2021-08-07T00:00:00.000Z	KENMASTER Sarung Tangan Karet - Orange	Kenmaster	Blibli	enak dipake, lembut gak tebal banget jd gak kaku	delicious to use, soft not too thick so not stiff	positif

Gambar 5. Hasil Klasifikasi Label Sentimen pada Data Review

Data *review* yang sudah diberi label sentimen terdiri dari field *id*, *review*, *review_date*, *product_name*, *product_brand*, *ecommerce*, *review_bersih*, *translate*, dan *sentimen*. Data *review* yang sudah diberi label atau diklasifikasi disimpan ke dalam database MongoDB.

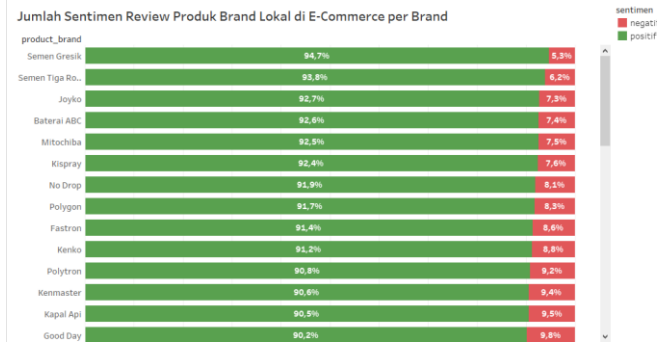
4.6 Hasil Visualisasi Data

Data *review* setelah melalui proses *pre-processing* dan *inference* dengan model XLNet berjumlah sebanyak 4.110.012 data, data tersebut kemudian divisualisasikan untuk menampilkan data dalam bentuk yang mudah dipahami dan dianalisis. Hasil visualisasi data ditampilkan dalam bentuk diagram *bar*, *line*, *pie*, dan *word cloud*.



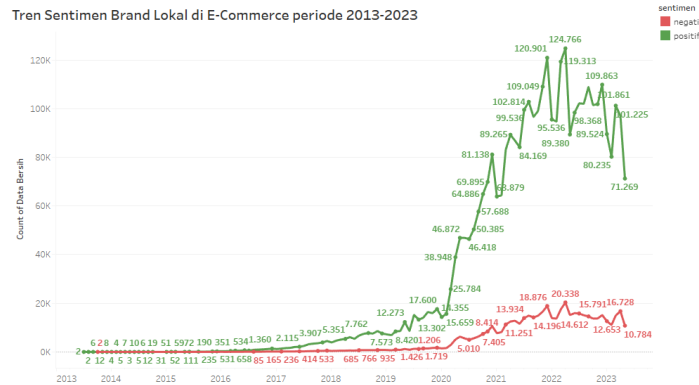
Gambar 6. Pie Chart Perbandingan Sentimen

Gambar 6 menunjukkan visualisasi hasil perbandingan sentimen pada data *review* produk *brand* lokal. Sentimen masyarakat terhadap *brand* lokal didominasi oleh sentimen positif dengan jumlah data *review* sebanyak 3.609.543 *review* dan (87,82%). Sentimen negatif pada *brand* lokal ditemukan lebih sedikit dibandingkan sentimen positif dimana jumlah data *review* negatif sebanyak 500.457 *review* (12,18%).



Gambar 7. Perbandingan Data Review Sentimen Positif dan Negatif dari Setiap Brand

Gambar 7 menunjukkan diagram *bar* yang menampilkan perbandingan jumlah dan persentase antara sentimen positif dengan sentiment negatif masyarakat pada *review* produk dari masing-masing *brand* lokal yang digunakan pada penelitian. Persentase *review* kelas sentimen positif tertinggi terdapat pada *brand* Semen Gresik (positif 94,65%; negatif 5,35%). Sementara itu, persentase *review* kelas sentimen negatif tertinggi terdapat pada *brand* Sepeda Listrik Selis (negatif 21,43%; positif 78,57%).



Gambar 8. Grafik Tren Sentimen Review Seluruh Produk Brand Lokal

Gambar 8 menunjukkan tren sentimen masyarakat terhadap seluruh produk *brand* lokal dari tahun 2013 hingga 2023. Jumlah *review* produk *brand* lokal di situs *e-commerce* yang semakin meningkat dari bulan Maret tahun 2020 hingga seterusnya berhubungan dengan semakin tingginya minat belanja masyarakat dalam berbelanja *online* di situs *e-commerce* ketika masyarakat mengalami keterbatasan kegiatan di luar rumah pada saat pandemi COVID-19.



Gambar 9. Word Cloud Sentimen Review Positif Seluruh Produk Brand Lokal

Gambar 9 menunjukkan visualisasi *word cloud review* positif yang menggambarkan kata-kata yang sering muncul dan berkaitan satu sama lain, seperti: “Really Good”, “Good stuff”, “Good”, “Original product”, “Good performance”, dan lain-lain. Berdasarkan visualisasi *word cloud review* positif, dapat disimpulkan bahwa beberapa pembeli produk *brand* lokal menyukai barang

- [7] S. Farzadnia dan I. Raeesi Vanani, "Identification of opinion trends using sentiment analysis of airlines passengers' reviews," *J Air Transp Manag*, vol. 103, hlm. 102232, Agu 2022, doi: 10.1016/j.jairtraman.2022.102232.
- [8] R. L. Atimi dan Enda Esyudha Pratama, "Implementasi Model Klasifikasi Sentimen Pada Review Produk Lazada Indonesia," *Jurnal Sains dan Informatika*, vol. 8, no. 1, hlm. 88–96, Jul 2022, doi: 10.34128/jsi.v8i1.419.
- [9] E. H. Muktafin, K. Kusriani, dan E. T. Luthfi, "Analisis Sentimen pada Ulasan Pembelian Produk di Marketplace Shopee Menggunakan Pendekatan Natural Language Processing," *Jurnal Eksplora Informatika*, vol. 10, no. 1, hlm. 32–42, Sep 2020, doi: 10.30864/eksplora.v10i1.390.
- [10] D. Bhoi dan A. Thakkar, "Sentiment Analysis Performance and Reliability Evaluation Using an XLNet-based Deep Learning Approach," *Reliability: Theory & Applications*, vol. 17, no. 1 (67), hlm. 391–397, Mar 2022, doi: 10.24412/1932-2321-2022-167-391-397.
- [11] Top Brand Award, "Top Brand Index," 2022. https://www.topbrand-award.com/top-brand-index/?tbi_year=2022 (diakses 16 Juni 2023).
- [12] I. K. Sastrawan, I. P. A. Bayupati, dan D. M. S. Arsa, "Detection of fake news using deep learning CNN–RNN based methods," *ICT Express*, vol. 8, no. 3, hlm. 396–408, Sep 2022, doi: 10.1016/j.icte.2021.10.003.
- [13] I. N. K. Bayu, I. M. A. D. Suarjaya, dan P. W. Buana, "Classification of Indonesian Population's Level Happiness on Twitter Data Using N-Gram, Naïve Bayes, and Big Data Technology," *Int J Adv Sci Eng Inf Technol*, vol. 12, no. 5, hlm. 1944, Okt 2022, doi: 10.18517/ijaseit.12.5.14387.
- [14] D. A. Savita, I. K. G. D. Putra, dan N. K. D. Rusjyanthi, "Public Sentiment Analysis of Online Transportation in Indonesia through Social Media Using Google Machine Learning," *Jurnal Ilmiah Merpati (Menara Penelitian Akademika Teknologi Informasi)*, vol. 9, no. 2, hlm. 153–164, Apr 2021, doi: 10.24843/JIM.2021.v09.i02.p06.
- [15] A. Yadav dan D. K. Vishwakarma, "Sentiment analysis using deep learning architectures: a review," *Artif Intell Rev*, vol. 53, no. 6, hlm. 4335–4385, Agu 2020, doi: 10.1007/s10462-019-09794-5.
- [16] Z. Yang, Z. Dai, Y. Yang, J. Carbonell, R. Salakhutdinov, dan Q. V. Le, "XLNet: Generalized Autoregressive Pretraining for Language Understanding," *Adv Neural Inf Process Syst*, vol. 32, Jun 2019, [Daring]. Tersedia pada: <http://arxiv.org/abs/1906.08237>
- [17] X.-R. Gong, J.-X. Jin, dan T. Zhang, "Sentiment Analysis Using Autoregressive Language Modeling and Broad Learning System," dalam *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, IEEE, Nov 2019, hlm. 1130–1134. doi: 10.1109/BIBM47256.2019.8983025.
- [18] I. P. B. G. Prasetyo Raharja, I. M. Suwija Putra, dan T. Le, "Kekarangan Balinese Carving Classification Using Gabor Convolutional Neural Network," *Lontar Komputer : Jurnal Ilmiah Teknologi Informasi*, vol. 13, no. 1, hlm. 1, Apr 2022, doi: 10.24843/LKJITI.2022.v13.i01.p01.
- [19] I. Ramadhani dan W. Maharani, "Predicting Depressive Disorder Based on DASS-42 on Twitter Using XLNet's Pretrained Model Classification Text," *Journal of Computer System and Informatics (JoSYC)*, vol. 3, no. 4, hlm. 379–385, Sep 2022, doi: 10.47065/josyc.v3i4.2157.