Journal of Applied Informatics and Computing (JAIC)
Vol.10, No.1, February 2026, pp. 630~639
e-ISSN: 2548-6861

Comparison of Random Forest and LSTM for Tokopedia Sentiment Analysis

Fahrizal Denta Saputra [1]*, Fikri Budiman [2]*
* Teknik Informatika, Universitas Dian Nuswantoro
111202214059@mhs.dinus.ac.id [1], fikri.budiman@dsn.dinus.ac.id [2]

Article Info

Article history:

Received 2025-12-16
Revised 2025-12-27
Accepted 2026-01-08

Keyword:

Tokopedia,
Sentiment Analysis,
Random Forest,
LSTM,
E-commerce.

ABSTRACT

Tokopedia is one of the largest *e-commerce* platforms in Indonesia, where every transaction generates user reviews containing opinions about the products or services received. These reviews provide important information about product quality, but the very large quantity makes manual analysis inefficient. This study aims to automatically classify Tokopedia review sentiment and compare the performance of machine learning and deep learning methods. The dataset used was obtained from Kaggle and has undergone an initial cleaning stage, including removing irrelevant columns and manually labeling into two sentiment classes, positive and negative. The research methodology includes several stages, namely data preprocessing (cleaning, case-folding, stopword removal, tokenization, normalization, and stemming), feature extraction using TF-IDF for Random Forest and word embedding for LSTM, implementation of Random Forest and Long Short-Term Memory (LSTM) models, and model evaluation using confusion matrix. Experimental results show that LSTM provides the best performance with 94% accuracy, while Random Forest achieves 92% accuracy. These findings indicate that LSTM is more effective in understanding language context, resulting in more accurate sentiment classification and is useful for decision making in the *e-commerce* field.

I. PENDAHULUAN

Kemajuan teknologi digital mendorong pertumbuhan pesat perdagangan di Indonesia melalui hadirnya berbagai platform *e-commerce*. Peningkatan penggunaan *e-commerce* tidak terlepas dari tingginya akses internet oleh masyarakat. Berdasarkan hasil survei dari Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) mencatat bahwa pada tahun 2024 jumlah pengguna internet di Indonesia mencapai sekitar 221 juta jiwa dari total populasi 278 juta jiwa [1]. Tingkat penetrasi internet nasional pada periode tersebut tercatat sebesar 79,5%, mengalami kenaikan 1,4% dibandingkan tahun sebelumnya [1]. Peningkatan jumlah pengguna internet berkontribusi langsung terhadap bertambahnya basis pelanggan pada berbagai platform *e-commerce* di Indonesia. Hal ini sejalan dengan temuan penelitian terdahulu yang menekankan bahwa penetrasi internet menjadi salah satu faktor utama dalam mendorong pertumbuhan transaksi digital dan aktivitas belanja daring [2][3].

Meskipun pertumbuhan *e-commerce* di Indonesia menunjukkan perkembangan pesat, proses pembelian daring masih menghadapi tantangan, terutama terkait ketidakpastian kualitas produk yang diterima konsumen [4][5]. Tokopedia, sebagai salah satu platform *e-commerce* terbesar di Indonesia, memfasilitasi transaksi jual beli barang maupun jasa, di mana ulasan pelanggan menjadi elemen penting dalam mempengaruhi keputusan pembelian [6]. Ulasan produk memiliki peran signifikan dalam membentuk persepsi konsumen, seiring dengan meningkatnya pemanfaatan media digital dan sosial [7][8]. Mengingat ulasan berisi opini positif maupun negatif yang dapat mempengaruhi calon konsumen, diperlukan metode otomatis untuk mengklasifikasikannya, karena penyortiran manual memerlukan waktu yang cukup lama [9].

Dalam konteks pertumbuhan *e-commerce*, analisis sentimen berperan penting untuk memahami opini konsumen yang terkandung dalam ulasan produk maupun interaksi di media sosial. Ulasan konsumen sering kali memuat opini

positif, negatif, maupun netral yang memberikan gambaran mengenai kepuasan, kualitas produk, maupun pelayanan dari penjual. Metode analisis sentimen memungkinkan transformasi informasi tidak terstruktur, seperti ulasan di forum, blog, maupun media sosial, menjadi data terstruktur yang lebih mudah dianalisis dan diproses secara otomatis [10]. Dengan demikian, analisis sentimen membantu pengguna *e-commerce* dan pelaku bisnis untuk memperoleh wawasan yang lebih jelas terkait persepsi konsumen, mendukung pengambilan keputusan strategis, serta memfasilitasi perbaikan kualitas layanan dan produk [11][12].

Penerapan analisis sentimen tidak hanya bergantung pada pemrosesan bahasa alami, tetapi juga memerlukan metode klasifikasi yang mampu membedakan opini positif dan negatif secara akurat. Berbagai pendekatan telah digunakan, termasuk algoritma machine learning tradisional seperti Random Forest, yang dikenal efektif dalam menangani data berukuran besar dan kompleks [13][14]. Di sisi lain, perkembangan deep learning menghadirkan model seperti Long Short-Term Memory (LSTM), yang memiliki kemampuan memahami urutan kata secara kontekstual dan mempertahankan informasi penting dari teks panjang melalui mekanisme memori jangka panjang [15][16]. Dengan karakteristik tersebut, Random Forest dan LSTM mewakili dua pendekatan yang berbeda dalam pemrosesan data teks. Random Forest merupakan algoritma berbasis pohon keputusan yang memiliki kemampuan generalisasi tinggi terhadap berbagai jenis data, sedangkan LSTM merupakan jaringan saraf yang unggul dalam memahami konteks bahasa alami dan menangkap hubungan antar-kata secara mendalam.

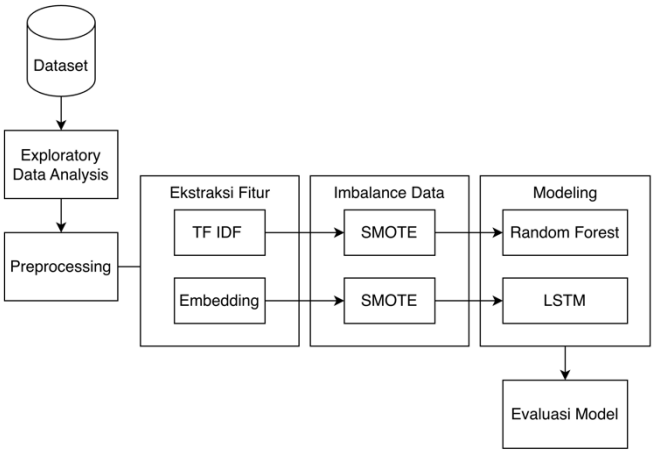
Walaupun Random Forest dan LSTM telah banyak digunakan dalam analisis sentimen, sebagian besar penelitian terdahulu masih terbatas pada penerapan salah satu metode saja. Beberapa penelitian berikut menggambarkan hasil penerapan algoritma tersebut dalam berbagai konteks analisis sentimen. Beberapa studi menyoroti penggunaan Random Forest pada ulasan Tokopedia di Play Store dengan akurasi 93% dan F1-Score 96%, serta mengidentifikasi kualitas produk dan kinerja aplikasi sebagai masalah utama [17]. Studi lain menganalisis sentimen Twitter tentang Tokopedia menggunakan Naïve Bayes dan Random Forest, dengan penerapan SMOTE meningkatkan akurasi menjadi 86,93% pada Naïve Bayes dan 88,44% pada Random Forest [18]. Penelitian Analisis Sentimen Berbasis Aspek (ABSA) pada ulasan *e-commerce* Indonesia membandingkan CNN dan LSTM, menunjukkan LSTM unggul dengan akurasi 86,10% dan F1-score 85,80% [19]. Studi terkait ulasan dan rating *e-commerce* selama pandemi COVID-19 membandingkan SVM, BNB, LR, CNN, dan LSTM, dengan LSTM mencapai akurasi tertinggi 94% [20]. Selain itu, analisis sentimen publik Pilkada 2024 di Twitter/X menunjukkan LSTM memiliki akurasi 94,21%, mengungguli Random Forest 92,78% dan Naive Bayes 89,43%, dengan sentimen negatif dominan terkait kebijakan dan kinerja kandidat [21].

Meskipun berbagai penelitian sebelumnya menerapkan Random Forest maupun LSTM dalam analisis sentimen, sebagian besar studi tersebut masih berfokus pada satu pendekatan atau konteks dataset tertentu. Penelitian yang secara langsung membandingkan metode machine learning dan deep learning pada ulasan *e-commerce* Indonesia, khususnya pada platform Tokopedia, masih relatif terbatas. Berbeda dengan penelitian terdahulu, penelitian ini secara khusus melakukan perbandingan langsung antara Random Forest dan LSTM pada dataset ulasan Tokopedia dengan skema evaluasi yang sama, sehingga perbedaan kinerja kedua pendekatan dapat dianalisis secara lebih objektif.

Berdasarkan celah penelitian tersebut, rumusan masalah dalam penelitian ini adalah bagaimana performa algoritma Random Forest dan Long Short-Term Memory (LSTM) dalam mengklasifikasikan ulasan produk *e-commerce* Tokopedia serta metode mana yang lebih efektif dalam menangkap opini konsumen. Penelitian ini bertujuan untuk membandingkan kinerja kedua algoritma tersebut melalui penerapan ekstraksi fitur TF-IDF pada Random Forest dan Embedding Layer pada LSTM, dengan evaluasi performa menggunakan metrik akurasi, precision, recall, dan F1-score. Manfaat penelitian ini diharapkan dapat membantu pelaku *e-commerce* dalam memahami opini konsumen secara lebih cepat dan akurat, mendukung pengambilan keputusan strategis, serta menjadi dasar pengembangan sistem analisis sentimen yang lebih efektif.

II. METODE

Tahapan penelitian meliputi dataset, Exploratory Data Analyst, preprocessing, ekstraksi fitur, imbalance data, modeling dan evaluasi model. Secara keseluruhan, alur penelitian dapat dilihat pada Gambar 1.



Gambar 1. Metode Penelitian

A. Dataset

Dataset yang digunakan dalam penelitian ini berasal dari platform Kaggle dan berisi 5.400 baris data ulasan konsumen terhadap berbagai produk di Tokopedia. Dataset ini memiliki 11 kolom, antara lain Category, Product Name, Location,

Price, Overall Rating, Number Sold, Total Review, Customer Rating, Customer Review, Sentiment, dan Emotion. Namun, penelitian ini hanya berfokus pada dua kolom utama, yaitu Customer Review dan Customer Rating. Kolom Customer Review berisi teks ulasan yang menggambarkan pengalaman pengguna dalam bertransaksi, sedangkan Customer Rating merupakan nilai numerik (skala 1–5) yang mencerminkan tingkat kepuasan pelanggan terhadap produk atau layanan.

B. Exploratory Data Analyst

Tahap Exploratory Data Analysis (EDA) dilakukan untuk memperoleh pemahaman awal terhadap karakteristik dataset sebelum preprocessing dan pemodelan. Proses ini mencakup pemeriksaan struktur data, jumlah baris dan kolom, tipe data, serta pengecekan nilai kosong atau duplikat. Analisis statistik deskriptif dilakukan untuk mengetahui sebaran data dan jumlah nilai unik pada setiap atribut, termasuk distribusi Customer Rating dalam rentang 1 hingga 5, yang divisualisasikan menggunakan diagram batang. Hasil EDA menunjukkan adanya ketidakseimbangan kelas, di mana ulasan dengan rating tinggi atau sentimen positif lebih dominan dibandingkan ulasan negatif, serta beberapa data duplikat yang perlu dibersihkan. Selain itu, penelitian ini melakukan proses pelabelan sentimen secara manual berdasarkan nilai Customer Rating, dengan ulasan rating 1–3 dikategorikan sebagai sentimen negatif (0) dan rating 4–5 dikategorikan sebagai sentimen positif (1).

C. Preprocessing Data

Pada penelitian ini, tahap preprocessing dilakukan untuk memastikan data teks dalam kondisi siap olah sebelum ekstraksi fitur dan pemodelan. Tahap ini bertujuan membersihkan, menstandarkan, dan menyiapkan struktur data agar dapat diproses dengan baik oleh model machine learning maupun deep learning. Adapun tahapan preprocessing adalah sebagai berikut:

1) *Cleaning*: Proses penghapusan tanda baca, angka, dan karakter yang tidak relevan.
2) *Case Folding*: Mengubah setiap kata menjadi huruf kecil untuk menjadikan teks seragam.
3) *Stopword Removal*: Proses penghapusan kata-kata umum yang tidak penting.
4) *Stemming*: Proses mengubah setiap kata menjadi bentuk dasar.
5) *Tokenizing*: Proses pemecahan teks menjadi unit kata atau token.
6) *Sequence*: Proses mengubah kata (token) menjadi urutan angka.
7) *Padding*: Proses menyamakan panjang input teks agar sesuai dengan kebutuhan model.

Langkah-langkah ini membantu mengurangi kebisingan data dan membuat model lebih fokus pada kata-kata penting yang benar-benar berpengaruh terhadap hasil klasifikasi [22]. Untuk tahapan tokenizing, sequence, dan padding hanya diterapkan pada metode LSTM, karena model ini membutuhkan input teks dalam bentuk urutan angka dengan panjang yang seragam agar proses pembelajaran sekuensial dapat berlangsung secara optimal. Setelah melalui seluruh tahapan tersebut, data teks menjadi lebih bersih, terstandarisasi, dan siap digunakan dalam proses ekstraksi fitur dan pemodelan.

D. Ekstraksi Fitur

Penelitian ini menggunakan dua jenis ekstraksi fitur, yaitu TF-IDF dan word embedding. Pemilihan metode ekstraksi fitur disesuaikan dengan algoritma yang digunakan, dengan rincian sebagai berikut:

1) *TF-IDF*: Pada model machine learning, tahap ekstraksi fitur bertujuan untuk mengubah data teks mentah menjadi representasi numerik agar dapat diproses oleh algoritma klasifikasi. Metode yang digunakan adalah TF-IDF (Term Frequency–Inverse Document Frequency), yang memberikan bobot pada kata berdasarkan dua komponen utama, yaitu Term Frequency (TF) dan Inverse Document Frequency (IDF). TF mengukur frekuensi kemunculan suatu kata dalam dokumen, di mana semakin sering kata tersebut muncul maka semakin besar nilai TF-nya, sebagaimana ditunjukkan pada persamaan (1). Sementara itu, IDF berfungsi untuk mengukur tingkat keunikan kata dalam seluruh korpus dokumen, di mana kata yang jarang muncul akan memiliki nilai IDF yang lebih tinggi, sebagaimana ditunjukkan pada persamaan (2). Bobot akhir TF-IDF diperoleh dari hasil perkalian nilai TF dan IDF sesuai dengan persamaan (3).

$$TF_{t,d} = \frac{frekuensi\ kata\ t\ pada\ dokumen\ d}{Jumlah\ kata\ dalam\ dokumen\ d} \quad (1)$$

$$IDF_t = log\ \frac{N}{n_t} \quad (2)$$

$$TF\text{-}IDF_{t,d} = TF_{t,d} \times IDF_t \quad (3)$$

dengan t kata (term) yang dihitung, d dokumen tempat kata muncul, $TF_{t,d}$ nilai Term Frequency dari kata t pada dokumen d, IDF_t nilai Inverse Document Frequency dari kata t, $TF\text{-}IDF_{t,d}$ bobot akhir TF-IDF dari kata pada dokumen, N jumlah total dokumen, n_t jumlah dokumen yang mengandung kata t.

Kata yang sering muncul dalam satu dokumen tetapi jarang di dokumen lain akan memiliki bobot TF-IDF tinggi, membantu model mengenali kata paling relevan untuk klasifikasi sentimen ulasan Tokopedia [23].

2) *Embedding*: Sementara itu, model berbasis deep learning seperti Long Short-Term Memory (LSTM) menggunakan representasi teks berupa word embedding. Pada penelitian ini, embedding dibangun menggunakan Embedding layer bawaan pada framework Keras dan dipelajari secara langsung dari data pelatihan. Embedding merepresentasikan kata dalam vektor berdimensi tetap yang menangkap makna semantik serta hubungan antar kata dalam konteks kalimat. Berbeda dengan TF-IDF yang

hanya mempertimbangkan frekuensi kata, embedding memungkinkan LSTM memahami konteks dan keterkaitan kata secara lebih efektif dalam analisis sentimen [24]. Metode ini dipilih karena mampu mempertahankan informasi urutan kata dan makna semantik, sehingga model dapat menangkap nuansa bahasa dengan lebih baik. Pendekatan ini sejalan dengan penelitian sebelumnya yang membandingkan machine learning (Naive Bayes dan SVM) menggunakan ekstraksi fitur TF-IDF dan deep learning (LSTM) menggunakan embedding [25].

E. Imbalance Data

Untuk mengatasi ketidakseimbangan data, diterapkan teknik resampling menggunakan SMOTE (Synthetic Minority Over-sampling Technique). Metode ini menambah jumlah data pada kelas minoritas agar distribusi kelas lebih seimbang. SMOTE dipilih karena populer dan efektif, dengan cara menghasilkan data sintetis baru melalui interpolasi antar sampel pada kelas minoritas di ruang fitur, sehingga meningkatkan kualitas pelatihan model dibandingkan oversampling acak biasa. Pendekatan ini bertujuan menyeimbangkan distribusi kelas sehingga model dapat mempelajari pola dari kelas positif dan negatif secara lebih proporsional serta mengurangi bias terhadap kelas mayoritas, sejalan dengan penelitian sebelumnya yang juga menggunakan SMOTE [25].

F. Modeling

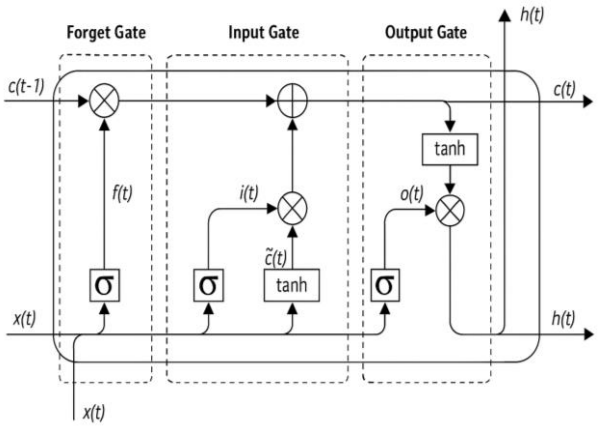
Pada tahap pemodelan, algoritma digunakan untuk mengklasifikasikan sentimen ulasan pelanggan Tokopedia berdasarkan fitur yang telah diekstraksi. Penelitian ini menerapkan dua pendekatan yang berbeda, yaitu :

1) *Random Forest*: Metode Random Forest merupakan algoritma klasifikasi berbasis *ensemble* yang dikembangkan dari Decision Tree. Metode ini membangun sejumlah pohon keputusan dengan pemilihan atribut secara acak pada setiap node [26]. Setiap pohon keputusan dibangun dari *root node* hingga *leaf node*, dan hasil klasifikasi ditentukan berdasarkan mekanisme *majority voting* dari seluruh pohon yang terbentuk. Strategi penggabungan beberapa decision tree ini membuat Random Forest mampu meningkatkan akurasi serta menghasilkan model yang lebih stabil dan robust terhadap *overfitting* [27]. Secara umum, proses prediksi Random Forest dapat dinyatakan pada persamaan (4).

$$f(x) = Average(f_1(x), f_2(x), \ldots, f_n(x)) \quad (4)$$

dimana $f(x)$ hasil prediksi, $f_1(x), f_2(x), \ldots, f_n(x)$ hasil prediksi dari setiap pohon keputusan, dan x inputan data.

2) *Long Short-Term Memory (LSTM)*: Metode Long Short-Term Memory (LSTM) merupakan salah satu algoritma deep learning yang dikembangkan dari arsitektur Recurrent Neural Network (RNN) dan dirancang untuk mengatasi permasalahan long-term dependency pada data berurutan. LSTM memiliki struktur khusus berupa memory cell yang memungkinkan model menyimpan dan mempertahankan informasi dalam jangka waktu yang panjang, sehingga efektif digunakan pada berbagai aplikasi Natural Language Processing (NLP), termasuk analisis sentimen [28]. Setiap unit LSTM terdiri dari tiga gerbang utama, yaitu *input gate*, *forget gate*, dan *output gate*, yang berfungsi mengatur aliran informasi ke dalam dan ke luar *memory cell*. Pada setiap waktu *t*, LSTM menerima input berupa vektor *xt* dan keadaan tersembunyi sebelumnya *ht-1*, kemudian menghasilkan keadaan sel *ct* dan keadaan tersembunyi baru *ht*. Mekanisme ini memungkinkan LSTM mempertahankan informasi yang relevan dan mengabaikan informasi yang tidak penting, sehingga menghasilkan prediksi yang lebih stabil dan akurat pada data teks berurutan. Struktur arsitektur LSTM dilihat pada Gambar 2.



Gambar 2. Arsitektur LSTM

G. Evaluasi Model

Evaluasi ini bertujuan untuk membandingkan kinerja kedua model berdasarkan hasil pengujian terhadap data uji yang telah disiapkan sebelumnya. Hasil evaluasi tersebut kemudian menjadi dasar dalam menentukan model yang paling optimal untuk analisis sentimen pada ulasan pelanggan Tokopedia. Evaluasi model dilakukan untuk menilai tingkat keandalan dan performa model dalam mengklasifikasikan sentimen ulasan pelanggan. Pada penelitian ini, evaluasi diterapkan pada dua jenis model, yaitu machine learning (Random Forest) dan deep learning (LSTM). Proses evaluasi dilakukan dengan membandingkan hasil prediksi model dan label aktual menggunakan Confusion Matrix, yang merepresentasikan kinerja klasifikasi melalui empat komponen utama, yaitu True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN).

Tahap evaluasi dilakukan dengan menghitung beberapa metrik kinerja, yaitu Accuracy, Precision, Recall, dan F1-Score. Accuracy merepresentasikan perbandingan antara jumlah prediksi yang sesuai dengan label sebenarnya terhadap keseluruhan data pengujian, sebagaimana dirumuskan pada persamaan (5), sehingga mencerminkan tingkat ketepatan model secara umum. Precision mengukur tingkat ketepatan

prediksi positif, yakni rasio antara jumlah True Positive terhadap seluruh hasil prediksi positif, seperti ditunjukkan pada persamaan (6). Recall menggambarkan kemampuan model dalam mengenali data positif dengan benar, yaitu perbandingan antara jumlah True Positive dan total data positif yang ada, sebagaimana diformulasikan pada persamaan (7). Adapun F1-Score merupakan ukuran gabungan yang diperoleh dari rata-rata harmonis antara Precision dan Recall, sebagaimana dijelaskan pada persamaan (8).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\,\% \quad (5)$$

$$Precision = \frac{TP}{TP + FP} \times 100\% \quad (6)$$

$$Recall = \frac{TP}{TP + FN} \times 100\% \quad (7)$$

$$F1score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \times 100\% \quad (8)$$

dengan TP merupakan True Positive, TN merupakan True Negative, FP merupakan False Positive, dan FN merupakan False Negative.

III. HASIL DAN PEMBAHASAN

A. Dataset

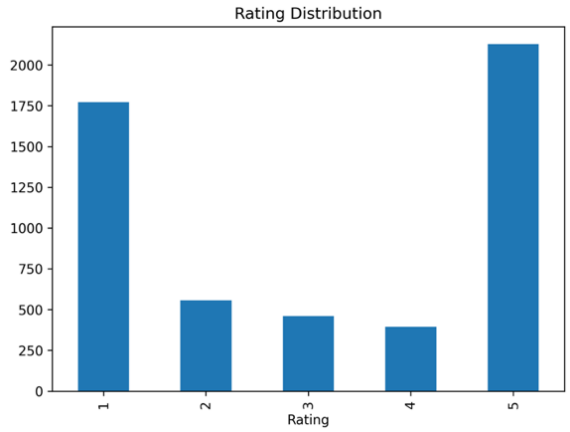
Dataset yang digunakan dalam penelitian ini adalah PRDECT-ID Dataset, yang berisi 5.400 ulasan produk. Struktur lengkap dataset telah dibahas pada bab sebelumnya. Penelitian ini melakukan penyederhanaan kolom pada dataset dengan mempertahankan dua kolom utama, yaitu Customer Review dan Customer Rating. Untuk Customer Review berisi teks ulasan yang menjadi input utama dalam proses preprocessing dan ekstraksi fitur, sedangkan Customer Rating digunakan sebagai dasar pemberian label sentimen dalam dua kategori, yaitu positif dan negatif. Pemilihan kedua kolom ini didasarkan pada hubungan langsung antara opini pelanggan dan penilaian yang diberikan, sehingga lebih representatif untuk klasifikasi sentimen. Dengan demikian, meskipun beberapa informasi lain pada dataset tidak digunakan, fokus pada kedua kolom ini membuat analisis lebih jelas dan terfokus. Contoh potongan dataset yang memuat kedua kolom dapat dilihat pada Gambar 3.

	Customer Rating	Customer Review
0	5	Alhamdulillah berfungsi dengan baik. Packaging...
1	5	barang bagus dan respon cepat, harga bersaing ...
2	5	barang bagus, berfungsi dengan baik, seler ram...
3	5	bagus sesuai harapan penjual nya juga ramah. t...
4	5	Barang Bagus, pengemasan Aman, dapat Berfungsi...

Gambar 3. Struktur Dataset yang Digunakan

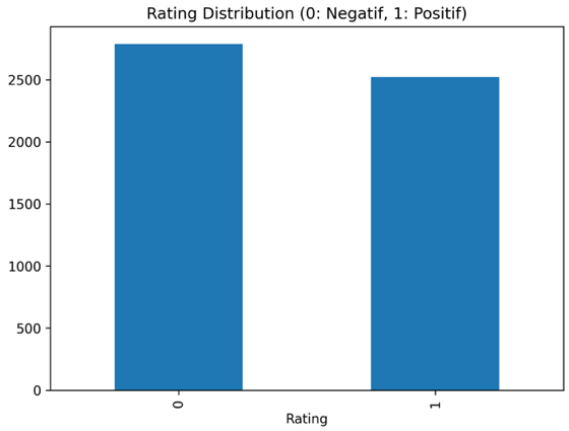
B. Exploratory Data Analyst

Pada tahap Exploratory Data Analysis (EDA) dilakukan pemeriksaan awal untuk memahami karakteristik dataset. Analisis meliputi peninjauan struktur data serta pengecekan format dua kolom utama, yaitu Customer Review dan Customer Rating, termasuk keberadaan data duplikat dan nilai kosong. Hasil EDA menunjukkan terdapat 88 data duplikat dan tidak ditemukan nilai kosong, sehingga data duplikat dihapus untuk menjaga kualitas dataset. Selain itu, analisis distribusi rating menunjukkan bahwa rating 5 mendominasi dengan sekitar 2.129 ulasan, diikuti rating 1 sebanyak 1.772 ulasan. Rating 2 dan 3 masing-masing berjumlah sekitar 557 dan 460 ulasan, sementara rating 4 memiliki jumlah paling sedikit, yaitu sekitar 394 ulasan. Pola ini mengindikasikan adanya ketidakseimbangan kelas pada data ulasan, yang visualisasinya dapat dilihat pada Gambar 4.



Gambar 4. Distribusi Rating

Berdasarkan distribusi rating pada Gambar 4, data tersebut digunakan sebagai dasar pelabelan ulang sentimen. Nilai rating dikonversi menjadi dua kategori, yaitu rating 1–3 sebagai sentimen negatif (0) yang merepresentasikan ketidakpuasan pelanggan, serta rating 4–5 sebagai sentimen positif (1) yang mencerminkan kepuasan pelanggan. Pelabelan ulang ini dilakukan untuk menyederhanakan klasifikasi sentimen menjadi dua kelas utama agar sesuai dengan tujuan penelitian dan memudahkan penerapan serta evaluasi model. Hasilnya diperoleh bahwa terdapat 2.789 data yang berlabel negatif dan 2.523 data berlabel positif, sebagaimana ditunjukkan pada Gambar 5. Dengan demikian, tahap EDA berperan penting tidak hanya dalam memahami distribusi data, tetapi juga dalam menetapkan skema pelabelan sentimen yang konsisten.



Gambar 5. Distribusi Pelabelan Rating

C. *Preprocessing Data*

Pada tahap preprocessing, data ulasan dipersiapkan sebelum dilakukan ekstraksi fitur menggunakan TF-IDF. Proses ini mencakup beberapa tahapan, yaitu cleaning, case folding, tokenizing, stopword removal, serta stemming dengan memanfaatkan pustaka Sastrawi. Tahapan tersebut bertujuan untuk mengurangi noise dan variasi kata yang tidak relevan, sehingga teks menjadi lebih bersih dan konsisten. Melalui tokenisasi dan normalisasi, representasi teks dapat disederhanakan sehingga membantu model Random Forest dalam mempelajari pola sentimen secara lebih efektif. Contoh hasil preprocessing ditampilkan pada Tabel 1.

TABEL I
HASIL PREPROCESSING

Proses	**Hasil**
Text Awal	Barang sesuai pesanan, kondisi berfungsi dengan baik. Seller responsif proses kirim cepat. Recommended!!!
Cleaning	Barang sesuai pesanan kondisi berfungsi dengan baik Seller responsif proses kirim cepat Recommended
Case Folding	barang sesuai pesanan kondisi berfungsi dengan baik seller responsif proses kirim cepat recommended
Tokenizing	[‘barang’, ‘sesuai’, ‘pesanan’, ‘kondisi’, ‘berfungsi’, ‘dengan’,’baik’, ‘seller’, ‘responsif’, ‘proses’, ‘kirim’, ‘cepat’, ‘recommended’]
Stopword Removal	[‘barang’, ‘sesuai’, ‘pesanan’, ‘kondisi’, ‘berfungsi’, ‘baik’, ‘seller’, ‘responsif’, ‘proses’, ‘kirim’, ‘cepat’, ‘recommended’]
Stemming	[‘barang’, ‘sesuai’, ‘pesan’, ‘kondisi’, ‘fungsi’, ‘baik’, ‘seller’, ‘responsif’, ‘proses’, ‘kirim’, ‘cepat’, ‘recommended’]
Hasil Akhir	barang sesuai pesan kondisi fungsi baik seller responsif proses kirim cepat recommended

Berbeda dengan model Random Forest yang menggunakan representasi fitur berbasis TF-IDF, model LSTM memerlukan tahapan preprocessing tambahan karena memproses data dalam bentuk sekuens. Oleh karena itu, pada model LSTM dilakukan tahap tambahan berupa konversi teks menjadi sekuens numerik melalui proses tokenisasi, di mana setiap kata dipetakan ke dalam indeks numerik berdasarkan kamus (vocabulary) yang dibangun selama proses tokenisasi dengan jumlah maksimum 5.000 kata. Selanjutnya, sekuens tersebut dilakukan padding hingga mencapai panjang seragam sebesar 100 token (MAX_LEN = 100) agar seluruh input memiliki dimensi yang sama dan dapat diproses oleh jaringan LSTM. Setiap indeks kata kemudian diubah menjadi vektor embedding berdimensi 128, sehingga kata direpresentasikan dalam ruang vektor numerik yang mampu menangkap hubungan semantik antar kata serta konteks urutan kata sebagai masukan bagi layer LSTM.

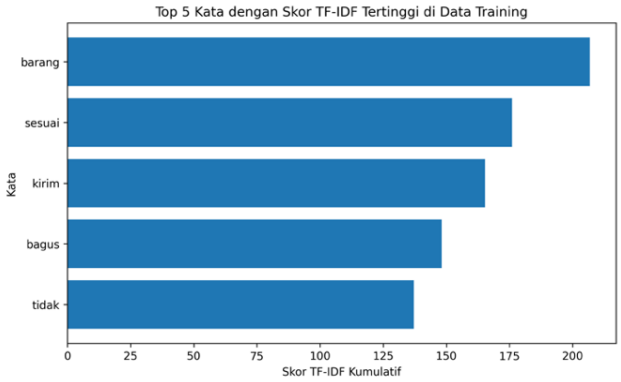
TABEL II
HASIL PREPROCESSING TAMBAHAN LSTM

Proses	**Hasil**
Tokenizing	[‘barang’, ‘sesuai’, ‘pesanan’, ‘kondisi’, ‘berfungsi’, ‘dengan’,’baik’, ‘seller’, ‘responsif’, ‘proses’, ‘kirim’, ‘cepat’, ‘recommended’]
Sequence (Word Index)	[1, 4, 16, 168, 50, 13, 22, 19, 244, 58, 2, 10, 94]
Padding (MAX_LEN= 100)	[1, 4, 16, 168, 50, 13, 22, 19, 244, 58, 2, 10, 94, 0, 0, 0, 0, 0, …. sampai total 100 elemen]
Embedding (dimensi 128)	[[-0.0015, -0.0717, 0.0263, ..., -0.0122], ...]

Dengan demikian, preprocessing pada model LSTM tidak hanya berfokus pada pembersihan dan normalisasi teks, tetapi juga mencakup transformasi data ke dalam bentuk numerik dan vektor yang sesuai dengan arsitektur deep learning.

D. *Ekstraksi Fitur*

Pada tahap ekstraksi fitur, teks ulasan yang telah melalui proses stemming direpresentasikan ke dalam vektor numerik menggunakan metode TF-IDF dengan bantuan *TfidfVectorizer*. Parameter max_df = 0.5 digunakan untuk mengabaikan kata yang terlalu umum, sedangkan min_df = 2 bertujuan menghilangkan kata yang sangat jarang muncul, sehingga diperoleh matriks fitur TF-IDF yang merepresentasikan tingkat kepentingan setiap kata dalam ulasan. Selain fitur berbasis teks, ditambahkan dua fitur numerik, yaitu panjang ulasan (review_length) dan persentase tanda baca (punct), untuk memperkaya representasi data sebelum dimasukkan ke dalam model Random Forest. Hasil ekstraksi fitur menunjukkan bahwa kata “barang” memiliki skor TF-IDF kumulatif tertinggi, diikuti oleh “sesuai”, “kirim”, “bagus”, dan “tidak”, yang mencerminkan aspek utama dalam ulasan pelanggan dan berperan penting dalam membedakan sentimen positif dan negatif. Visualisasi kata dengan skor TF-IDF tertinggi dapat dilihat pada Gambar 6, yang menunjukkan kata-kata dominan dengan bobot tertinggi dan berkontribusi signifikan dalam proses klasifikasi sentimen.

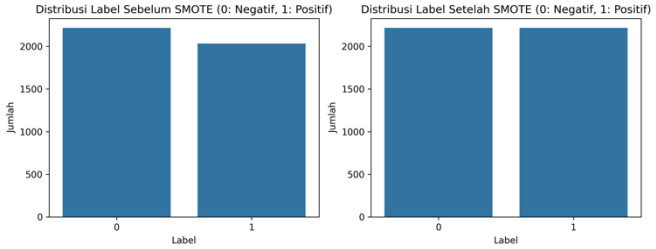


Gambar 6. Distribusi Kata dengan Skor TF-IDF tertinggi

Berbeda dengan model Random Forest yang menggunakan representasi fitur berbasis TF-IDF, model LSTM memerlukan tahapan preprocessing tambahan karena memproses data dalam bentuk sekuens. Oleh karena itu, pada model LSTM dilakukan tahap tambahan berupa konversi teks menjadi sekuens numerik melalui proses tokenisasi, di mana setiap kata dipetakan ke dalam indeks numerik berdasarkan kamus (vocabulary) yang dibangun selama proses tokenisasi dengan jumlah maksimum 5.000 kata. Selanjutnya, sekuens tersebut dilakukan padding hingga mencapai panjang seragam sebesar 100 token (MAX_LEN = 100) agar seluruh input memiliki dimensi yang sama dan dapat diproses oleh jaringan LSTM. Setiap indeks kata kemudian diubah menjadi vektor embedding berdimensi 128, sehingga kata direpresentasikan dalam ruang vektor numerik yang mampu menangkap hubungan semantik antar kata serta konteks urutan kata sebagai masukan bagi layer LSTM.

E. Imbalance Data (SMOTE)

Pada tahapan awal, terlihat bahwa distribusi sentimen pada dataset tidak seimbang, dengan kelas Negatif (0) berjumlah 2.216 sampel dan kelas Positif (1) sebanyak 2.033 sampel. Meskipun perbedaan jumlah sampel antar kelas tidak terlalu besar, kondisi ini dapat dikategorikan sebagai *minor imbalance* yang berpotensi mempengaruhi kinerja model, khususnya dalam memprediksi kelas minoritas. Untuk mengatasi permasalahan tersebut, metode *Synthetic Minority Over-sampling Technique* (SMOTE) pada data pelatihan. SMOTE bekerja dengan menghasilkan sampel sintetis pada kelas minoritas melalui interpolasi berbasis k-nearest neighbors, sehingga distribusi kelas pada training set menjadi lebih seimbang. Penerapan teknik ini dilakukan secara eksklusif pada data pelatihan (x_train_vect dan X_train_dl) untuk mencegah terjadinya kebocoran data (data leakage). Setelah proses SMOTE, jumlah sampel pada kedua kelas menjadi seimbang, masing-masing 2.216 sampel. Dengan demikian, data latih hasil penyeimbangan (x_train_resampled dan X_train_dl_resampled) siap digunakan dalam pelatihan model klasifikasi yang lebih stabil dan tidak bias terhadap salah satu kelas sentimen. Visualisasi distribusi kelas setelah penerapan SMOTE dapat dilihat pada Gambar 7.



Gambar 7. Distribusi Pelabelan Rating Sebelum dan Sesudah SMOTE

F. Modeling

Sebelum tahap pemodelan, dataset dibagi menjadi data pelatihan dan data pengujian dengan perbandingan 80% data latih dan 20% data uji. Selanjutnya, penelitian ini menerapkan dua model klasifikasi, yaitu Random Forest dan Long Short-Term Memory (LSTM), untuk mengklasifikasikan sentimen ulasan ke dalam kategori Negatif dan Positif guna membandingkan kinerja kedua model.

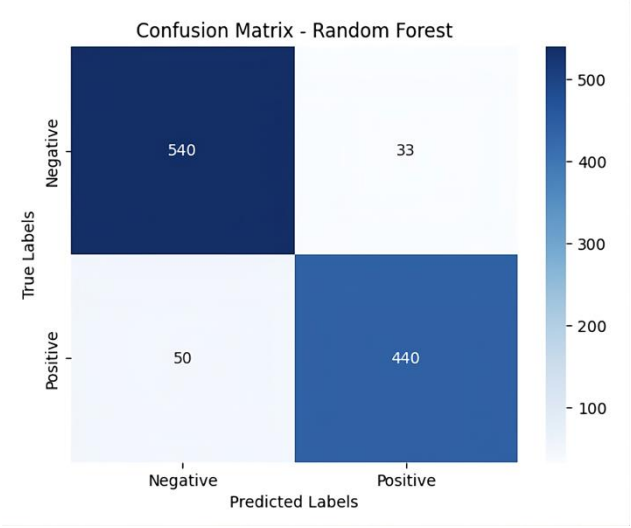
1) *Random Forest*: Model Random Forest diterapkan menggunakan fitur TF-IDF yang telah dihasilkan pada tahap sebelumnya, serta memanfaatkan data pelatihan yang telah diseimbangkan menggunakan SMOTE (x_train_resampled). Model diinisialisasi dengan parameter bawaan dari library *scikit-learn* dan menggunakan random_state = 42 untuk menjaga konsistensi hasil pelatihan. Selain fitur TF-IDF, model juga memanfaatkan fitur numerik tambahan, seperti panjang ulasan (*review_length*) dan persentase tanda baca (*punct*). Setelah proses pelatihan, model diuji menggunakan data uji untuk memprediksi sentimen ulasan, dan kinerjanya dievaluasi menggunakan metrik akurasi, precision, recall, F1-score, serta confusion matrix guna menganalisis pola prediksi yang dihasilkan.
2) *Long Short-Term Memory (LSTM)*: Untuk model LSTM dirancang untuk menangani data dalam bentuk sekuens. Data input berupa sekuens numerik hasil tokenisasi teks ulasan yang telah diproses melalui *padding* hingga panjang seragam sebesar 100 token (MAX_LEN = 100) dan direpresentasikan dalam bentuk vektor embedding berdimensi 128. Model LSTM menggunakan arsitektur yang terdiri dari *SpatialDropout1D* sebesar 0,4, satu layer LSTM dengan 196 unit serta parameter *dropout* dan *recurrent dropout* masing-masing sebesar 0,2, dan diakhiri dengan layer output Dense berfungsi aktivasi sigmoid. Model dilatih menggunakan data pelatihan yang telah diseimbangkan dengan SMOTE (X_train_dl_resampled) selama 7 *epochs* dengan *batch size* 32 dan *validation split* sebesar 0,2. Arsitektur ini memungkinkan LSTM untuk menangkap konteks dan dependensi urutan kata dalam teks ulasan secara efektif.

G. Evaluasi Model

Pada tahap evaluasi model Random Forest dan Long Short-Term Memory (LSTM) diuji menggunakan data uji yang terpisah dari data pelatihan untuk mengukur kemampuan klasifikasi sentimen ulasan Tokopedia ke dalam kelas Positif

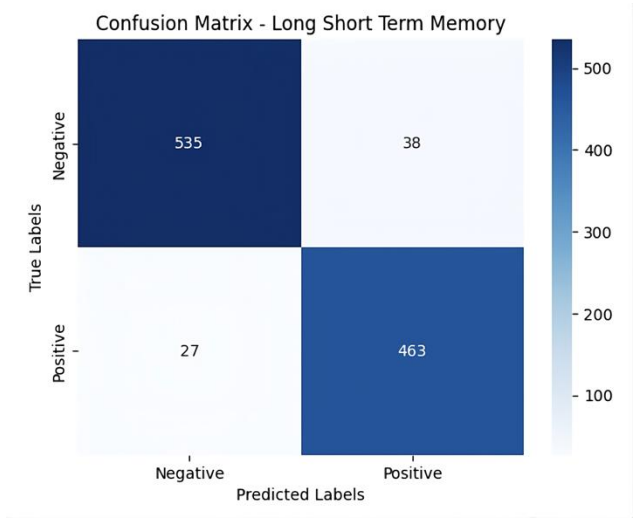
dan Negatif. Untuk memperoleh evaluasi yang lebih komprehensif dan tidak hanya bergantung pada nilai akurasi, penelitian ini menggunakan metrik accuracy, precision, recall, dan F1-score, serta confusion matrix.

1) *Evaluasi Model Random Forest*: Berdasarkan hasil evaluasi, Random Forest menunjukkan kinerja baik dalam mengklasifikasikan sentimen ulasan Tokopedia dengan akurasi 92% pada data uji. Untuk kelas Negatif, model mencapai precision 0,92, recall 0,94, dan F1-score 0,93, sedangkan untuk kelas Positif diperoleh precision 0,93, recall 0,90, dan F1-score 0,91. Confusion matrix menunjukkan bahwa model berhasil mengklasifikasikan 540 ulasan Negatif dan 440 ulasan Positif dengan benar, meski terdapat 33 ulasan Negatif yang salah diklasifikasikan sebagai Positif dan 50 ulasan Positif yang salah sebagai Negatif. Secara keseluruhan, nilai macro average dan weighted average F1-score masing-masing 0,92 menunjukkan bahwa Random Forest memiliki performa stabil dan relatif seimbang pada kedua kelas. Visualisasi hasil klasifikasi dapat dilihat pada Gambar 8.



Gambar 8. Confusion Matrix Random Forest

2) *Evaluasi Model Long Short-Term Memory*: Berdasarkan hasil evaluasi, menunjukkan bahwa LSTM memiliki performa sangat baik dalam mengklasifikasikan sentimen ulasan Tokopedia, dengan akurasi 94% pada data uji. Untuk kelas Negatif, model mencapai precision 0,95, recall 0,93, dan F1-score 0,94, sedangkan kelas Positif memperoleh precision 0,92, recall 0,94, dan F1-score 0,93. Confusion matrix menunjukkan bahwa LSTM berhasil mengklasifikasikan 535 ulasan Negatif dan 463 ulasan Positif dengan benar, meski terdapat 38 ulasan Negatif yang salah diklasifikasikan sebagai Positif dan 27 ulasan Positif yang salah sebagai Negatif. Nilai macro average dan weighted average F1-score masing-masing 0,94 menandakan performa yang konsisten dan seimbang pada kedua kelas. Visualisasi hasil klasifikasi dapat dilihat pada Gambar 9.



Gambar 9. Confusion Matrix LSTM

H. Perbandingan Algoritma Random Forest dan LSTM

Berdasarkan hasil pengujian yang telah dilakukan, Tabel III menampilkan perbandingan kinerja model Random Forest dan Long Short-Term Memory (LSTM) dalam mengklasifikasikan sentimen ulasan Tokopedia. Evaluasi dilakukan menggunakan data uji yang terpisah dari data pelatihan dengan menerapkan metrik akurasi, presisi, recall, dan F1-score untuk masing-masing kelas sentimen, sehingga performa kedua model dapat dianalisis secara lebih komprehensif.

TABEL III
PERBANDINGAN PERFORMA ALGORITMA

Algoritma	Akurasi	Kelas	Presisi	Recall	F1-score
Random Forest	92%	Positif	93%	90%	91%
		Negatif	92%	94%	93%
LSTM	94%	Positif	92%	94%	93%
		Negatif	95%	93%	94%

Hasil evaluasi menunjukkan bahwa model LSTM memperoleh akurasi tertinggi sebesar 94%, sedikit lebih unggul dibandingkan Random Forest dengan akurasi sebesar 92%. Pada model Random Forest, kelas positif memiliki nilai presisi sebesar 93% dan recall sebesar 90%, sedangkan kelas negatif menunjukkan nilai recall yang lebih tinggi, yaitu 94%. Temuan ini mengindikasikan bahwa Random Forest cenderung lebih efektif dalam mengenali ulasan bersentimen negatif dibandingkan ulasan positif. Sementara itu, model LSTM menunjukkan performa yang lebih seimbang pada kedua kelas sentimen. Kelas positif memiliki nilai recall sebesar 94%, yang mencerminkan kemampuan LSTM dalam mengidentifikasi ulasan positif secara tepat. Pada kelas negatif, LSTM mencapai nilai presisi sebesar 95%, yang

menandakan tingkat kesalahan prediksi negatif yang relatif rendah. Konsistensi nilai F1-score pada kedua kelas menunjukkan bahwa LSTM mampu menjaga keseimbangan antara presisi dan recall secara lebih stabil dibandingkan Random Forest.

Meskipun LSTM mencatat akurasi sedikit lebih tinggi dibandingkan Random Forest dengan selisih 2%, perbedaan ini tergolong marginal dan belum diuji secara statistik, sehingga kedua model dapat dianggap kompetitif dalam klasifikasi sentimen ulasan Tokopedia. LSTM unggul dalam menangkap urutan kata dan konteks semantik melalui embedding, sehingga lebih efektif pada kalimat panjang atau ulasan dengan pergeseran sentimen. Sebagai contoh, pada ulasan yang diawali kata "bagus" namun diikuti keluhan seperti "pengiriman lama", Random Forest berbasis TF-IDF cenderung salah memprediksi sebagai positif, sementara LSTM mampu mempertahankan konteks negatif hingga akhir kalimat. Di sisi lain, Random Forest tetap stabil dalam mendeteksi sentimen negatif, menunjukkan bahwa pendekatan berbasis fitur statistik masih relevan untuk analisis ulasan *e-commerce* yang lebih lugas dan efisien.

Dari segi kompleksitas dan efisiensi, Random Forest lebih cepat dilatih dan membutuhkan sumber daya komputasi yang lebih sedikit, sehingga sangat cocok untuk sistem dengan keterbatasan perangkat keras atau kebutuhan pemrosesan cepat. Sebaliknya, LSTM memiliki arsitektur yang lebih kompleks dan memerlukan daya komputasi lebih besar karena proses pelatihan jaringan saraf dan embedding, namun model ini mampu menangkap konteks bahasa dan hubungan antar kata secara lebih mendalam. Oleh karena itu, pemilihan model sebaiknya disesuaikan dengan kebutuhan sistem, tujuan analisis, dan skala implementasinya.

Secara praktis, hasil analisis sentimen ini dapat membantu platform *e-commerce*, khususnya Tokopedia, dalam memantau opini konsumen secara otomatis. Informasi mengenai sentimen positif dan negatif pada ulasan produk bisa digunakan untuk mengidentifikasi produk atau penjual dengan tingkat keluhan tinggi, mendukung evaluasi kualitas layanan, serta menjadi dasar peningkatan sistem rekomendasi dan layanan pelanggan. Dengan demikian, penelitian ini tidak hanya bernilai akademis, tetapi juga relevan untuk memperbaiki pengelolaan dan kualitas layanan *e-commerce*.

Penelitian ini memiliki beberapa keterbatasan. Dataset yang digunakan bersumber dari Kaggle dan hanya mencakup dua kelas sentimen, yaitu positif dan negatif, tanpa mempertimbangkan keberadaan sentimen netral. Selain itu, perbandingan model hanya menggunakan Random Forest dan LSTM, sehingga belum mencakup metode lain yang relevan. Pengujian signifikansi statistik untuk menilai perbedaan performa antar model juga belum dilakukan, sehingga hasil perbandingan masih bersifat deskriptif.

IV. KESIMPULAN

Berdasarkan hasil perbandingan kinerja algoritma Random Forest dan Long Short-Term Memory (LSTM) dalam klasifikasi sentimen ulasan Tokopedia, dapat disimpulkan bahwa kedua metode mampu memberikan performa yang baik. Model LSTM menunjukkan kinerja terbaik dengan akurasi sebesar 94%, sedikit lebih tinggi dibandingkan Random Forest yang mencapai akurasi 92%. Selain itu, LSTM juga menunjukkan keseimbangan yang lebih stabil antara nilai presisi, recall, dan F1-score pada kedua kelas sentimen, yang mengindikasikan kemampuan model dalam mengklasifikasikan ulasan secara lebih konsisten.

Keunggulan LSTM terlihat pada kemampuannya dalam memahami urutan kata dan konteks semantik melalui representasi embedding, sehingga lebih efektif dalam menangani ulasan dengan struktur kalimat panjang atau mengandung pergeseran sentimen. Namun, efektivitas tersebut dicapai dengan konsekuensi meningkatnya kompleksitas model serta waktu pelatihan yang lebih lama dibandingkan metode machine learning tradisional. Sebaliknya, Random Forest yang menggunakan representasi fitur TF-IDF tetap menunjukkan performa yang kompetitif, khususnya dalam mendeteksi sentimen negatif, serta memiliki keunggulan dari sisi efisiensi komputasi dan waktu pelatihan yang lebih singkat, sehingga lebih sesuai untuk sistem dengan keterbatasan sumber daya.

Meskipun perbedaan performa antara kedua model relatif kecil dan belum diuji secara statistik, hasil penelitian ini menunjukkan bahwa pendekatan deep learning berbasis LSTM memiliki keunggulan dalam memahami konteks bahasa alami pada ulasan pelanggan, sedangkan Random Forest tetap menjadi alternatif yang andal ketika efisiensi komputasi dan kesederhanaan model menjadi pertimbangan utama. Oleh karena itu, pemilihan algoritma tidak hanya bergantung pada tingkat akurasi, tetapi juga perlu mempertimbangkan kebutuhan sistem, ketersediaan sumber daya, serta skala implementasi. Secara keseluruhan, penelitian ini memberikan gambaran komparatif yang dapat dijadikan acuan dalam pengembangan sistem analisis sentimen untuk pengolahan ulasan pelanggan pada layanan *e-commerce*.

DAFTAR PUSTAKA

[1] A. A. H. Siregar, R. R. Apriliani, dan N. Nurhasanah, "Analisis Korelasi Statistik Antara Populasi Jumlah Penduduk dan Pengguna Internet Di Indonesia," *RIGGS J. Artif. Intell. Digit. Bus.*, vol. 4, no. 3, hlm. 4776–4781, Sep 2025, doi: 10.31004/riggs.v4i3.2684.

[2] U. Ajnura, I. Ikramuddin, C. Chalirafi, dan M. Subhan, "Pengaruh Faktor Pendorong Belanja Online Terhadap Niat Perilaku Konsumen Di Kota Lhokseumawe Dengan Metode Pembayaran Cash-on-Delivery Sebagai Variabel Mediasi," *J. Manaj. Pemasar.*, vol. 18, no. 1, hlm. 25–39, Apr 2024, doi: 10.9744/pemasaran.18.1.25-39.

[3] N. Emantonio, R. S. Magdalena, dan A. Wulandari, "Analisis Pertumbuhan Platform Bisnis Digital di Indonesia," vol. 11, 2025.

[4] A. S. Kembau, J. G. E. Putri, dan R. J. Nehemia, "Customer Indecisiveness pada E-Commerce Indonesia: Peran Price Sensitivity, Product Involvement, Risk, dan Social Inference".

[5] Y. Sutarso, B. Suminar, dan A. Maschudah Ilfitriah, "Do shopping anxiety and data leakage risks matter to e-commerce customers? Evidence from the largest economy in Southeast Asia," *J. Manaj. Dan Pemasar. Jasa*, vol. 17, no. 1, hlm. 97–116, Apr 2024, doi: 10.25105/v17i1.18673.

[6] D. Fadila dan M. Ikhsan, "Analisis Sentimen Pada Aplikasi Tokopedia Menggunakan Metode Support Vector Machine," vol. 21, no. 1.

[7] H. K. Tunjungsari, "Pengaruh Persepsi Kegunaan Dari Ulasan Online, Kepercayaan Konsumen, Dan Persepsi Risiko Pada Intensi Membeli Produk Busana Secara Online," vol. 9, no. 2.

[8] Regina Dwi Amelia, M. Michael, dan R. Mulyandi, "Analisis Online Consumer Review Terhadap Keputusan Pembelian pada E-Commerce Kecantikan," *J. Indones. Sos. Teknol.*, vol. 2, no. 2, hlm. 274–280, Feb 2021, doi: 10.36418/jist.v2i2.80.

[9] N. A. Salsabila, U. Sa'adah, dan F. Fauzi, "Analisis Sentimen Pada Ulasan Aplikasi Tokopedia Menggunakan Klasifikasi Naïve Bayes," vol. 7, 2024.

[10] D. Darwis, N. Siskawati, dan Z. Abidin, "Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter BMKG Nasional," *J. Tekno Kompak*, vol. 15, no. 1, hlm. 131, Feb 2021, doi: 10.33365/jtk.v15i1.744.

[11] R. L. Atimi dan Enda Esyudha Pratama, "Implementasi Model Klasifikasi Sentimen Pada Review Produk Lazada Indonesia," *J. Sains Dan Inform.*, vol. 8, no. 1, hlm. 88–96, Jul 2022, doi: 10.34128/jsi.v8i1.419.

[12] A. H. Hasugian, M. Fakhriza, dan D. Zukhoiriyah, "Analisis Sentimen Pada Review Pengguna E-Commerce Menggunakan Algoritma Naïve Bayes," *J-SISKO TECH J. Teknol. Sist. Inf. Dan Sist. Komput. TGD*, vol. 6, no. 1, hlm. 98, Jan 2023, doi: 10.53513/jsk.v6i1.7400.

[13] A. R. Azis, "Analisis Komparasi Algoritma Machine Learning dalam Prediksi Performa Akademik Mahasiswa: Literature Review," *J. Ilmu Komput. Dan Inform.*, vol. 4, no. 2, hlm. 143–148, Jan 2025, doi: 10.54082/jiki.212.

[14] A. Syah, F. Nurdiyansyah, dan A. Y. Rahman, "Analisis Sentimen Aplikasi Shopee, Tokopedia, Lazada Dan Blibli Menggunakan Leksikon Dan Random Forest," *J. Inform. Dan Tek. Elektro Terap.*, vol. 12, no. 3S1, Okt 2024, doi: 10.23960/jitet.v12i3S1.5155.

[15] G. Tamami, W. A. Triyanto, dan S. Muzid, "Sentiment Analysis Mobile JKN Reviews Using SMOTE Based LSTM," *IJCCS Indones. J. Comput. Cybern. Syst.*, vol. 19, no. 1, hlm. 13, Jan 2025, doi: 10.22146/ijccs.101910.

[16] A. R. Gunawan dan R. F. A. Aziza, "Sentiment Analysis Using LSTM Algorithm Regarding Grab Application Services in Indonesia," vol. 9, no. 2.

[17] F. M. Rayhan, S. H. Wijoyo, dan W. H. N. Putra, "Analisis Sentimen Root Cause Analisis Kepuasan Pengguna Aplikasi Tokopedia Pada Ulasan Menggunakan Metode Random Forest".

[18] A. Andreyestha dan Q. N. Azizah, "Analisa Sentimen Kicauan Twitter Tokopedia Dengan Optimalisasi Data Tidak Seimbang Menggunakan Algoritma SMOTE," *Infotek J. Inform. Dan Teknol.*, vol. 5, no. 1, hlm. 108–116, Jan 2022, doi: 10.29408/jit.v5i1.4581.

[19] C. Y. Adhelina, M. H. Dar, dan M. N. S. Hasibuan, "Implementation of Deep Learning Models in Conducting Aspect-Based Sentiment Analysis".

[20] D. Purnamasari, A. B. Aji, S. Madenda, I. M. Wiryana, dan S. Harmanto, "Sentiment Analysis Methods for Customer Review of Indonesia E-Commerce," 2024, *ICIC International* 学会 01. doi: 10.24507/ijicic.20.01.47.

[21] S, Panggabean dan A, Junika, "Sentiment Analysis on Public Opinions Regarding the 2024 Regional Elections Using Long Short-Term Memory (LSTM), Random Forest, and Naive Bayes," 2024.

[22] M. Siino, I. Tinnirello, dan M. La Cascia, "Is text preprocessing still worth the time? A comparative survey on the influence of popular preprocessing methods on Transformers and traditional classifiers," *Inf. Syst.*, vol. 121, hlm. 102342, Mar 2024, doi: 10.1016/j.is.2023.102342.

[23] R. Rahmadani, A. Rahim, dan R. Rudiman, "Analisis Sentimen Ulasan 'Ojol the Game' Di Google Play Store Menggunakan Algoritma Naive Bayes Dan Model Ekstraksi Fitur Tf-Idf Untuk Meningkatkan Kualitas Game," *J. Inform. Dan Tek. Elektro Terap.*, vol. 12, no. 3, Agu 2024, doi: 10.23960/jitet.v12i3.4988.

[24] N. A. Dirfas dan V. R. S. Nastiti, "Perbandingan Kinerja Pre-Trained Word Embedding Terhadap Performa Klasifikasi Sentimen Ulasan Produk Tokopedia Dengan Long Short-Term Memory(LSTM)," *Build. Inform. Technol. Sci. BITS*, vol. 6, no. 2, Sep 2024, doi: 10.47065/bits.v6i2.5634.

[25] S. Dermawan dan A. T. Ayunda, "Sentiment Analysis of Coretax on Social Media X Using Naive Bayes, SVM, and LSTM for Service Improvement," vol. 9, no. 6.

[26] S. Agustiani, Y. T. Arifin, A. Junaidi, S. K. Wildah, dan A. Mustopa, "Klasifikasi Penyakit Daun Padi menggunakan Random Forest dan Color Histogram," *J. Komputasi*, vol. 10, no. 1, hlm. 65–74, Apr 2022, doi: 10.23960/komputasi.v10i1.2961.

[27] Suci Amaliah, M. Nusrang, dan A. Aswi, "Penerapan Metode Random Forest Untuk Klasifikasi Varian Minuman Kopi di Kedai Kopi Konijiwa Bantaeng," *VARIANSI J. Stat. Its Appl. Teach. Res.*, vol. 4, no. 3, hlm. 121–127, Des 2022, doi: 10.35580/variansiunm31.

[28] Y. Romadhoni dan K. F. H. Holle, "Analisis Sentimen Terhadap PERMENDIKBUD No.30 pada Media Sosial Twitter Menggunakan Metode Naive Bayes dan LSTM," *J. Inform. J. Pengemb. IT*, vol. 7, no. 2, hlm. 118–124, Mei 2022, doi: 10.30591/jpit.v7i2.3191.