

Aspect Based Sentiment Analysis Terhadap Ulasan Produk Skincare di E-Commerce Menggunakan CNN-LSTM

I Made Treshnanda Mas^{a1}, AAIN Eka Karyawati^{a2}

^aProgram Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
¹treshnanda@gmail.com
²eka.karyawati@unud.ac.id

Abstract

Skincare product is any type of formulation intended to be topically applied with the intention of improving skin appearance, texture, or health. The burgeoning skincare product market in Indonesia, particularly in E-Commerce, underscores a shift towards online purchasing habits. With platforms like Shopee, Tokopedia, and Lazada dominating, user reviews play a pivotal role when making decision. However the sheer volume of reviews necessitates efficient processing methods. Aspect Based Sentiment Analysis (ABSA) emerges as a solution, delving deeper into sentiment analysis by identifying specific aspects within user feedback. Previous studies have shown Convolutional Neural Network based algorithms to have high performance in similar contexts. In this study, the author used CNN mixed with Long Short Term Memory (LSTM) alongside Term Frequency-Inverse Document Frequency (TF-IDF) and Bag of Words (BoW) as feature selection techniques to classify aspects into 5 classes and sentiment into 2 classes. The result of aspect classification produced accuracy, precision, recall, and F1-Score of 87.3%, 87%, 87%, and 85.9%. Meanwhile sentiment classification produced accuracy, precision, recall, and F1-Score of 87%, 87%, 87%, and 87%.

Keywords: Skincare Product, E-Commerce, Aspect Based Sentiment Analysis, CNN, LSTM, TF-IDF, BoW

1. Pendahuluan

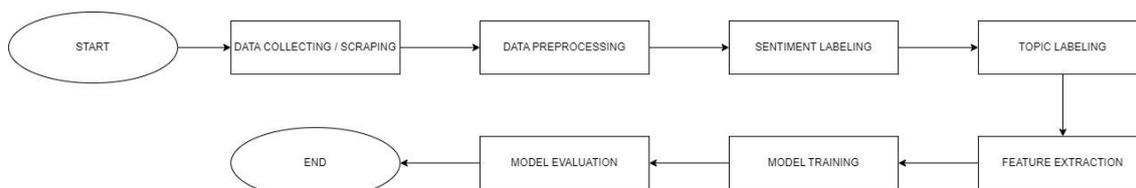
Produk skincare adalah formulasi apa pun yang dimaksudkan untuk dioleskan pada kulit dengan tujuan memperbaiki penampilan, tekstur, atau kesehatannya. Produk skincare tersedia dalam berbagai bentuk seperti krim, lotion, gel, masker, pembersih, dan pengelupas kulit. Produk skincare sering kali mengandung kombinasi bahan aktif, pelembab, emolien, dan senyawa lain yang dirancang untuk mengatasi permasalahan kulit tertentu. Penggunaan produk skincare di Indonesia semakin meningkat seiring tahun. Menurut laporan yang dirilis oleh [1] terhadap Produk FMCG (Fast Moving Consumer Goods) di E-Commerce Indonesia pada tahun 2023, produk kecantikan dan perawatan menempati posisi paling atas dengan persentase pembeliannya sebesar 49% dengan produk skincare sebagai sebagian besar dari produk kecantikan dan perawatan dengan persentasenya sebesar 39.4%. Menurut [2] E-Commerce merujuk pada proses distribusi, penjualan, promosi, serta pembelian barang atau layanan menggunakan teknologi elektronik seperti internet, televisi, jaringan komputer, dan lainnya. Ini juga mencakup transaksi keuangan online, pengelolaan stok secara otomatis, pertukaran informasi digital, dan sistem pengumpulan data secara otomatis. Pada tahun 2023, jumlah penduduk Indonesia yang merupakan pengguna E-Commerce adalah sebanyak 196 juta, angka tersebut meningkat dari tahun 2022 dengan jumlah pengguna E-Commerce adalah 178.94 juta [3]. Sebagian besar masyarakat Indonesia memanfaatkan platform E-Commerce seperti Shopee, Tokopedia, dan Lazada untuk berbelanja secara online. [4] melakukan analisis terhadap data SimilarWeb dengan hasilnya adalah Shopee merupakan platform E-Commerce dengan peringkat teratas paling sering dikunjungi, Tokopedia mendapat posisi kedua sedangkan Lazada menempati posisi

ketiga. Pengguna E-Commerce dapat memberikan suatu ulasan terhadap produk yang mereka beli setelah mereka menerima atau menggunakan produk tersebut, ulasan yang diberikan oleh pengguna berupa teks dan nilai rating dengan rentan 1 sampai 5. Ulasan-ulasan yang diberikan oleh pengguna terhadap suatu produk dapat digunakan oleh pengguna lainnya sebagai bahan pertimbangan ketika mereka ingin membeli suatu produk. Tetapi ulasan-ulasan ini cenderung berada dalam jumlah yang banyak dan setiap ulasan dapat membahas topik atau aspek yang berbeda-beda, sehingga jika menggunakan metode manual yaitu dengan membaca ulasan satu per satu akan memerlukan waktu yang cukup panjang. Oleh karena itu dibutuhkan suatu pendekatan untuk memproses ulasan-ulasan dari pengguna terhadap suatu produk agar informasi yang terdapat dalam ulasan-ulasan tersebut dapat dimanfaatkan dengan baik oleh pengguna. Aspect Based Sentiment Analysis (ABSA) merupakan area khusus dalam Text Sentiment Analysis yang bertujuan untuk menentukan polaritas sentimen pengguna terhadap istilah aspek tertentu dalam sebuah kalimat [5]. Tidak seperti Sentiment Analysis pada umumnya, informasi yang dapat diperoleh dari ABSA lebih banyak karena ulasan-ulasan tidak hanya dibagi menjadi dua kategori, yaitu "Positif" dan "Negatif", tetapi juga pada aspek apa kedua kelas tersebut. Sebagai contoh, kalimat "Harga produk ini sangat mahal" pada Sentiment Analysis biasa akan diklasifikasikan menjadi "Negatif" tetapi pada ABSA akan diklasifikasikan menjadi "Aspek Harga" dan "Negatif". Oleh karena itu, penggunaan ABSA diharapkan dapat memudahkan pengguna untuk mendapatkan informasi dari ulasan-ulasan terhadap suatu produk skincare. Penelitian yang dilakukan [6] menggunakan algoritma Convolutional Neural Network (CNN) untuk melakukan ABSA pada ulasan aplikasi Peduli Lindungi, hasil pengujian yang dilakukan menghasilkan performa yang bagus dengan nilai f1 score untuk klasifikasi aspek adalah 92.23% dan untuk klasifikasi sentimen adalah 95.13%. Dan penelitian yang dilakukan [7] penggunaan CNN-LSTM dalam analisis sentimen produk halar di Twitter menghasilkan nilai akurasi tertinggi sebesar 93.78% ketika dibandingkan dengan Long Short Term Memory (LSTM), CNN, dan CNN+Bi-RNN+Bi-LSTM. Sebab itu, ABSA dengan metode deep learning menggunakan algoritma CNN-LSTM akan dilakukan pada penelitian ini.

2. Metode Penelitian

2.1. Alur Penelitian

Gambar 1. Menunjukkan alur penelitian yang akan dilakukan pada penelitian ini.



Gambar 1. Alur Penelitian ABSA menggunakan CNN-LSTM

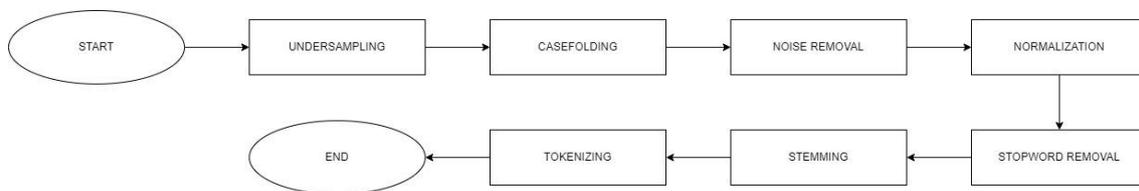
Pada penelitian ini pengumpulan data pertama dilakukan dengan scraping ulasan-ulasan pengguna terhadap produk skincare di platform E-Commerce. Selanjutnya setelah data terkumpul preprocessing terhadap data dilakukan dengan tujuan membersihkan data ulasan-ulasan pengguna tersebut. Pelabelan sentimen dilakukan dengan cara memberi kelas "Negatif" dan "Positif" kepada ulasan dengan menggunakan nilai rating yang diberi pengguna. Pelabelan topik atau aspek dilakukan menggunakan algoritma Latent Dirichlet Allocation (LDA) dengan data ulasan-ulasan pengguna yang sudah dilakukan preprocessing. Pada penelitian ini, fitur diekstraksi dengan metode Term Frequency-Inverse Document Frequency (TF-IDF) dan Bag of Words (BoW). Setelah melakukan ekstraksi fitur, langkah selanjutnya adalah melatih model berbasis CNN untuk klasifikasi aspek dan klasifikasi sentimen. Kemudian kedua model klasifikasi yang telah dilatih selanjutnya akan dievaluasi performanya.

2.2. Data Collecting

Data ulasan-ulasan pengguna terhadap produk skincare di platform E-Commerce yang digunakan dalam penelitian ini dikumpulkan melalui penerepan teknik scraping. Scraping dilakukan dengan cara mengirim request atau permintaan kepada Application Programming Interface (API) E-Commerce, response atau respons yang didapat dari pengiriman permintaan tersebut berupa file dengan ekstensi json. Respons tersebut kemudian diproses untuk mengambil data teks dan rating saja dan disimpan dalam format csv. Data ulasan-ulasan pengguna diambil dari 1400 produk dengan kategori skincare dan secara keseluruhan terdapat 252 ribu ulasan yang diperoleh.

2.3. Data Preprocessing

Adapun tahapan-tahapan preprocessing yang akan dilakukan pada penelitian ini terlampir pada Gambar 2.



Gambar 2. Data Preprocessing

Undersampling adalah langkah pertama yang dilakukan, di mana jumlah data ulasan-ulasan dikurangi dari 252 ribu menjadi 12 ribu. Selanjutnya casefolding dilakukan untuk mengubah semua kata kapital menjadi kata kecil agar semua kata direpresentasikan dalam bentuk yang seragam yaitu dengan kata kecil. Sehabis itu dilakukan noise removal yang bertujuan untuk menghapuskan elemen-elemen selain kata seperti tanda baca, angka, dan elemen html. Lalu normalization dilakukan untuk mengubah kata-kata yang ditulis dengan salah, dengan singkat, dan dengan bentuk tidak formal menjadi bentuknya yang seharusnya. Lalu dilakukan stopword removal dengan tujuan untuk menghapus kata-kata yang tidak mengandung makna atau arti. Setelah itu stemming akan dilakukan untuk mengubah kata-kata yang memiliki imbuhan menjadi bentuk dasarnya. Hasil dari proses stemming selanjutnya akan dilakukan proses tokenizing untuk memecah data yang berbentuk kalimat menjadi terpisah kata per kata.

2.4. Sentiment Labeling

Setelah melakukan preprocessing, selanjutnya ulasan-ulasan pengguna terhadap produk skincare akan diberi label sentimen berdasarkan nilai rating dari ulasan tersebut. Untuk ulasan yang memiliki nilai rating 1 sampai 3 maka akan diberikan label "Negatif", sedangkan untuk ulasan yang memiliki nilai rating 4 sampai 5 maka akan diberikan label "Positif". Secara keseluruhan, terdapat 6121 ulasan pengguna terhadap produk skincare dengan sentimen positif dan 6121 ulasan pengguna terhadap produk skincare dengan sentimen negatif.

2.5. Topic Labeling

Pelabelan topik dalam penelitian ini dilakukan dengan menggunakan algoritma LDA. LDA merupakan model probabilitas yang terkenal sebagai salah satu algoritma yang populer untuk pemodelan topik [8]. LDA mencoba menganalisis dokumen dengan melihat ke belakang untuk menemukan beragam topik yang mungkin ada. Pendekatan ini memungkinkan LDA untuk menyajikan dokumen dengan kumpulan topik yang diperkirakan berdasarkan probabilitasnya. Dalam penelitian ini, LDA digunakan untuk mengelompokkan ulasan-ulasan pengguna terhadap produk skincare ke dalam 5 topik yaitu "Delivery", "Skin Compatibility and Usage", "Ordering Experience", "Packaging Quality", "Product Experience" dengan jumlahnya masing-masing adalah 3873, 2601, 2309, 1761, dan 1596.

2.6. Feature Extraction

Dalam penelitian ini, dilakukan dua jenis ekstraksi fitur yaitu TF-IDF dan BoW. TF-IDF adalah metode ekstraksi fitur dengan algoritma yang bekerja dengan cara menghitung nilai bobot bagi setiap kata dalam sebuah kalimat, untuk menetapkan nilai bobot untuk setiap kata dalam dokumen, pendekatan ini akan menghitung seberapa sering kata tersebut muncul pada dokumen tersebut [9]. Dalam TF-IDF, semakin sering suatu kata itu muncul dalam dokumen maka semakin rendah nilai pembobotannya.

$$w_{ij} = tf_{ij} \cdot \log\left(\frac{D}{df_j}\right) \quad (1)$$

Sedangkan BoW adalah sebuah metode representasi yang mengubah data teks menjadi vektor dengan panjang yang telah ditentukan, dengan cara melakukan perhitungan frekuensi kata yang berulang. Dalam representasi ini, kalimat atau dokumen dianggap sebagai himpunan kata tanpa memperhatikan urutan dan direpresentasikan dalam vektor dengan ukuran tetap yang mencatat jumlah kemunculan kata, tanpa memperhatikan struktur tata bahasa atau urutan kata.

2.7. Convolutional Neural Network

Model yang akan digunakan dalam penelitian ini adalah CNN. CNN merupakan sebuah teknologi kecerdasan buatan yang berbasis pada jaringan saraf berlapis-lapis yang mampu mengenali, mengidentifikasi, dan mengklasifikasikan objek, serta mampu mendeteksi dan memisahkan objek dalam gambar. Meskipun CNN umumnya digunakan dalam Computer Vision seperti deteksi objek dan klasifikasi gambar tetapi CNN telah terbukti efektif dan memiliki performa bagus dalam klasifikasi teks [10]. CNN umumnya terdiri dari 4 layer, yaitu Convolution Layer, Embedding Layer, Pooling Layer, dan Fully Connected Layer. Convolution Layer adalah layer utama di mana proses konvolusi terjadi untuk input dan filter, ketika input memasuki lapisan ini, operasi konvolusi yang melibatkan filter diterapkan ke jendela kata untuk menghasilkan fitur baru, filter diterapkan berulang kali ke setiap jendela kata dalam kalimat untuk menghasilkan peta fitur [10]. Embedding Layer adalah layer yang berfungsi untuk memetakan input menjadi vektor berdimensi rendah, hasil pemetaan di Embedding Layer selanjutnya diberikan kepada Convolution Layer. Pooling Layer adalah layer yang memiliki fungsi untuk mengurangi jumlah parameter secara bertahap, kompleksitas komputasi model, dan kontrol overfitting [11]. Fully Connected Layer adalah layer dalam CNN yang membentuk neuron satu dimensi, layer ini terdiri dari neuron-neuron yang saling terhubung dengan neuron-neuron pada layer sebelum dan seterusnya [12].

2.8. Confusion Matrix

Confusion Matrix akan digunakan sebagai metode evaluasi untuk model klasifikasi aspek dan klasifikasi sentimen. Confusion Matrix merupakan sebuah tabel yang digunakan untuk menilai performa suatu model klasifikasi, dalam Confusion Matrix, setiap baris mewakili kelas sebenarnya dari data, sedangkan setiap kolom mewakili kelas prediksi oleh model. Terdapat 4 perhitungan yang dilakukan dalam Confusion Matrix yaitu akurasi, presisi, recall, dan F1-score.

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

$$\text{Presisi} = \frac{TP}{TP+FP} \quad (3)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (4)$$

$$\text{Recall} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

3. Hasil dan Diskusi

3.1. Data Collecting

Adapun contoh respons dari request data ulasan-ulasan pengguna terhadap produk dengan kategori skincare kepada API E-Commerce sebagai berikut.

Tabel 1. Hasil Data Collecting

index	Review Message	Rating
0	Sangat rekomen, pemakaian 1 bln terbukti ampuh	5
1	maap baru kasih ulasan, thanks bgt pesanan semuanya sesuai, good seller !	4
2	maaf di tipe kulit ku yg berminyak kurang cocok krna bikin tmbh berminyak	3
3	ku kira cocok buat pria ternyata oh ternyata putihnya pucet ga cocok banget malah timbul jerawat,kapok dah barang kebuang sia sia	2
4	tidak terjadi apa2 setelah pemakaian selama 3 bulan	1

3.2. Data Preprocessing

Setelah berhasil mengumpulkan data, selanjutnya jumlah baris yang akan digunakan adalah 12 ribu baris, jumlah baris tersebut ditentukan berdasarkan dari jumlah ulasan yang memiliki nilai rating 3 dan dibawahnya. Selanjutnya akan dilakukan Case Folding untuk mengubah semua kata kapital menjadi kata dasar, Noise Removal, Normalization, Stopword Removal, Stemming, dan Tokenizing. Berikut adalah hasil setelah

Tabel 2. Hasil Data Preprocessing

Tahapan	Text	Rating
Casefolding	sangat rekomen, pemakaian 1 bln terbukti ampuh	5
Noise Removal	sangat rekomen pemakaian bln terbukti ampuh	5
Normalization	sangat recommend pemakaian bulan terbukti ampuh	5
Stopword Removal	recommend pemakaian terbukti ampuh	5
Stemming	recommend pakai bukti ampuh	5
Tokenizing	[recommend, pakai, bukti, ampuh]	5

3.3. Sentiment Labeling

Setiap baris pada dataset yang telah dilakukan preprocessing kemudian akan diberikan label sentimen berdasarkan dari nilai rating yang dimiliki setiap baris. Untuk baris dengan nilai rating 4 sampai 5 maka akan diberikan sentimen label "Positif", sedangkan untuk baris dengan nilai rating 1 sampai 3 akan diberikan label sentimen "Negatif". Setelah pemberian label sentimen, terdapat 6121 baris dalam dataset dengan label sentimen "Positif", dan 6121 baris dalam dataset dengan label sentimen "Negatif".

Tabel 3. Hasil Sentiment Labeling

Text	Rating	Sentimen
[recommend, pakai, bukti, ampuh]	5	Positif
['enggak', 'cocok', 'muka']	3	Negatif

3.4. Topic Labeling

Pelabelan topik atau aspek dengan LDA dilakukan dengan menggunakan hasil dari ekstraksi fitur BoW. Adapun 5 topik dengan 10 kata kunci masing masing topik sebagai berikut.

Tabel 4. Pengelompokan Topik dengan Kata Kuncinya

Topic	Kata Kunci
Delivery	[kirim, barang, lama, cepat, aman, pesan, packing, banget, produk, proses]
Product Experience	[enggak, pakai, beli, kalo, kecewa, sih, tidak, beda, komplain, kayak]
Ordering Experience	[terima, kasih, barang, sesuai, harga, beli, tokopedia, produk, tidak, paket]
Skin Compatibility and Usage	[cocok, kulit, pakai, moga, coba, enggak, jerawat, muka, banget, wajah]
Packaging Quality	[enggak, tidak, kemas, isi, buka, tanggal, beli, botol, tutup, pas]

Kemudian setiap baris pada dataset akan diberi label topik atau aspeknya berdasarkan dari hasil pengelompokan topik berdasarkan kata-kata yang dilakukan dengan menggunakan model LDA. Distribusi jumlah baris setiap topik dapat dilihat pada tabel sebagai berikut.

Tabel 5. Distribusi Jumlah Topik

Topic	Jumlah
Delivery	3873
Product Experience	1596
Ordering Experience	2309
Skin Compatibility and Usage	2601
Packaging Quality	1761

3.5. Data Preparation

Sebelum diberikan kepada model untuk pelatihan, dataset yang telah dilakukan preprocessing menjalani tahap pemisahan untuk memastikan keberhasilan pelatihan, validasi, dan evaluasi model yang efektif. Proses ini dimulai dengan pembagian dataset menjadi tiga subset utama yaitu pelatihan, validasi, dan evaluasi. Pembagian dataset dilakukan dengan proporsi yang telah ditetapkan sebelumnya, di mana 80% dari data digunakan untuk pelatihan model, 10% dialokasikan untuk validasi selama proses pelatihan, dan 10% tersisa untuk mengukur performa model setelah pelatihan selesai.

3.6. TF-IDF Feature Extraction

Teknik ekstraksi fitur TF-IDF dilakukan untuk mengubah data tekstual menjadi representasi dalam bentuk numerikal untuk diberikan sebagai input dalam model CNN. Proses ini dimulai dengan tokenisasi dan pengkodean urutan dari data teks input, teks disegmentasi menjadi token atau kata individual dan mengkodekannya ke dalam urutan panjang tetap menggunakan padding untuk memastikan keseragaman dalam dimensi input. Setelah itu, representasi TF-IDF dibuat untuk data teks. Adapun hasil dari ekstraksi fitur TF-IDF pada gambar berikut.

```
print(tfidf_train.shape)
print(tfidf_train)

(9712, 6116)
(0, 5652) 0.6650708789751087
(0, 5658) 0.31338233187547343
(0, 362) 0.44377947657805805
(0, 3130) 0.41388246414479085
(0, 2039) 0.302048542537632
(1, 3190) 0.8665414472226423
(1, 2688) 0.4991051194340615
(2, 1144) 0.11683185976136362
(2, 3218) 0.21011689495370203
```

Gambar 3. Hasil Ekstraksi TF-IDF

3.7. BoW Feature Extraction

Selain TF-IDF, ekstraksi fitur dengan menggunakan teknik BoW juga dilakukan. Mirip dengan proses ekstraksi fitur TF-IDF, dalam ekstraksi fitur BoW pertama dilakukan tokenisasi dan mengkodekan urutan teks. Setelah itu representasi BoW dihasilkan untuk data teks dengan mengubah urutan teks yang telah diberi token menjadi vektor dengan menghitung frekuensi setiap istilah dalam kosakata. Hasil ekstraksi fitur ini kemudian akan digunakan sebagai input dalam model CNN. Adapun hasil ekstraksi fitur BoW dapat terlihat pada gambar sebagai berikut

```
print(bow_train.shape)
print(bow_train)

(9712, 6116)
(0, 2039) 1
(0, 3130) 1
(0, 362) 1
(0, 5658) 1
(0, 5652) 1
(1, 2688) 1
(1, 3190) 1
(2, 4344) 1
```

Gambar 4. Hasil Ekstraksi BoW

3.8. Pelatihan Model CNN-LSTM Klasifikasi Aspek

Model CNN yang akan dilatih untuk klasifikasi aspek pada penelitian ini digabungkan dengan LSTM untuk secara efektif menangkap pola spasial dan temporal dalam data teks. Model ini menggabungkan representasi TF-IDF dan BoW dari data teks input sebagai fitur tambahan. Representasi ini digabungkan dengan output dari lapisan LSTM dan dimasukkan ke dalam lapisan padat yang terhubung sepenuhnya untuk klasifikasi. Arsitektur model CNN-LSTM yang digunakan untuk pelatihan terdiri dari tiga komponen utama: Embedding Layer yang diikuti oleh dua Convolution Layer, LSTM Layer, dan Dense Layer berikutnya untuk klasifikasi. Berikut adalah parameter yang digunakan dalam pelatihan model CNN-LSTM untuk klasifikasi aspek.

Tabel 6. Parameter Pelatihan Model Klasifikasi Aspek

Parameter	Nilai
Epochs	20

Parameter	Nilai
Batch Size	128
Conv1D (1)	Filters=16, kernel_size=4. Activation = relu
Conv1D (2)	Filters=4, kernel_size=2. Activation = relu
LSTM	Units=128. Dropout=0.5, recurrent_dropout=0.2
Dense (1)	Units=128, dropout=0.5, activation=relu
Learning Rate	0.001

3.9. Pelatihan Model CNN Klasifikasi Sentimen

Model CNN untuk klasifikasi sentimen juga menggabungkan LSTM untuk secara efektif menangkap pola spasial dan temporal dalam data teks. Model ini juga menggabungkan representasi TF-IDF dan BoW dari data teks input sebagai fitur tambahan. Representasi ini digabungkan dengan output dari lapisan LSTM dan dimasukkan ke dalam lapisan padat yang terhubung sepenuhnya untuk klasifikasi. Arsitektur model CNN-LSTM yang digunakan untuk analisis sentimen terdiri dari Embedding Layer yang diikuti oleh 1 Convolution Layer, LSTM Layer, dan Dense Layer berikutnya untuk klasifikasi. Berikut merupakan nilai parameter yang digunakan dalam pelatihan model CNN-LSTM untuk klasifikasi sentimen.

Tabel 7. Parameter Pelatihan Model Klasifikasi Sentimen

Parameter	Nilai
Epochs	30
Batch Size	128
Conv1D (1)	Filters=128, kernel_size=5. Activation = relu
LSTM	Units=128. Dropout=0.2, recurrent_dropout=0.2
Dense (1)	Units=128, dropout=0.2, activation=relu
Learning Rate	0.0001

3.10. Evaluasi Model CNN-LSTM Klasifikasi Aspek

Setelah model CNN-LSTM untuk klasifikasi aspek selesai dilatih, langkah selanjutnya adalah mengukur dan mengevaluasi performa model tersebut menggunakan data evaluasi yang sebelumnya telah dipisah dari keseluruhan dataset. Berikut adalah hasil evaluasi dari model CNN-LSTM untuk klasifikasi aspek.



Gambar 5. Confusion Matrix Model Klasifikasi Aspek

Berdasarkan dari Confusion Matrix pada gambar 5. Model klasifikasi aspek memiliki nilai akurasi sebesar 87.3%, presisi sebesar 87%, recall sebesar 87%, dan F1-Score sebesar 85,9%.

3.11. Evaluasi Model CNN-LSTM Klasifikasi Sentimen

Pengukuran dan evaluasi performa model CNN-LSTM untuk klasifikasi sentimen dilakukan setelah pelatihan selesai. Evaluasi performa model akan dilakukan menggunakan data evaluasi yang telah dipisah dari keseluruhan dataset. Hasil evaluasi dari model CNN-LSTM untuk klasifikasi sentimen dapat dilihat sebagai berikut.



Gambar 6. Confusion Matrix Model Klasifikasi Sentimen

Berdasarkan Confusion Matrix gambar 6. Yang didapat dengan menggunakan data evaluasi, didapatkan bahwa model CNN-LSTM untuk klasifikasi sentimen memiliki nilai akurasi sebesar 87%, nilai presisi sebesar 87%, nilai recall sebesar 87% dan nilai F1-Score sebesar 87%.

4. Kesimpulan

Dapat disimpulkan bahwa penggunaan model CNN-LSTM dalam Aspect Based Sentiment Analysis terhadap ulasan produk skincare di E-Commerce memberikan hasil yang cukup bagus. Dengan menggunakan teknik ekstraksi fitur TF-IDF dan BoW, model CNN-LSTM untuk klasifikasi aspek mendapatkan nilai akurasi sebesar 87.3%, presisi sebesar 87%, recall sebesar 87%, dan F1-Score sebesar 85,9%. Sedangkan, model CNN-LSTM untuk klasifikasi sentimen mendapatkan nilai akurasi sebesar 87%, nilai presisi sebesar 87%, nilai recall sebesar 87% dan nilai F1-Score sebesar 87%.

Daftar Pustaka

- [1] Kompas, "Kompas Market Insight: Indonesian FMCG E-commerce Report 2023," Kompas. <https://kompas.co.id/article/kompas-market-insight-indonesia-fmcg-e-commerce-report-2023/> (accessed May. 9, 2024)
- [2] H. Harmayani, D. Marpaung, A. Hamzah, N. Mulyani, and J. Hutahaeen, E-Commerce: Suatu Pengantar Bisnis Digital. Yayasan Kita Menulis, 2020.
- [3] S. Deny, "Masih Digandrungi, pengguna e-commerce Indonesia capai 196 Juta di 2023," liputan6, <https://www.liputan6.com/bisnis/read/5447108/masih-digandrungi-pengguna-e-commerce-indonesia-capai-196-juta-di-2023?page=2> (accessed May 9, 2024).
- [4] A. Ahdiat, "5 e-commerce Dengan Pengunjung terbanyak sepanjang 2023," Databoks: Pusat Data Ekonomi dan Bisnis Indonesia, <https://databoks.katadata.co.id/datapublish/2024/01/10/5-e-commerce-dengan-pengunjung-terbanyak-sepanjang-2023> (accessed May 9, 2024).
- [5] K. Ahmed, M. I. Nadeem, Z. Zheng, D. Li, I. Ullah, M. Assam, Y. Y. Ghadi, and H. G. Mohamed, "Breaking down linguistic complexities: A structured approach to aspect-based sentiment analysis," J. King Saud Univ. - Comput. Inf. Sci., vol. 35, no. 8, pp. 101651, 2023. [Online]. Available: <https://doi.org/10.1016/j.jksuci.2023.101651>

- [6] P. Aritonang, M. Johan, and I. Prasetiawan, "Aspect-Based Sentiment Analysis on Application Review using Convolutional Neural Network," *Ultima InfoSys: J. Ilmu Sist. Inf.*, vol. 13, pp. 54-61, 2022. [Online]. Available: <https://doi.org/10.31937/si.v13i1.2684>.
- [7] A. Feizollah, S. Ainin, N. Anuar, N. Abdullah, and H. Hanif, "Halal Products on Twitter: Data Extraction and Sentiment Analysis Using Stack of Deep Learning Algorithms," *IEEE Access*, vol. 7, pp. 1-1, 2019. [Online]. Available: <https://doi.org/10.1109/ACCESS.2019.2923275>.
- [8] R. Albalawi, T. H. Yeap, and M. Benyoucef, "Using Topic Modeling Methods for Short-Text Data: A Comparative Analysis," *Front. Artif. Intell.*, vol. 3, pp. 42, 2020. [Online]. Available: <https://doi.org/10.3389/frai.2020.00042>
- [9] S. A. Pratomo, S. Al Faraby, and M. D. Purbolaksono, "Analisis Sentimen Pengaruh Kombinasi Ekstraksi Fitur TF-IDF dan Lexicon Pada Ulasan Film Menggunakan Metode KNN," 2021.
- [10] Y. Kim, "Convolutional Neural Networks for Sentence Classification," 2014. [Online]. Available: <https://arxiv.org/abs/1408.5882>.
- [11] K. O'Shea and R. Nash, "An Introduction to Convolutional Neural Networks," 2015. [Online]. Available: <https://arxiv.org/abs/1511.08458>.
- [12] M. Bangsa, S. Priyanta, and Y. Suyanto, "Aspect-Based Sentiment Analysis of Online Marketplace Reviews Using Convolutional Neural Network," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 14, pp. 123, Apr. 2020. [Online]. Available: <https://doi.org/10.22146/ijccs.51646>.