

Pengenalan Emosi Ulasan Pelanggan E-Commerce Menggunakan Deep Learning Berbasis Transformer

*Emotion Recognition of E-Commerce Customer Reviews
Using Transformer-Based Deep Learning*

Ahmad Sabil Deva Pratama¹, Nova Rijati^{2*}

Fakultas Ilmu Komputer, Universitas Dian Nuswantoro, Indonesia
E-mail: 111202013098@mhs.dinus.ac.id, nova.rijati@dsn.dinus.ac.id

**Corresponding author*

Abstrak

Penelitian ini mengeksplorasi penerapan arsitektur deep learning berbasis transformer untuk mengidentifikasi emosi dari ulasan pelanggan e-commerce berbahasa Indonesia. Menggunakan dataset yang terdiri dari 5.400 ulasan pelanggan, model ini dirancang untuk mengklasifikasikan lima kategori emosi: Happy, Sadness, Anger, Love, dan Fear. Hasil analisis menunjukkan kecenderungan pelanggan untuk berbagi pengalaman positif dengan dominasi emosi Happy dalam ulasan. Model Transformer berhasil mencapai akurasi klasifikasi keseluruhan sebesar 77,2%, dengan efisiensi waktu pelatihan yang optimal sekitar 90 detik. Evaluasi performa menggunakan confusion matrix dan metrik presisi, recall, dan F1-score memberikan wawasan tentang keefektifan model dalam membedakan emosi kompleks. Temuan ini merefleksikan potensi pemanfaatan teknologi deep learning dalam meningkatkan pemahaman terhadap perilaku pelanggan dan mendukung pengembangan strategi bisnis yang responsif.

Kata kunci: Transformer, Ulasan Pelanggan, Klasifikasi Emosi

Abstract

This research explores the application of a Transformer-based deep learning architecture to identify emotions from customer reviews in Indonesian-language e-commerce. Using a dataset of 5,400 customer reviews, the model is designed to classify five categories of emotions: Happy, Sadness, Anger, Love, and Fear. The analysis results show the tendency of customers to share positive experiences with the dominance of Happy emotion in the reviews. The Transformer model achieved an overall classification accuracy of 77,2%, with an optimal training time efficiency of about 90 seconds. Performance evaluation using confusion matrix and precision, recall, and F1-score metrics provided insights into the model's effectiveness in distinguishing complex emotions. These findings reflect the potential of utilizing deep learning technologies in improving understanding of customer behavior and supporting the development of responsive business strategies.

Keywords: Transformer, Sentiment Analysis, Customer Reviews, Emotion Classification

1. PENDAHULUAN

Dampak perkembangan e-commerce yang pesat menyebabkan pelanggan semakin cenderung mengandalkan *platform* pemasaran *online* untuk melakukan pembelian produk. Setelah melakukan pembelian, mereka kemudian berbagi ulasan yang mencerminkan pengalaman pribadi, perasaan, dan respons emosional mereka terhadap produk tersebut [1]. Kebebasan dunia maya menghapuskan batasan geografis fisik, memungkinkan pelanggan untuk menjelajahi berbagai pilihan dari berbagai merek dan penjual di seluruh dunia. Perubahan ini mempengaruhi pola perilaku konsumen dalam proses pembelian, memberikan mereka akses yang lebih luas terhadap informasi yang komprehensif tentang produk atau layanan yang sedang mereka pertimbangkan, yang tersedia dalam bentuk ulasan yang dibagikan oleh sesama pelanggan [2].

Emosi adalah unsur yang tak terpisahkan dalam kehidupan manusia dan memainkan peran yang signifikan dalam pengambilan keputusan individu. Dalam konteks bisnis dan konsumen, ulasan pelanggan online telah diakui sebagai faktor yang memiliki pengaruh besar terhadap keputusan pembelian [3]. Dalam era *e-commerce* yang tumbuh pesat, ulasan pelanggan telah menjadi komponen penting dalam membentuk preferensi pembelian, dan konsumen sendiri percaya bahwa ulasan yang mencerminkan emosi yang kuat akan berdampak lebih besar pada orang lain serta memengaruhi keputusan pembelian mereka [4]. Menurut penelitian [5], ulasan produk *online* telah berkembang menjadi sumber informasi yang berharga yang mempengaruhi keputusan pembelian dengan meningkatnya media digital dan sosial. Pengenalan emosi mencoba untuk mengidentifikasi emosi yang lebih spesifik yang mungkin terkandung dalam ulasan, seperti senang, sedih, atau marah. Sebuah ulasan mungkin memiliki sentimen positif tetapi mengungkapkan emosi seperti kegembiraan atau kelegaan. Sebaliknya, ulasan dengan sentimen negatif mungkin mencerminkan emosi seperti frustrasi, kekecewaan, atau kemarahan [6]. Oleh karena itu, pemahaman yang mendalam tentang bagaimana emosi tercermin dalam ulasan pelanggan dan bagaimana emosi tersebut memengaruhi perilaku pembelian adalah hal yang sangat penting dalam memahami dinamika pasar *e-commerce* yang kompetitif saat ini.

Natural Language Processing (NLP) muncul sebagai teknologi yang mampu mengekstraksi makna dan emosi dari teks. Melalui penggunaan algoritma yang canggih dan model pelatihan yang mendalam, NLP memiliki kapabilitas untuk mengenali pola bahasa, memahami konteks, dan bahkan menafsirkan emosi yang terkandung dalam teks [7]. Terutama dalam konteks ulasan produk, NLP membuka peluang yang sangat menarik untuk memahami pandangan pelanggan dalam skala besar. Ini memungkinkan penyedia layanan untuk memberikan analisis sentimen yang akurat dan wawasan mendalam yang sebelumnya sulit dicapai dalam jumlah besar. Penggunaan NLP dalam menganalisis ulasan pelanggan membuka pintu untuk pemahaman yang lebih baik tentang preferensi dan opini pelanggan, yang dapat menginformasikan pengambilan keputusan bisnis yang lebih cerdas dan efektif.

Seiring dengan pesatnya perkembangan dalam bidang NLP, model berbasis Transformer muncul sebagai salah satu pendekatan yang paling canggih dan berpengaruh dalam NLP. Penelitian [8] memperkenalkan model berbasis Transformer yang telah merevolusi bidang NLP dengan pendekatan inovatifnya terhadap pemrosesan sekuensial. Model berbasis Transformer memperkenalkan paradigma baru dalam pemrosesan bahasa yang telah mengubah cara kita memahami dan mengolah teks. Salah satu fitur paling menonjol dari Transformer adalah struktur *attention* yang memungkinkan model untuk fokus pada bagian-bagian tertentu dari teks sumber saat menganalisis atau menghasilkan teks target [9]. Kemampuan struktur *attention* ini memiliki dampak yang sangat besar dalam pemahaman kontekstual, yang sangat penting dalam menangkap nuansa emosi dalam teks.

Dalam beberapa tahun terakhir, teknologi NLP mengalami kemajuan pesat dalam mengenali emosi dari teks. Ulasan produk di *e-commerce* oleh konsumen menjadi sumber informasi yang kaya untuk mendapatkan wawasan tentang persepsi dan emosi pelanggan terhadap produk atau layanan. Analisis sentimen, khususnya, telah menjadi metode yang populer untuk mengekstrak dan memahami emosi dari ulasan tersebut. Peran analisis sentimen sangat signifikan dalam interpretasi ulasan konsumen terkait produk, melalui penerapan *machine learning*, serta perbandingannya dengan teknik *machine learning* lainnya [10]. Terlebih lagi, pendekatan yang berfokus pada teks dalam memprediksi emosi telah menunjukkan potensi yang signifikan. Sebagai contoh, [11] mengembangkan sistem prediksi emosi yang berlandaskan teks menggunakan *machine learning*, yang berhasil mengenali dan mengkategorikan emosi dari teks. Lebih lanjut, [12] menekankan penggunaan *deep learning* dalam pengenalan emosi berbasis teks, yang menunjukkan hasil yang lebih baik dalam menangkap nuansa emosi yang kompleks dalam teks dibandingkan dengan metode-metode tradisional. Sebuah penelitian [13] menunjukkan bahwa pendekatan berbasis *deep learning* dapat digunakan untuk mengenali emosi dari teks tertulis dan teks lisan. Hasil penelitian menunjukkan bahwa sistem yang mereka kembangkan mampu mengenali emosi dengan akurasi yang lebih baik, terutama ketika menggunakan dataset

yang lebih besar. Dengan demikian, pendekatan *deep learning* menawarkan potensi besar dalam meningkatkan akurasi dan efektivitas pengenalan emosi dari teks.

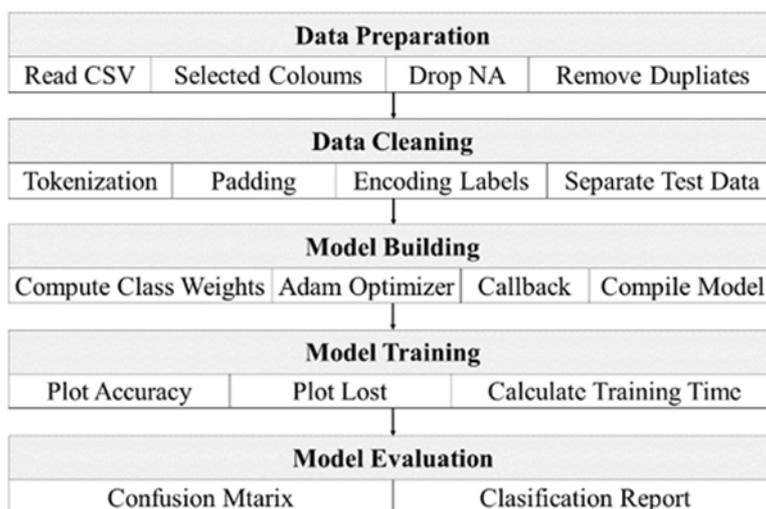
Dalam beberapa penelitian terkini, model berbasis Transformer, seperti BERT, menunjukkan keunggulan dalam tugas-tugas NLP, termasuk dalam menganalisis ulasan pelanggan. Peneliti [14] mengusulkan penggunaan BERT untuk memprediksi sentimen pelanggan dengan menggunakan ulasan pelanggan dari platform *Twitter*, *IMDB Movie Reviews*, *Yelp*, dan *Amazon*. Hasilnya menunjukkan bahwa model BERT unggul dalam kinerjanya dengan skor akurasi sebesar 90%, *recall* sebesar 91%, presisi sebesar 90%, dan F1 sebesar 90%, dibandingkan dengan model-model lain seperti LSVM, *fastText*, dan BiLSTM yang skor evaluasinya lebih rendah. Selain itu, penelitian [15] menggunakan BERT dan RoBERTa dalam sentiment analisis pada *customer review* di e-commerce dan menghasilkan akurasi sebesar 90% untuk BERT dan akurasi sebesar 97,6% untuk RoBERTa. Hasil ini melampaui akurasi dari Naive Bayes dan SVM yang menggunakan pembobotan *bag of words*.

Penelitian ini menggunakan arsitektur Transformer dalam upaya untuk memperdalam pemahaman tentang bagaimana emosi tercermin dalam ulasan pelanggan pada perilaku pembelian dalam konteks *e-commerce* yang sangat kompetitif. Diharapkan bahwa hasil penelitian ini memberikan wawasan berharga bagi pengambilan keputusan bisnis yang lebih cerdas dan efektif, membantu perusahaan dalam memahami preferensi dan opini pelanggan dengan lebih baik untuk menghadapi persaingan sengit dalam dunia *e-commerce* secara lebih baik.

2. METODE PENELITIAN

Penelitian ini menggunakan *Product Reviews Dataset for Emotions Classification Tasks Indonesian (PRDECT-ID) Dataset*. Dataset ini dikumpulkan dari ulasan produk di platform *e-commerce* Tokopedia, yang merupakan representasi dari beragam ekspresi emosi pelanggan terhadap berbagai produk, seperti *anger*, *fear*, *happy*, *love*, dan *sadness*, berdasarkan kriteria anotasi yang dikembangkan oleh [16].

Berdasarkan Gambar 1 proses penelitian dimulai dengan pengumpulan dan pembersihan data yang ekstensif, menggunakan dataset yang diperoleh dari ulasan produk di platform *e-commerce* Tokopedia. Data awal dimuat dari file CSV, dipilih kolom yang relevan, menghilangkan baris dengan nilai yang hilang, dan menghapus duplikat untuk memastikan integritas data. Pembersihan data melibatkan konversi teks ulasan menjadi huruf kecil dan mengganti karakter non-alfabetik dengan spasi, serta menghilangkan spasi berlebih. Kemudian mengekstrak subset dari data untuk dijadikan data uji, memisahkan 100 sampel acak dari setiap kategori emosi, memastikan bahwa setiap kategori diwakili secara setara dalam pengujian.



Gambar 1 Alur Kerja Model Pengenalan Emosi

Proses persiapan data lebih lanjut melibatkan *tokenisasi teks*, yaitu teks dipecah menjadi unit-unit yang lebih kecil yang dapat diproses oleh model. Setelah *tokenisasi*, urutan *token dipad* untuk memastikan konsistensi panjang, dan label emosi dikodekan untuk persiapan pelatihan. Data kemudian dibagi menjadi set pelatihan dan validasi, dengan 80% data untuk pelatihan dan 20% untuk validasi, menggunakan teknik *splitting* yang mempertimbangkan keacakan dan distribusi kelas yang seimbang.

Model klasifikasi emosi dibangun berdasarkan arsitektur Transformer yang menggunakan teknik *Multi Head Attention*. Selanjutnya parameter model ditentukan, seperti dimensi *embedding*, jumlah *attention heads*, dan ukuran lapisan tersembunyi dalam jaringan *feed-forward*. *Regularisasi L2* juga diterapkan untuk mengurangi *overfitting*. Dengan arsitektur ini, model dikompilasi menggunakan *optimizer Adam* dengan *learning rate* yang sangat rendah untuk memfasilitasi *konvergensi* yang stabil.

Pelatihan model melibatkan penggunaan bobot kelas yang dihitung untuk menangani ketidakseimbangan distribusi kelas dalam data. Proses pelatihan dipantau menggunakan grafik yang *plot* akurasi dan *loss* dari set pelatihan dan validasi, memberikan visualisasi yang jelas dari kemajuan dan kinerja model seiring waktu. Waktu pelatihan juga dipantau untuk mengevaluasi efisiensi komputasional dari model. Setelah pelatihan selesai, model dievaluasi menggunakan data uji yang sebelumnya telah dipisahkan. Prediksi model dianalisis menggunakan *confusion matrix* dan laporan klasifikasi yang menyediakan metrik kinerja seperti *presisi*, *recall*, dan skor F1 untuk setiap kelas.

3. HASIL DAN PEMBAHASAN

Dengan total 5.400 entri, dataset ini mencakup ulasan dalam Bahasa Indonesia, yang memberikan wawasan unik tentang pengalaman dan persepsi pelanggan. Setiap entri dalam dataset ini terdiri dari 11 kolom meliputi berbagai kategori produk, diperkaya dengan atribut seperti lokasi, harga, rating keseluruhan, jumlah terjual, total ulasan, dan rating pelanggan, menyediakan basis data komprehensif untuk analisis emosi dalam ulasan pelanggan. Fokus penelitian ini hanya pada kolom '*Customer Review*' dan '*Emotion*', yang masing-masing menyediakan teks ulasan dan label emosi yang terkait. Keberagaman data ini memastikan representasi ekspresi emosi yang luas, membuka jalan bagi analisis mendalam terkait pengaruh ulasan terhadap perilaku pembelian di *e-commerce* Indonesia

3.1 Data Preparation

Data Preparation dilakukan untuk mengurangi beban komputasi yang tidak perlu pada model dan meningkatkan akurasi pembelajaran dengan fokus pada ciri-ciri teks yang paling relevan untuk analisis emosi. Proses yang dilakukan pada dataset ulasan pelanggan meliputi:

1. Seleksi Kolom. Dataset kompleks dengan banyak kolom disederhanakan menjadi dua kolom yang relevan yaitu '*Customer Review*' dan '*Emotion*'. Kolom ini dipilih karena langsung berkaitan dengan tujuan penelitian, yaitu untuk menganalisis teks ulasan dan mengidentifikasi emosi yang terkait.
2. Pembersihan Awal. Data yang tidak lengkap (*null values*) dihapus menggunakan metode '*dropna()*'. Data redundan dihilangkan dengan '*drop_duplicates()*', sehingga menghasilkan dataset yang lebih bersih dan akurat untuk analisis.
3. Normalisasi Teks. Semua ulasan dikonversi ke huruf kecil menggunakan metode '*str.lower()*'. Normalisasi ini penting karena mengurangi kompleksitas model dengan menghilangkan perbedaan antara huruf besar dan kecil yang tidak mempengaruhi makna teks.
4. Pembersihan Karakter Non-Alfabetik. Karakter yang bukan huruf, seperti angka dan simbol, dianggap tidak relevan untuk analisis emosi dan oleh karena itu diganti dengan spasi. Hal ini dilakukan dengan metode '*str.replace()*' dan ekspresi reguler '*^[^a-z\s]*', yang mencari karakter non-alfabetik dan menggantinya.

5. Penghilangan Spasi Berlebih. Teks ulasan sering kali mengandung spasi ekstra yang tidak perlu, yang jika tidak dihilangkan, dapat menyebabkan inkonsistensi dalam pemrosesan data lebih lanjut. Oleh karena itu, penghilangan spasi berlebih dilakukan dengan menerapkan fungsi *lambda* yang memanfaatkan `re.sub()` untuk menggantikan satu atau lebih spasi (`\s+`) dengan spasi tunggal dan kemudian memangkas spasi di awal dan akhir *string*.

3.2 Data Cleaning

Persiapan data adalah tahapan penting yang memfasilitasi transisi dari data mentah ke format yang dapat diolah oleh model pembelajaran mesin. Langkah ini harus dilakukan dengan teliti untuk memastikan bahwa input yang diberikan ke model mampu merefleksikan informasi penting dari teks asli sambil menghilangkan aspek yang tidak penting atau berpotensi mengaburkan pembelajaran yang efektif. Proses ini melibatkan konversi teks ulasan menjadi *sekuens* angka yang kemudian di-*padding* untuk memastikan bahwa semua input memiliki panjang yang sama sebelum dilatih oleh model, yang dijelaskan pada Gambar 2.

```

01 BEGIN
02     // Memuat pustaka yang diperlukan
03     IMPORT numpy, pandas, scipy.stats
04     IMPORT keras.preprocessing.text, keras.preprocessing.sequence
05     IMPORT sklearn.preprocessing
06
07     // Mengasumsikan 'data' adalah DataFrame yang berisi kolom 'Customer
08     Review' dan 'Emotion'
09
10     // Mengonversi 'Customer Review' menjadi daftar string
11     SET text = data["Customer Review"].astype(str).tolist()
12
13     // Mengonversi label 'Emotion' menjadi daftar
14     SET labels = data['Emotion'].tolist()
15
16     // Parameter
17     SET vocab_size = 7000
18     SET maxlen = 40
19
20     // Tokenisasi teks
21     SET tokenizer =
22     keras.preprocessing.text.Tokenizer(num_words=vocab_size)
23     tokenizer.fit_on_texts(text)
24     SET sequences = tokenizer.texts_to_sequences(text)
25
26     // Padding urutan
27     SET x_data = keras.preprocessing.sequence.pad_sequences(sequences,
28     maxlen=maxlen, padding='post')
29
30     // Mengkodekan label
31     SET label_encoder = sklearn.preprocessing.LabelEncoder()
32     SET y_data = label_encoder.fit_transform(labels)
33     SET y_data = y_data.reshape(-1, 1)
34
35     // Membagi dataset menjadi set pelatihan dan validasi
36     SET x_train, x_val, y_train, y_val =
37     sklearn.model_selection.train_test_split(x_data, y_data, test_size=0.2,
38     random_state=42)
39     END

```

Gambar 2 Pseudocode untuk Data Cleaning

Setelah pra-pemrosesan, dataset dibagi menjadi set pelatihan dan validasi menggunakan fungsi `train_test_split` dengan `test_size=0.2` dan `random_state=42`, memastikan pembagian data yang konsisten untuk reproduktibilitas eksperimen. Ini menghasilkan 80% data untuk pelatihan dan 20% untuk validasi, dengan 500 sampel tambahan yang disiapkan untuk pengujian.

Pembagian ini penting untuk memvalidasi kinerja model dan mencegah overfitting. Ini menetapkan dasar yang solid untuk pelatihan model, dengan total 5400 sampel awal yang disiapkan, 3876 untuk pelatihan, 969 untuk validasi, dan 500 untuk pengujian, memastikan dataset siap untuk tahap pelatihan model selanjutnya.

3.3 Pembangunan Model Transformer

Proses pembangunan model Transformer untuk tugas pengenalan emosi dari ulasan pelanggan e-commerce menggabungkan teknik-teknik pemrosesan bahasa terkini dengan arsitektur jaringan saraf yang inovatif. Proses pembangunan model, dengan fokus pada setiap komponen dan parameter yang digunakan untuk menyusun arsitektur Transformer yang efektif dan adaptif, mengikuti *pseudocode* pada Gambar 3

```

01 BEGIN
02     // Define model parameters
03     SET embed_dim = 32
04     SET num_heads = 2
05     SET ff_dim = 32
06     SET regularizer = 12(0.001)
07
08     // Build the model
09     SET inputs = layers.Input(shape=maxlen)
10     SET embedding_layer = TokenAndPositionEmbedding(maxlen, vocab_size,
11 embed_dim)
12     SET x = embedding_layer(inputs)
13     SET transformer_block = TransformerBlock(embed_dim, num_heads,
14 ff_dim)
15     SET x = transformer_block(x)
16     SET x = layers.GlobalAveragePooling1D()(x)
17     SET x = layers.Dropout(0.2)(x)
18     SET x = layers.Dense(128, activation="relu")(x)
19     SET outputs = layers.Dense(5, activation="SoftMax")(x)
20     SET model = keras.Model(inputs=inputs, outputs=outputs)
21 END

```

Gambar 3 Pseudocode untuk Membangun Model Transformer

Dalam rangka memastikan efektivitas model Transformer dalam mengklasifikasikan emosi dari ulasan pelanggan *e-commerce*, menggunakan strategi pelatihan yang meliputi pengaturan *class weights*, *inisialisasi callback* untuk *early stopping*, penyesuaian *optimizer*, dan proses *fitting* model.

3.3.1 Penetapan Class Weights

Karena distribusi kelas dalam dataset tidak seimbang, penetapan *class weights* menjadi langkah penting untuk memastikan bahwa model tidak bias terhadap kelas yang lebih dominan. Dengan menggunakan fungsi `compute_class_weight` dari *scikit-learn*, *class weights* dihitung berdasarkan distribusi frekuensi kelas dalam data pelatihan. Strategi ini menyeimbangkan kelas dengan memberikan bobot lebih kepada kelas yang kurang terwakili, memastikan bahwa setiap kelas memiliki pengaruh yang setara selama proses pelatihan. *Class weights* ini kemudian disimpan dalam `class_weight_dict`, memungkinkan mereka untuk diterapkan langsung ke model selama proses *fitting*.

3.3.2 Callback Early Stopping

Untuk menghindari *overfitting* dan memastikan bahwa model tidak terlalu menyesuaikan diri dengan data pelatihan, *callback early stopping* diterapkan. Ini diinisialisasi dengan parameter `patience=15`, yang berarti pelatihan akan dihentikan jika tidak ada peningkatan pada *validation loss* selama 15 *epoch* berturut-turut. *Callback* ini juga dikonfigurasi untuk mengembalikan bobot model terbaik yang tercapai sebelum terjadinya peningkatan, berkat parameter

``restore_best_weights=True``.

3.3.3 Optimizer dan Laju Pembelajaran

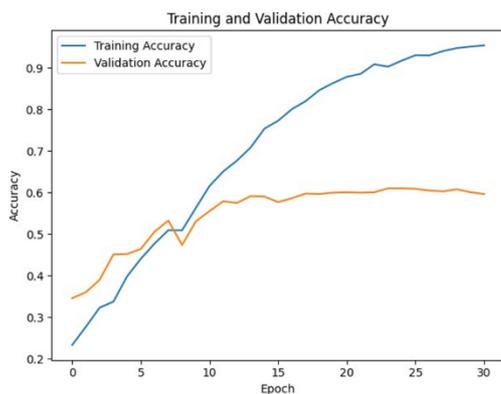
Optimizer Adam dipilih karena kemampuannya yang telah terbukti dalam menangani berbagai tugas optimasi dalam *deep learning*. Dengan laju pembelajaran yang diatur ke 0.0001, optimizer ini bertujuan untuk membuat perubahan bertahap dan konsisten pada bobot model, mencari titik optimal dalam ruang parameter model dengan cara yang efisien dan efektif.

3.3.4 Kompilasi dan Pelatihan Model

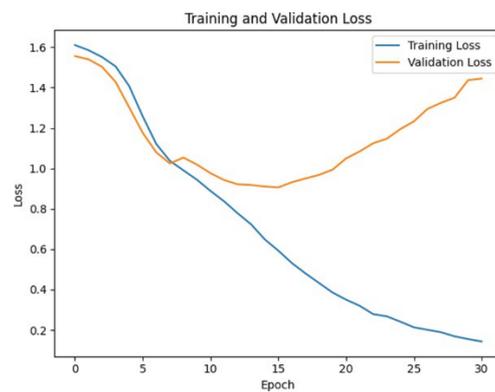
Model dikompilasi menggunakan *loss function sparse_categorical_crossentropy*, yang cocok untuk masalah klasifikasi dengan label kelas sebagai bilangan bulat, dan metrik *accuracy* untuk memantau kinerja model. Model kemudian dilatih menggunakan metode *fit*, dengan data pelatihan dan validasi yang telah dipersiapkan sebelumnya, *batch size* sebesar 32, dan jumlah *epoch* sebanyak 200. Selama pelatihan, *class weights* diterapkan untuk menyesuaikan pengaruh sample dari setiap kelas, dan *callback early stopping* dimasukkan untuk memonitor kinerja model dan menghentikan pelatihan jika diperlukan.

3.4 Performa Model

Analisis performa model merupakan tahapan krusial untuk mengukur efektivitas algoritma Transformer dalam klasifikasi emosi dari ulasan pelanggan *e-commerce*. Hasil evaluasi akurasi terlihat pada Gambar 4 yang menampilkan tren yang meningkat selama pelatihan, mengindikasikan bahwa model secara konsisten memperbaiki kemampuan prediksinya setiap *epoch*. Namun, validasi akurasi mengalami fluktuasi dan menunjukkan peningkatan yang lebih sedang, menggambarkan tantangan model dalam menggeneralisasi pembelajaran ke data yang belum pernah dilihat. Selisih antara akurasi pelatihan dan validasi mengindikasikan adanya *overfitting* yang mungkin terjadi, meskipun ini diatasi sebagian dengan penggunaan *callback arly stopping*. Sedangkan hasil evaluasi loss selama pelatihan dan validasi tercermin dalam Gambar 5, yang memberikan perspektif lain dari pembelajaran model. *Loss* pelatihan yang terus menurun mengindikasikan bahwa model secara efektif meminimalkan kesalahan prediksi pada data pelatihan. Namun, *loss* validasi tidak menunjukkan penurunan yang konsisten, melainkan cenderung stabil setelah penurunan awal. Ini dapat menunjukkan bahwa model mungkin telah mencapai batas kemampuannya dalam menangkap pola dari data validasi, atau mungkin memerlukan penyesuaian lebih lanjut pada arsitektur atau *hyperparameter* untuk meningkatkan generalisasi.



Gambar 4 Grafik Akurasi Model Transformer



Gambar 5 Grafik Loss Model Transformer

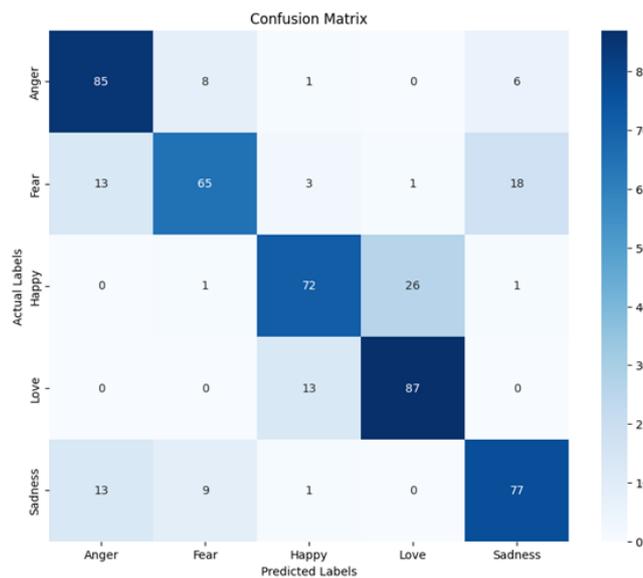
Dari hasil yang diamati, model menunjukkan kemampuan yang kuat dalam memahami data pelatihan, tetapi menghadapi tantangan dalam mengaplikasikan pembelajaran tersebut secara luas. Ini menggarisbawahi pentingnya teknik pelatihan yang tepat, termasuk penyetelan *hyperparameter*, *regularisasi* yang tepat, dan strategi pelatihan untuk menghindari *overfitting*.

3.5 Model Evaluation

Dalam proses evaluasi model dengan data pengujian, fokus utama adalah untuk mengukur sejauh mana model Transformer yang telah dilatih mampu mengenali dan mengklasifikasikan emosi yang terkandung dalam ulasan pelanggan secara akurat. Evaluasi ini penting untuk memahami bagaimana model akan beroperasi dalam lingkungan nyata, di mana ketepatan dalam mengidentifikasi emosi dari teks adalah kunci untuk memperoleh insight yang berharga dari data pelanggan. Evaluasi model dengan data pengujian ini dilakukan menggunakan dataset testing sebanyak 500 data dengan masing-masing kelasnya 100 data.

3.5.1 Confusion Matrix

Confusion matrix adalah alat yang sangat penting dalam machine learning untuk mengukur kinerja model klasifikasi. *Confusion matrix* yang dihasilkan oleh model klasifikasi emosi dari ulasan pelanggan *e-commerce* tersaji pada Gambar 6.



Gambar 6 Confusion Matrix Model Transformer

Pada Gambar 6, model menunjukkan kemampuan yang baik dalam mengidentifikasi 'Anger' dan 'Love' dengan *True Positives* (TP) yang tinggi, namun menemui kesulitan dalam membedakan antar emosi serupa, terutama 'Fear', 'Happy', dan 'Sadness'. Misalnya, model efektif mendeteksi 'Anger' dengan 85 TP, tetapi memiliki *False Negatives* (FN) 15, menandakan beberapa nuansa kemarahan tidak teridentifikasi. Demikian pula, meskipun 'Love' dikenali dengan 87 TP, kesalahan positif pada emosi lain menunjukkan adanya overlap dalam ekspresi linguistik. *False Positives* (FP) dan FN dalam kategori lain menunjukkan tantangan dalam menginterpretasikan ekspresi emosi yang kompleks, seperti kesedihan yang sering kali tertukar dengan emosi negatif lain.

Analisis ini menyoroti kekuatan model dalam memahami beberapa emosi sambil menandai kebutuhan untuk meningkatkan sensitivitasnya terhadap ekspresi emosi yang lebih halus. Misalnya, 'Sadness' diketahui dengan 77 TP, namun 23 FN mengindikasikan bahwa konteks atau kata-kata kunci tertentu yang menandakan kesedihan mungkin tidak sepenuhnya diperhitungkan. Ini menegaskan pentingnya pengembangan lebih lanjut pada tahap pra-pemrosesan data dan adaptasi teknik ekstraksi fitur yang lebih kompleks untuk memperkaya pemahaman model terhadap berbagai nuansa emosi dalam ulasan, serta penyesuaian model dengan dataset yang lebih variatif untuk meningkatkan generalisasi dan keakuratan dalam klasifikasi emosi.

3.5.2 Metrik Evaluasi Model

Setelah evaluasi menggunakan *confusion matrix*, selanjutnya adalah evaluasi hasil model untuk mengetahui mengenai performa model klasifikasi di setiap kelas. Evaluasi ini membandingkan output prediksi model dengan label sebenarnya dan menghasilkan metrik kunci: *presisi*, *recall*, dan skor F1 untuk setiap kelas emosi.

Tabel 1 Evaluasi Hasil Model Transformer

Emosi	Akurasi	Precision	Recall	F1-Score
Anger	0.850	0.766	0.850	0.805
Fear	0.650	0.783	0.650	0.711
Happy	0.720	0.800	0.720	0.758
Love	0.870	0.763	0.870	0.812
Sadness	0.770	0.755	0.770	0.763
Total	0.772	0.772	0.772	0.772

Evaluasi hasil model klasifikasi emosi dari ulasan pelanggan e-commerce menunjukkan performa variatif di antara kelas emosi, dengan metrik utama seperti presisi, recall, dan skor F1 yang masing-masing mencerminkan kekuatan dan kelemahan model dalam mengidentifikasi emosi tertentu. Model ini paling efektif dalam mengenali 'Love' dengan akurasi 87% dan skor F1 81.2%, menunjukkan kehandalan dalam mengidentifikasi ekspresi kasih sayang. Namun, model menemui tantangan lebih besar dalam mengklasifikasikan 'Fear' dan 'Happy', dengan akurasi 65% dan 72% serta skor F1 71.1% dan 75.8%, mengindikasikan kesulitan dalam membedakan nuansa ekspresi ketakutan dan kebahagiaan dari emosi lain.

Kinerja keseluruhan model, dengan akurasi rata-rata 77.2% dan skor F1 yang sebanding, menunjukkan kemampuan yang cukup baik dalam mengenali emosi dalam ulasan, tetapi juga menggarisbawahi area untuk peningkatan, khususnya dalam meningkatkan sensitivitas model terhadap ekspresi emosi yang lebih halus dan kompleks. Tantangan seperti meminimalkan kesalahan positif dalam prediksi 'Fear' dan 'Happy' menegaskan kebutuhan untuk strategi pra-pemrosesan data yang lebih canggih atau penyesuaian arsitektur model untuk meningkatkan pemahaman terhadap berbagai nuansa emosi.

4. KESIMPULAN DAN SARAN

Kesimpulan harus mengindikasikan secara jelas hasil-hasil yang diperoleh, kelebihan dan Penelitian ini berhasil menggambarkan efektivitas model deep learning berbasis Transformer dalam klasifikasi emosi ulasan pelanggan e-commerce berbahasa Indonesia, dengan temuan menarik bahwa emosi positif seperti 'Happy' lebih sering diungkapkan dalam ulasan. Model menunjukkan performa yang kuat secara keseluruhan dengan akurasi 77,2%, dengan 'Love' sebagai emosi dengan pengenalan terbaik. Variabilitas dalam performa antar kategori emosi menyoroti potensi model dalam memahami nuansa emosional teks dan kebutuhan akan peningkatan dalam mengenali emosi seperti 'Fear' dan 'Happy'. Temuan ini memperkuat potensi model berbasis Transformer untuk analisis sentimen dalam e-commerce, meskipun terdapat ruang untuk optimasi dan peningkatan keakuratan.

Rekomendasi untuk penelitian masa depan mencakup pengembangan model dengan dataset yang lebih besar dan variatif untuk meningkatkan keakuratan dan keandalan, optimisasi *hyperparameter* untuk *eksplorasi konfigurasi* optimal model Transformer, dan pendekatan multimodal untuk menggabungkan berbagai jenis data dalam pengenalan emosi. Eksplorasi ini diharapkan tidak hanya memajukan pemahaman teknis tentang pengolahan bahasa alami dan analisis emosi tetapi juga mendorong implementasi praktis yang dapat meningkatkan interaksi pelanggan dalam e-commerce dan sektor lainnya.

DAFTAR PUSTAKA

- [1] S. N. Alsubari, "Data Analytics for the Identification of Fake Reviews Using Supervised Learning," *Comput. Mater. Contin.*, vol. 70, no. 2, pp. 3189–3204, 2022, doi: 10.32604/cmc.2022.019625.
- [2] E. A. Grigoreva, L. F. Garifova, and E. A. Polovkina, "Consumer Behavior in the Information Economy: Generation Z," *Int. J. Financ. Res.*, vol. 12, no. 2, p. 164, Jan. 2021, doi: 10.5430/ijfr.v12n2p164.
- [3] J. Guo, X. Wang, and Y. Wu, "Positive emotion bias: Role of emotional content from online customer reviews in purchase decisions," *J. Retail. Consum. Serv.*, vol. 52, p. 101891, Jan. 2020, doi: 10.1016/j.jretconser.2019.101891.
- [4] M. D. Rocklage and R. H. Fazio, "The Enhancing Versus Backfiring Effects of Positive Emotion in Consumer Reviews," *J. Mark. Res.*, vol. 57, no. 2, pp. 332–352, Apr. 2020, doi: 10.1177/0022243719892594.
- [5] H. J. Katole, "A research article on Importance of Online Customer Reviews on Customer Purchase," *Bus. Manag. Rev.*, vol. 13, no. 03, pp. 10–11, 2022, doi: 10.24052/bmr/v13nu03/art-04.
- [6] C. Changchit and T. Klaus, "Determinants and Impact of Online Reviews on Product Satisfaction," *J. Internet Commer.*, vol. 19, no. 1, pp. 82–102, Jan. 2020, doi: 10.1080/15332861.2019.1672135.
- [7] D. Hussen Maulud, S. R. M. Zeebaree, K. Jacksi, M. A. Mohammed Sadeeq, and K. Hussein Sharif, "State of Art for Semantic Analysis of Natural Language Processing," *Qubahan Acad. J.*, vol. 1, no. 2, pp. 21–28, Mar. 2021, doi: 10.48161/qaj.v1n2a44.
- [8] A. Vaswani *et al.*, "Attention is all you need," *Adv. Neural Inf. Process. Syst.*, vol. 2017-December, no. Nips, pp. 5999–6009, 2017.
- [9] J. Sun, J. Xie, and H. Zhou, "EEG Classification with Transformer-Based Models," in *2021 IEEE 3rd Global Conference on Life Sciences and Technologies (LifeTech)*, Mar. 2021, pp. 92–93. doi: 10.1109/LifeTech52111.2021.9391844.
- [10] A. Purohit, "Sentiment Analysis of Customer Product Reviews using deep Learning and Compare with other Machine Learning Techniques," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 9, no. VII, pp. 233–239, Jul. 2021, doi: 10.22214/ijraset.2021.36202.
- [11] A. F. Ab. Nasir *et al.*, "Text-based emotion prediction system using machine learning approach," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 769, no. 1, p. 012022, Feb. 2020, doi: 10.1088/1757-899X/769/1/012022.
- [12] S. K. Bharti *et al.*, "Text-Based Emotion Recognition Using Deep Learning Approach," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–8, Aug. 2022, doi: 10.1155/2022/2645381.
- [13] B. T. Atmaja, K. Shirai, and M. Akagi, "Deep Learning-based Categorical and Dimensional Emotion Recognition for Written and Spoken Text," *IPTEK J. Proc. Ser.*, pp. 1–10, 2019.
- [14] A. K. Durairaj and A. Chinnalagu, "Transformer based Contextual Model for Sentiment Analysis of Customer Reviews: A Fine-tuned BERT," *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 11, 2021, doi: 10.14569/IJACSA.2021.0121153.
- [15] L. Davoodi and J. Mezei, "a Comparative Study of Machine Learning Models for Sentiment Analysis: Customer Reviews of E-Commerce Platforms," *35th Bled eConference Digit. Restruct. Hum. (Re)action, BLED 2022 - Proc.*, pp. 217–231, 2022, doi: 10.18690/um.fov.4.2022.13.
- [16] R. Sutoyo, S. Achmad, A. Chowanda, E. W. Andangsari, and S. M. Isa, "PRDECT-ID: Indonesian product reviews dataset for emotions classification tasks," *Data Br.*, vol. 44, p. 108554, Oct. 2022, doi: 10.1016/j.dib.2022.108554.