MULTI TASK HYBRID INDOBERT CNN UNTUK KLASIFIKASI SENTIMEN DAN EMOSI PADA ULASAN E-COMMERCE BERBAHASA INDONESIA

Tira Parida[1], Taswanda Taryo[2], Sudarno Wiharjo[3]

[12,]Program Studi Teknik Informatika S-2, Universitas Pamulang

Universitas Pamulang, Jl. Raya Puspitek, Buaran, Kec. Pamulang, Kota Tangerang Selatan, Banten 15310.
[1] tiraprd29@gmail.com, [2] taswandataryo@gmail.com, [3] sudarnogm@gmail.com

Abstract

Consumer review analysis on e-commerce platforms plays an important role in identifying customer satisfaction and evaluating product quality. However, this task is challenged by the large volume of textual data, language ambiguity, and semantic complexity contained in customer reviews. This study aims to simultaneously identify sentiment polarity and classify specific emotion categories from customer opinions regarding electronic products on the tokopedia e-commerce platform. The proposed method employs a hybrid multi-task learning approach that combines the IndoBERT architecture with a convolutional neural network (CNN), enabling sentiment and emotion classification to be performed simultaneously within a single model. The research object consists of Indonesian-language customer review documents obtained from the PRDECT-ID dataset, representing customer reviews of electronic products on the tokopedia platform, with a total of 5,400 data instances divided into 3,779 training data, 810 validation data, and 811 testing data. The data were processed through several text preprocessing stages, including case folding, cleaning, and excessive whitespace removal. Subsequently, the processed data were analyzed using the proposed hybrid multi-task learning approach, which integrates IndoBERT embeddings as a shared feature extractor, parallel feature extraction through the IndoBERT encoder for capturing global contextual information and a convolutional neural network (CNN) for extracting local spatial features, followed by a feature fusion mechanism. The study evaluated a multi-task learning model optimized using a joint loss function. Experimental results based on core metrics computation demonstrated that the proposed model, trained for five epochs, achieved a sentiment classification accuracy of 98.15% and an emotion classification accuracy of 65.19%. Furthermore, the model obtained a multi-task hamming loss of 0.1833, indicating its effectiveness in simultaneously performing sentiment and emotion classification. The findings conclude that the proposed hybrid multi-task learning approach based on IndoBERT and CNN enhances semantic understanding of customer reviews while contributing to the advancement of natural language processing (NLP), particularly in sentiment analysis and emotion classification for customer reviews of electronic products on the tokopedia e-commerce platform.

Keywords : *Sentiment Analysis, Convolutional Neural Network, IndoBERT, Emotion Classification, Multi-task Learning.*

Abstrak

Analisis terhadap ulasan konsumen pada platform *e-commerce* penting dilakukan untuk mengidentifikasi tingkat kepuasan pelanggan serta mengevaluasi kualitas produk, namun tantangan muncul akibat tingginya volume teks, ambiguitas bahasa, serta kompleksitas semantik yang terkandung di dalamnya. Penelitian ini bertujuan untuk mengidentifikasi polaritas sentimen dan mengklasifikasikan jenis emosi spesifik secara simultan dari opini pelanggan terhadap produk elektronik pada platform *e-commerce* tokopedia. Metode yang digunakan dalam penelitian ini adalah *hybrid multi-task learning* dengan arsitektur gabungan IndoBERT dan *convolutional neural network* (CNN), yang memungkinkan klasifikasi sentimen dan emosi dilakukan secara bersamaan dalam satu model. Objek penelitian berupa kumpulan dokumen teks ulasan berbahasa Indonesia yang bersumber dari dataset PRDECT-ID, yang merepresentasikan ulasan pelanggan terhadap produk elektronik pada platform tokopedia, dengan total 5.400 data yang terbagi menjadi data latih (3.779), data validasi (810), dan data uji (811). Data diproses melalui tahapan prapemrosesan teks yang meliputi *case folding, cleaning,* dan eliminasi spasi berlebih. Data yang telah diproses kemudian dianalisis menggunakan metode *hybrid multi-task learning* yang memadukan

embedding IndoBERT sebagai *shared feature extractor*, teknik *feature extraction* paralel melalui *encoder* IndoBERT untuk menangkap konteks global dan *convolutional neural network* (CNN) untuk mengekstraksi fitur lokal, serta mekanisme *feature fusion*. Penelitian ini menguji skenario pemodelan tugas ganda (*multi-task*) yang dioptimalkan menggunakan *joint loss function*. Hasil evaluasi berdasarkan *core metrics computation* menunjukkan bahwa model gabungan yang dilatih selama lima *epoch* menghasilkan akurasi klasifikasi sentimen sebesar 98,15% dan akurasi klasifikasi emosi sebesar 65,19%. Selain itu, model menghasilkan nilai *multi-task hamming loss* sebesar 0,1833, yang menunjukkan kemampuan model dalam melakukan klasifikasi sentimen dan emosi secara simultan. Hasil penelitian ini menyimpulkan bahwa pendekatan *hybrid multi-task learning* berbasis IndoBERT dan CNN mampu meningkatkan pemahaman semantik terhadap ulasan pelanggan secara paralel serta memberikan kontribusi pada pengembangan metode *natural language processing* (NLP), khususnya dalam analisis sentimen dan klasifikasi emosi pada ulasan produk elektronik di platform *e-commerce* tokopedia.

Kata kunci : *Analisis Sentimen, Convolutional Neural Network, IndoBERT, Klasifikasi Emosi, Multi-task Learning.*

1. PENDAHULUAN

Perkembangan pesat teknologi informasi dan komunikasi telah mendorong perubahan signifikan dalam pola perilaku konsumen. Platform *e-commerce* di Indonesia mengalami pertumbuhan yang luar biasa dalam satu dekade terakhir. Konsumen kini tidak hanya melakukan transaksi jual beli, tetapi juga aktif memberikan ulasan (*review*) yang menjadi sumber informasi penting bagi pembeli lain dan bahan evaluasi strategis bagi penjual. Menurut laporan *We Are Social* (2024), jumlah pengguna internet Indonesia telah melampaui 213 juta, dan sebagian besar aktivitas digital terkait dengan perdagangan daring. Proporsi data yang dapat dialokasikan dalam penelitian ini terdiri dari 80% data latih untuk proses pembentukan model. Selama fase pelatihan, arsitektur CNN diimplementasikan guna mengekstraksi pola-pola informatif dari data yang nantinya berfungsi untuk klasifikasi sentimen. Pemanfaatan volume data latih yang dominan ini memungkinkan model untuk menangkap dan mempelajari fitur serta karakteristik unik yang terkandung dalam setiap ulasan pelanggan secara *komprehensif* [1].

Dataset yang digunakan dalam penelitian ini adalah PRDECT-ID, yang berisi 5.400 ulasan produk *e-commerce* berbahasa Indonesia yang diambil dari berbagai kategori produk di platform *e-commerce* lokal. Setiap entri mencakup informasi seperti kategori produk, nama produk, lokasi penjual, harga, rating, serta isi ulasan pelanggan. Setiap ulasan telah dilengkapi dengan dua label utama, yaitu polaritas sentimen (terdiri dari 2.579 ulasan positif dan 2.821 ulasan negatif) serta kategori emosi (seperti senang, kecewa, marah, dan sebagainya) yang menggambarkan respons emosional konsumen terhadap produk. Dataset ini diproses melalui tahapan pembersihan teks, tokenisasi, serta normalisasi kata tidak baku untuk memastikan kualitas data sebelum dimasukkan ke dalam model. Pemilihan dataset PRDECT-ID dilakukan karena dataset ini secara khusus dikembangkan untuk penelitian analisis sentimen dan emosi dalam konteks bahasa Indonesia pada domain *e-commerce*, sehingga sangat relevan dengan tujuan penelitian ini.

Ulasan produk berperan krusial dalam memengaruhi keputusan pembelian. Konsumen cenderung mempercayai pengalaman pengguna lain dibandingkan promosi langsung dari penjual. Oleh karena itu, analisis ulasan produk tidak hanya membantu memahami persepsi konsumen terhadap kualitas produk dan layanan, tetapi juga mendukung perusahaan dalam meningkatkan strategi bisnis berbasis data. Namun, volume ulasan yang sangat besar dan keragaman ekspresi bahasa Indonesia termasuk penggunaan kosakata informal, singkatan, emotikon, hingga campuran Bahasa (*code-mixing*) menjadi tantangan dalam melakukan analisis sentimen secara akurat. Lebih lanjut lagi, Ulasan *online* pada *e-commerce* berfungsi sebagai bentuk *electronic Word of Mouth* (e-WOM) yang efektif dalam mengurangi ketidakpastian konsumen saat hendak membeli suatu produk. Studi terbaru menunjukkan bahwa ulasan toko atau produk tidak hanya memiliki dampak langsung terhadap keputusan pembelian, tetapi juga secara tidak langsung membangun kepercayaan konsumen terhadap produk. Kepercayaan ini pada akhirnya dapat meningkatkan kemungkinan pembelian ulang (*re-order*). Namun demikian, ulasan-ulasan pada

e-commerce bisa sangat beragam, mulai dari ulasan singkat hingga panjang dan ekspresif. Ketika sebuah produk memiliki ratusan hingga ribuan ulasan, tentu tidak realistis bagi pelaku usaha untuk membaca semuanya satu per satu . Oleh karena itu, dibutuhkan metode otomatis yang mampu memahami dan meringkas opini pelanggan secara ringkas namun tetap informatif, terutama dalam mempertahankan makna emosional dari setiap ulasan [2].

Pendekatan tradisional berbasis *lexicon* atau metode pembelajaran mesin sederhana (misalnya *Naive Bayes*, SVM) sering kali kurang mampu menangkap konteks makna kata secara mendalam. Untuk mengatasi keterbatasan ini, teknologi *deep learning* seperti IndoBERT dan *Convolutional Neural Network* (CNN) menjadi solusi yang menjanjikan. IndoBERT unggul dalam memahami hubungan antar-kata secara global melalui mekanisme *self-attention*, sedangkan CNN efektif dalam mengekstraksi fitur lokal yang relevan, terutama pola frasa pendek yang kerap muncul dalam ulasan. Penelitian ini mengeksplorasi potensi algoritma CNN untuk mengetahui tingkat akurasi pada algoritma CNN. Hasil penelitian ini diharapkan dapat dijadikan acuan perusahaan untuk memilih algoritme dengan akurasi tertinggi untuk menganalisis sentimen pengguna agar lebih akurat dan efisien dalam mengelompokan ulasan agar terus konsisten memberikan layanan terbaik atau meningkatkan layanan perusahaan berdasarkan data ulasan pengguna di *Google Playstore* [3].

Lebih lanjut, pendekatan *multi-task learning* memungkinkan sebuah model dilatih untuk menyelesaikan dua tugas sekaligus. Dalam konteks ini, model dapat dikembangkan untuk mengklasifikasikan polaritas sentimen (positif, negatif, netral) dan kategori emosi (misalnya senang, marah, kecewa) secara simultan. Dalam pengambilan keputusan, emosi manusia memiliki peranan mendasar yang tidak dapat dikesampingkan. Pada lanskap bisnis digital saat ini, ulasan pelanggan di platform *e-commerce* telah diakui sebagai variabel krusial yang menentukan keputusan belanja. Konsumen umumnya meyakini bahwa ulasan yang sarat dengan muatan emosional memberikan pengaruh persuasif yang lebih tinggi terhadap calon pembeli lain. Seiring dengan perkembangan media sosial dan digital, ulasan produk telah bertransformasi menjadi sumber informasi utama yang memandu keputusan konsumen. Oleh karena itu, identifikasi emosi yang spesifik seperti rasa senang, sedih, atau amarah menjadi sangat relevan untuk dilakukan. Sebagai contoh, ulasan berlabel sentimen positif dapat membawa emosi berupa rasa lega atau kegembiraan, sementara sentimen negatif bisa jadi mencerminkan kekecewaan atau frustrasi. Memahami dinamika ini secara mendalam sangat diperlukan untuk menelaah perilaku konsumen di tengah persaingan pasar *e-commerce* yang kian sengit.

Teknologi *natural language processing* (NLP) kini hadir sebagai solusi untuk mengekstrak makna dan muatan emosional dari data tekstual. Berkat dukungan model pelatihan yang kompleks serta algoritma mutakhir, NLP mampu mengidentifikasi pola *linguistik*, menangkap konteks, hingga menafsirkan emosi yang tersirat dalam suatu teks. Pemanfaatan NLP pada ulasan produk memberikan kesempatan berharga untuk memahami sentimen pelanggan secara masif, yang sebelumnya sulit dilakukan melalui analisis manual. Dengan menerapkan NLP, perusahaan dapat memperoleh wawasan mendalam mengenai opini dan preferensi konsumen, pada akhirnya menjadi landasan strategis bagi pengambilan keputusan bisnis yang lebih cerdas dan efektif [4].

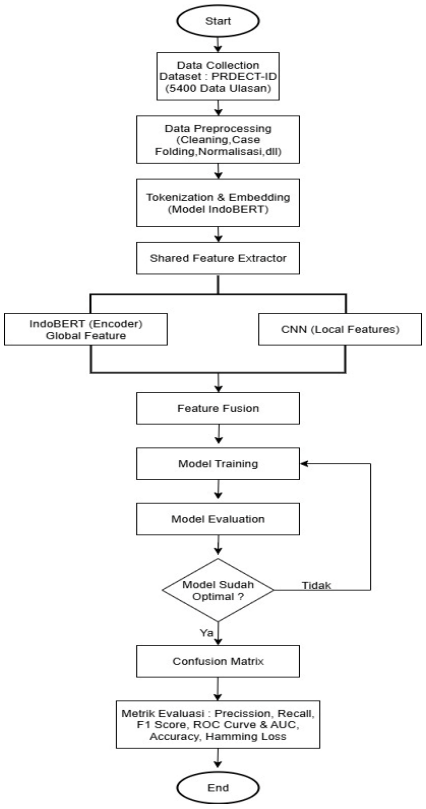
Beberapa penelitian terdahulu telah berhasil menerapkan model IndoBERT maupun *convolutional neural network* (CNN) untuk analisis sentimen secara terpisah pada data ulasan berbahasa Indonesia. Meskipun mampu menghasilkan performa yang baik, sebagian besar penelitian masih berfokus pada satu tugas klasifikasi (*single-task*), yaitu hanya mengidentifikasi polaritas sentimen atau hanya melakukan klasifikasi emosi. Selain itu, penelitian yang mengintegrasikan representasi kontekstual IndoBERT dengan kemampuan CNN dalam mengekstraksi fitur lokal melalui pendekatan *multi-task learning* pada ulasan *e-commerce* berbahasa Indonesia masih sangat terbatas. Kondisi tersebut menunjukkan masih adanya kesenjangan penelitian (*research gap*) dalam pengembangan model yang mampu melakukan klasifikasi sentimen dan emosi secara simultan dalam satu arsitektur terpadu.[5]. Selain itu, integrasi *multi-task learning* dalam satu kerangka pemodelan untuk klasifikasi sentimen dan emosi jarang ditemukan dalam literatur nasional. Berdasarkan *research gap* tersebut, penelitian ini menawarkan kebaruan berupa pengembangan model *multi-task hybrid* IndoBERT-CNN yang mengintegrasikan kemampuan IndoBERT dalam memahami konteks semantik global dengan kemampuan CNN dalam mengekstraksi fitur lokal melalui mekanisme *feature fusion*. Berbeda dengan penelitian sebelumnya yang umumnya

hanya berfokus pada satu tugas klasifikasi, model yang diusulkan mampu melakukan klasifikasi sentimen dan emosi secara simultan menggunakan satu arsitektur *multi-task learning*. Pendekatan ini diharapkan mampu menghasilkan representasi teks yang lebih komprehensif sehingga meningkatkan efektivitas analisis ulasan *e-commerce* berbahasa Indonesia.

2. METODOLOGI PENELITIAN

2.1. Skema Alur Penelitian

Penelitian ini dirancang melalui serangkaian tahapan sistematis sebagaimana ditampilkan dalam Gambar 1. Tujuan utama dari alur penelitian ini adalah membangun model klasifikasi yang mampu mengidentifikasi polaritas sentimen dan kategori emosi pada ulasan produk *e-commerce* berbahasa Indonesia dengan pendekatan *multi-task hybrid* IndoBERT–CNN.



Gambar 1. Digram Rancangan Penelitian

2.2. Pengumpulan Data

Data yang digunakan dalam penelitian ini diperoleh dari repositori data publik *Kaggle* melalui dataset resmi bertajuk "PRDECT-ID : *Indonesian Emotion Classification*". Informasi yang dikumpulkan termasuk kategori produk, nama produk, lokasi penjual, harga, peringkat, total ulasan, serta teks ulasan konsumen. Semua ulasan yang diambil merupakan opini berbahasa Indonesia yang telah dianotasi ke dalam dua target label *multi-task*, yaitu target sentimen dan target emosi. Selama proses pengumpulan data, sebanyak 5.400 dokumen ulasan dikumpulkan berhasil diunduh dari https://www.kaggle.com/datasets/jocelyndumlao/prdect-id-indonesian emotion-classification, dan dikategorikan ke dalam 2 kategori utama yaitu sentimen dan emosi yang mencakup positif dan negatif (Sentimen), serta *happy, sadness, love, anger, fear* (Emosi). Dataset disimpan dalam format *.csv* untuk memudahkan pengolahan data.



	Number Sold	Total Review	Customer	Customer Review	Sentiment	Emotion
5381	2102	789		5 Barang sesuai deskripsi, bagusss lucu , bisa dibawa kemana mana, mudahÇ¦lian awet	Positive	Love
5382	2102	789		5 Barangnya buat kado sih bagus yg terima kelilatan happy :) Kualitas Product Bagus Harga terjangkau! Recomended	Positive	Love
5383	2102	789		5 Paket cepat sampai dan terbungkus rapih dengan bubble. Terima Kasih Toped, Kunr dan SellerKualitas Product Bagus. Material bagus. Harga terjangkau	Positive	Love
5384	2800	2360		5 Packing ok, seller lekas kirim, benguna buat organize koper biar gak acak acakan, semoga awet, thank you seller	Positive	Happy
5385	2800	2360		5 barang sesuai pesanan.. tp seller luar biasa responnya cepat n barang sdh sampai lebih cepat dr estimasi.. sukses selalu n sehat selalu y gan... terimakasi	Positive	Happy
5386	2800	2360		5 Terima kasih untuk menjual travel pouch 1 set dgn berbagai ukuran dan bahan bagus tidak bau plastik. Barang cepat tiba dan dalam keadaan baik. Teri	Positive	Happy
5387	2800	2360		5 Barang sudah diterima..bagus...dim keadaan baik...warna gi sesuai dg yg diinginkan..bungkusnya juga bagus...mantap..suka deh...makasih penjual...mal	Positive	Happy
5388	2661	1201		5 Pengirimannya cepat, besok langsung sampai. Nggak expect apa-apa sebenernya, tapi ternyata bagusss. Bahannya lembut banget, gelap oke buat nutup	Positive	Happy
5389	2661	1201		5 Produk berfungsi dengan baik, harga sesuai dan kualitas baik. Packing rapi banget dan proses pengiriman juga cepat, seller sangat responsif dan ramah,	Positive	Love
5390	2661	1201		5 bahannya enak lembut, ada gel di dalamnya, talinya bisa diatur dan gk menggangggu pas tidur, packing rapih, pengiriman cepat, langsung proses ketika	Positive	Happy
5391	2072	949		5 Terima kasih Gan. Barang diterima aman dan berfungsi baik sesuai pesanan. Kondisi aman semua lampu menyala. Tambahan packing kardus dan bubble	Positive	Happy
5392	2072	949		5 sip, paket sampai dengan selamat, semoga berfungsi dengan baik. terima kasih seller fast responnya.	Positive	Happy
5393	1855	201		5 Waktu dites pake tongkat, jepretannya mantep bgt. Cuma pas beneran dikasi umpan buat tikus, setelah semalem, umpannya ngilang, tapi ga ada tikus	Positive	Happy
5394	1431	562		5 Barusan saya tes untuk ambil sampah di sungai dekat rumah, untuk kekuatan daya cekram si alatnya masuk akal untuk harga segini. untuk ambil plastik	Positive	Happy
5395	1431	562		5 Barangnya sudah sampai dan diterima dengan baik, sudah dicoba dan berfungsi dengan baik, cocok untuk ambil barang dibawah tidak perlu menunduk	Positive	Happy
5396	1431	562		5 produk sesuai,, berguna utk bersihin sampah di jalan depan rumah,, pengiriman cepat,, respon seller cepat,, packing dus+bubble wrap tebal.. tha	Positive	Happy
5397	1431	562		5 Harga dan barang, barang sesuai pesanan. Saya nambah extra packing kardus dan bubble wrap supaya barang aman dan benar sampai dengan selamat alh	Positive	Love
5398	13529	4074		5 Beli ini km Jenak & Istri mau liburan di Jakarta, cari2 ketemu toko ini dengan harga yg masih ramah kantong, untuk pelayanan toko bagus & cepet, pasha	Positive	Love
5399	13529	4074		5 pengemasan barang baik, kondisi barang yg utuh tanpa cacat, berfungsi normal dan lancar, respon dan servis dari seller juga baik	Positive	Happy
5400	13529	4074		5 Mungil tapi bekerja dng baik. Dan murahh terjangkau.. pas dng kebutuhan	Positive	Happy
5401	4785	3044		5 Produk sesuai deskripsi, packing aman terlindung, pengiriman cepat, toko/ penjual responsif, alat sudah dicoba dan berfungsi, sangat direkomendasikan	Positive	Love

Gambar 2. Dataset Ulasan *E-Commerce*.

Data Ulasan yang telah diperoleh dari proses pengumpulan data akan diteruskan untuk diproses pada tahap *preprocessing*. Proses *preprocesing* berisikan beberapa tahapan terdiri dari:

1. *Case Folding*
   Tahapan *case folding* secara umum adalah dimulai dengan memeriksa ukuran setiap karakter dari awal sampai akhir karakter pada teks ulasan konsumen, kemudian jika ditemukan karakter yang menggunakan huruf kapital atau *uppercase*, maka huruf tersebut akan diubah menjadi huruf kecil atau *lowercase*. Hasil dapat dilihat di tabel 1.

TABEL I. HASIL *CASE FOLDING*

No	*Corpus*	*Case Folding*
1.	Produk beroperasi secara optimal, didukung dengan pengemasan yang aman, serta respons penjual dan kurir yang komunikatif, ramah, dan terpercaya .	produk beroperasi secara optimal, didukung dengan pengemasan yang aman, serta respons penjual dan kurir yang komunikatif, ramah, dan terpercaya.
2.	BARANG BAGUS DAN RESPON CEPAT, harga bersaing dengan yg lain.	barang bagus dan respon cepat, harga bersaing dengan yg lain.
3.	Memakainya mudah, hanya saja lumayan juga ya ternyata beratnya kalo sudah lama megangnya pegel juga hehe.. *Anyways*, terima kasih ya Seller :)	memakainya mudah, hanya saja lumayan juga ya ternyata beratnya kalo sudah lama megangnya pegel juga hehe..

2. *Cleaning*
Proses *cleaning* ini melibatkan tahapan yang dilakukan oleh sistem untuk memperoleh hasil yang maksimal, seperti membersihkan tanda baca, membersihkan angka yang berdiri sendiri, membersihkan simbol/emoji, hingga membersihkan kelebihan spasi (*trim space*). Hasil dapat dilihat di tabel 2.

TABEL II. HASIL *CLEANING*

No	*Corpus*	*Cleaning*
1.	produk berfungsi optimal dengan kemasan yang terjamin keamanannya; pelayanan penjual yang responsif dan ramah, serta sikap amanah dari pihak kurir menjadi poin utama dalam ulasan ini.	produk berfungsi optimal dengan kemasan yang terjamin keamanannya pelayanan penjual yang responsif dan ramah serta sikap amanah dari pihak kurir menjadi poin utama dalam ulasan ini
2.	barang bagus dan respon cepat, harga bersaing dengan yg lain.	barang bagus dan respon cepat harga bersaing dengan yg lain
3.	memakainya mudah, hanya saja lumayan juga ya ternyata beratnya kalo sudah lama megangnya pegel juga hehe.. *anyways*, terima kasih ya seller :)	memakainya mudah hanya saja lumayan juga ya ternyata beratnya kalo sudah lama megangnya pegel juga hehe *anyways* terima kasih ya seller

3. *Trim Space*
Proses *trim space* berguna untuk merapikan kerapatan teks korpus dengan cara menghapus spasi ganda atau spasi berlebih di awal, di tengah, maupun di akhir kalimat ulasan yang muncul akibat efek dari penghapusan tanda baca sebelumnya. Hasil dapat dilihat di tabel 3.

TABEL III. HASIL *TRIM SPACE*

No	*Corpus*	*Trim Space*
1.	barang dapat beroperasi secara optimal, didukung oleh kemasan yang aman, serta respons cepat dari pihak penjual dan keramahan serta integritas dari kurir pengiriman	barang dapat beroperasi secara optimal, didukung oleh kemasan yang aman, serta respons cepat dari pihak penjual dan keramahan serta integritas dari kurir pengiriman
2.	barang bagus dan respon cepat harga bersaing dengan yg lain	barang bagus dan respon cepat harga bersaing dengan yg lain

4. *Tokenizing*
Tokenizing merupakan tahap yang berfungsi untuk memisahkan antara kata yang terdapat pada *corpus*. Hasil dapat dilihat di tabel 4.

TABEL IV. HASIL *TOKENIZING*

No	*Corpus*	*Tokenization*
1.	kualitas barang berfungsi secara optimal didukung oleh kemasan yang aman pelayanan penjual yang responsif serta integritas yang ditunjukkan oleh kurir selama proses pengiriman	'kualitas', 'barang', 'berfungsi', 'secara', 'optimal', 'didukung', 'oleh', 'kemasan', 'yang', 'aman', 'pelayanan', 'penjual', 'yang', 'responsif', 'serta', 'integritas', 'yang', 'ditunjukkan', 'oleh', 'kurir', 'selama', 'proses', 'pengiriman'
2.	barang bagus dan respon cepat harga bersaing dengan yg lain	'barang', 'bagus', 'dan', 'respon', 'cepat', 'harga', 'bersaing', 'dengan', 'yg', 'lain'

5. *Normalization*
Tahapan *normalization* (normalisasi data) bertujuan untuk melakukan standarisasi teks ulasan konsumen dengan mengubah kata-kata tidak baku, bahasa gaul (*slang words*), singkatan, maupun kata yang mengalami salah ketik (*typo*) menjadi bentuk kata yang baku sesuai dengan Kamus Besar Bahasa Indonesia (KBBI). Hasil dapat dilihat di tabel 5.

TABEL V. HASIL *NORMALIZATION*

No	*Corpus*	*Normalization*
1.	'syukur', 'produk', 'beroperasi', 'optimal', 'pengemasan', 'terjamin', 'layanan', 'responsif', 'komunikatif', 'penjual', 'serta', 'kurir', 'dapat', 'dipercaya'	'syukur', 'produk', 'beroperasi', 'optimal', 'pengemasan', 'terjamin', 'pelayanan', 'responsif', 'komunikatif', 'penjual', 'serta', 'kurir', 'dapat', 'dipercaya'
2.	'barang', 'bagus', 'dan', 'respon', 'cepat', 'harga', 'bersaing', 'dengan', 'yg', 'lain'	'barang', 'bagus', 'dan', 'respons', 'cepat', 'harga', 'bersaing', 'dengan', 'yang', 'lain'

6. *Stopword Removal*
Proses *Stopword removal* merupakan tahapan prapemrosesan teks yang berfungsi untuk menyaring dan menghapus kata-kata umum (*stopwords*) yang frekuensi kemunculannya tinggi namun tidak memiliki nilai informasi atau kontribusi yang signifikan dalam menentukan bobot sentimen maupun kategori emosi. Kata-kata seperti "dan", "di", "dari", atau "ke". Hasil dapat dilihat di tabel 6.

TABEL VI. HASIL *STOPWORD REMOVAL*

No	*Corpus*	*Stopword Removal*
1.	'syukur', 'produk', 'beroperasi', 'optimal', 'pengemasan', 'terjamin', 'pelayanan', 'responsif', 'komunikatif', 'penjual', 'serta', 'kurir', 'tepercaya'	'syukur', 'produk', 'beroperasi', 'optimal', 'pengemasan', 'terjamin', 'pelayanan', 'responsif', 'komunikatif', 'penjual', 'kurir', 'tepercaya'
2.	'barang', 'bagus', 'dan', 'respons', 'cepat', 'harga', 'bersaing', 'dengan', 'yang', 'lain'	'barang', 'bagus', 'respons', 'cepat', 'harga', 'bersaing'

Setelah melalui proses *preprocessing* dan pemberian identitas, data akan diuji tingkat akurasinya. Langkah ini dilakukan untuk

mengukur efektivitas metode IndoBERT-CNN dalam mengklasifikasikan data ulasan.

2.3. Analisa Data

Data merupakan komponen utama dalam pengembangan model klasifikasi berbasis *deep learning*. Pada penelitian ini, dataset yang digunakan berupa ulasan produk *e-commerce* berbahasa Indonesia yang diperoleh dari sumber *kaggle* data 5400 kumpulan ulasan *e-commerce*. Dataset ini terdiri dari teks ulasan yang mencerminkan opini pengguna terhadap suatu produk, yang digunakan untuk analisis sentimen dan emosi.

Berikut adalah beberapa teknik analisis yang mungkin digunakan dalam penelitian ini :

1) Tahap pertama, pengumpulan data, data *collection*, Sumber Data : Data dikumpulkan dataset publik berbahasa Indonesia untuk klasifikasi emosi yang berasal dari PRDECT-ID Dataset https://www.kaggle.com/datasets/jocelyndumlao/prdect-id-indonesian emotion-classification yang akan digunakan sebagai data yang diolah dalam penelitian yang dibuat.
2) Tahap kedua, yaitu tahap pra-pemrosesan (*preprocessing*) data bertujuan untuk menstrukturkan, merapihkan, dan memastikan data teks ulasan konsumen siap dianalisis untuk klasifikasi *multi-task* sentimen dan emosi produk *e-commerce*. *Preprocessing* data pada penelitian ini dilakukan melalui tiga tahapan utama yang terintegrasi, yaitu *Case Folding, Cleaning, Trim Space, Tokenizing, Normalization*, dan *Stopword Removal.*
3) Tahap *tokenization & embedding,* tahap ini merupakan proses menjembatani data teks ulasan yang telah bersih agar dapat dipahami oleh arsitektur *deep learning*. Proses ini memanfaatkan model *pretrained* IndoBERT sebagai *Shared Feature Extractor* yang mengekstrak fitur semantik bahasa Indonesia.
4) Tahap *Shared Feature Extractor,* merupakan komponen arsitektur di mana luaran vektor (*output embedding*) dari model IndoBERT yang berisi informasi kontekstual ulasan pelanggan, bertindak sebagai penyedia fitur utama (*shared feature*) untuk diproses secara simultan oleh dua jalur arsitektur *deep learning independen*.
5) IndoBERT, pada tahap ini jalur pemrosesan pertama dari *Shared Feature Extractor*, di mana matriks vektor dari IndoBERT diproses menggunakan arsitektur IndoBERT untuk menangkap hubungan sekuensial jangka panjang (*long-range dependencies*) dan ketergantungan antar kata dalam teks ulasan.
6) CNN, tahap ini merupakan jalur pemrosesan kedua dari *Shared Feature Extractor*, di mana matriks vektor dari IndoBERT diproses menggunakan arsitektur *Convolutional Neural Network* (CNN) 1-Dimensi (1D) untuk menangkap fitur lokal, pola hubungan kata yang berdekatan (n-gram), serta kata kunci spesifik (*key phrases*) pada ulasan produk.
7) *Feature Fusion,* tahap ini merupakan proses integrasi atau penyatuan dua karakteristik fitur yang berbeda, yaitu fitur konteks global yang dihasilkan oleh arsitektur transformer dan fitur konteks lokal yang diekstrak oleh arsitektur CNN. Proses penggabungan ini dilakukan untuk menghasilkan representasi vektor baru yang jauh lebih kaya dan komprehensif.
8) Model *Training,* tahap ini merupakan proses melatih arsitektur model hibrida yang telah dirancang menggunakan data latih (*training* data) agar model dapat mengenali pola, mengekstrak bobot fitur yang optimal, dan memprediksi target label sentimen serta emosi secara akurat.
9) Model *Evaluation,* tahap ini merupakan langkah visualisasi dan pemetaan performa hasil pengujian model secara mendalam. Apabila hasil evaluasi menunjukkan bahwa performa model belum optimal, maka dilakukan proses *hyperparameter tuning* seperti penyesuaian jumlah *epoch, learning rate, dropout*, dan *class weight*, kemudian model dilatih kembali (*retraining*) hingga diperoleh performa yang memenuhi kriteria. Sebaliknya, apabila hasil evaluasi telah optimal, proses dilanjutkan ke tahap analisis akhir berdasarkan *confusion matrix* dan metrik evaluasi.
10) *Confusion Matrix* menyajikan tabel *kontingensi* yang membandingkan secara detail antara jumlah data aktual (*ground truth*) dari PRDECT-ID Dataset dengan hasil prediksi yang dikeluarkan oleh

arsitektur model hibrida. pada persamaan 4 kuadran utama, yaitu True Positif (TP), *True* Negatif (TN), False Positif (FP), dan False Negatif (FN). Berikut rumus dari 4 kuadaran :

a. *Accuracy* : mengukur seberapa banyak prediksi (baik positif maupun negatif) yang benar dari keseluruhan data uji. sebagaimana dirumuskan dalam persamaan (1) :

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

b. *Precision* : mengukur ketepatan model dalam memprediksi kelas positif. Artinya, dari semua ulasan yang diprediksi masuk ke kelas tertentu oleh model, berapa banyak yang sebenarnya benar-benar sesuai dengan data asli. sebagaimana dirumuskan dalam persamaan (2) :

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

c. *Recall* : mengukur kemampuan model dalam menemukan kembali seluruh sampel data dari kelas aktual yang ada. Artinya, dari semua ulasan yang seharusnya masuk ke suatu kelas emosi, berapa banyak yang berhasil ditemukan oleh model. sebagaimana dirumuskan dalam persamaan (3) :

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

d. F1 *Score* : rata-rata harmonik antara *Precision* dan *Recall*. Metrik ini memberikan keseimbangan performa yang akurat ketika jumlah persebaran data antar-kelas tidak seimbang (*imbalanced* dataset). sebagaimana dirumuskan dalam persamaan (4) :

$$F1\ Score = 2\ x\ \frac{Precision\ x\ Recall}{Precision + Recall} \quad (4)$$

11) Metrik Evaluasi, tahap ini merupakan bagian akhir dari rangkaian analisis, di mana efektivitas, keandalan, dan performa kuantitatif dari arsitektur model hibrida dihitung secara matematis. Parameter numerik yang telah diperoleh dari tahap *confusion matrix* diolah menggunakan beberapa metrik evaluasi standar untuk klasifikasi teks *multi*-kelas dan *multi*-label sebagai berikut :

a) Kurva ROC dan AUC :
   1. *Receiver Operating Characteristic* (ROC) *Curve* digunakan untuk memvisualisasikan performa klasifikasi pada berbagai tingkatan nilai ambang batas (*threshold*).
   2. *Area Under the Curve* (AUC) menghitung luasan daerah di bawah kurva ROC untuk mengukur kemampuan model dalam membedakan antara satu kelas emosi dengan kelas emosi lainnya secara keseluruhan (skor mendekati 1.00 menandakan kemampuan diskriminasi kelas yang sempurna).

b) *Hamming Loss* :
   1. Mengukur fraksi atau *persentase* label yang salah diprediksi oleh model terhadap total label keseluruhan.
   2. Semakin kecil nilai *Hamming Loss* (mendekati 0), maka semakin tinggi tingkat akurasi dan konsistensi model dalam memprediksi kombinasi label sentimen dan emosi secara bersamaan. sebagaimana dirumuskan dalam persamaan (5) :

$$Hamming\ Loss = \frac{1}{N.L} \sum_{i=1}^{N} \sum_{j=1}^{L} I(y_{ij} \neq y_{ij}) \quad (5)$$

dimana :

N : Total jumlah sampel data uji yang digunakan dalam proses evaluasi.

L : Total jumlah kategori label yang diprediksi (pada penelitian ini L = 5 untuk kategori emosi).

Yij : Nilai *biner* aktual (*ground truth*) pada data ke-i untuk label ke-j.

Yij : Nilai *biner* hasil prediksi model hibrida pada data ke-i untuk label ke-j.
I : Fungsi indikator (*indicator function*) yang akan bernilai 1 jika hasil prediksi salah, dan bernilai 0 jika hasil prediksi benar.

Pemilihan model *hybrid multi-task* IndoBERT-CNN pada penelitian ini didasarkan pada karakteristik data ulasan *e-commerce* berbahasa Indonesia yang memiliki variasi kosakata, bahasa tidak baku, serta mengandung informasi sentimen dan emosi secara bersamaan. IndoBERT dipilih karena mampu menghasilkan representasi kontekstual yang sesuai dengan karakteristik bahasa Indonesia sehingga dapat memahami hubungan semantik antar kata secara lebih baik. Selanjutnya, CNN digunakan untuk mengekstraksi fitur-fitur lokal berupa pola frasa dan kata kunci yang sering muncul pada ulasan pelanggan. Pendekatan *multi-task learning* diterapkan agar model dapat mempelajari klasifikasi sentimen dan emosi secara simultan dalam satu arsitektur, sehingga pembelajaran menjadi lebih efisien dan representasi fitur yang dihasilkan lebih komprehensif dibandingkan pendekatan *single-task*.

Deep learning merupakan metode pembelajaran berbasis jaringan saraf tiruan yang dirancang untuk menangkap pola-pola abstrak melalui pembelajaran representasi data secara *hierarkis*. Dalam memanfaatkan rangkaian modul sederhana, *deep learning* secara progresif mentransformasikan input mentah menjadi representasi data yang lebih kompleks pada setiap tingkatannya [6].

Natural language processing (NLP) merupakan disiplin ilmu kecerdasan buatan yang memfasilitasi kemampuan komputer dalam menginterpretasikan dan memahami bahasa alami manusia, dengan tujuan menciptakan interaksi yang lebih efektif antara mesin dan pengguna. Pemanfaatan metode NLP dalam pengembangan *chatbot* semakin populer karena fleksibilitasnya dalam memberikan respons yang lebih akurat serta kemudahan dalam integrasi pengolahan datanya [7].

Dalam pengolahan data teks, *preprocessing* diterapkan untuk mentransformasikan dokumen tidak terstruktur menjadi data yang siap diolah oleh model. Proses ini melibatkan pemecahan teks menjadi unit kata individu, yang kemudian disempurnakan melalui serangkaian prosedur seperti *case folding* untuk penyeragaman karakter, *stemming* untuk reduksi kata dasar, serta *stopword removal* untuk menyaring kata yang tidak relevan [8].

IndoBERT adalah varian BERT yang dapat dioptimalkan untuk bahasa Indonesia. Kinerja model ini didukung oleh proses *pre-training* pada korpus data yang luas dan bersih, terdiri atas 220 juta kata yang dikumpulkan dari Wikipedia bahasa Indonesia, Indonesia *Web Corpus*, serta artikel berita dari Kompas, Tempo, dan Liputan 6 [9].

Embedding merepresentasikan teks dalam bentuk vektor numerik sehingga kesamaan semantik dapat diukur tanpa bergantung pada pencocokan kata kunci. Perkembangan *embedding* modern didukung oleh *benchmark* seperti *Massive Text Embedding Benchmark* (MTEB) yang memungkinkan perbandingan kinerja berbagai model pada tugas *retrieval* lintas domain [10]. Tahap *tokenizing* bertujuan untuk memecah teks ulasan menjadi unit-unit token yang dapat diproses oleh model. tokenisasi dilakukan berbasis kata yang selanjutnya diubah menjadi indeks atau vektor *embedding*. Sementara itu, pada BERT digunakan *subword tokenization* (*WordPiece*) yang memungkinkan model menangani kata tidak baku dan variasi bahasa secara lebih efektif. Token ini dapat berupa kata atau frasa. Tujuan utama tokenizing adalah untuk membuat teks lebih mudah diolah, dianalisis, dan dipahami oleh sistem atau program komputer. Dengan memecah teks menjadi token, sistem dapat lebih mudah mengidentifikasi pola, menentukan makna, dan melakukan analisis sentimen terhadap teks tersebut [11].

CNN dikenal luas sebagai arsitektur *deep learning* tangguh dalam memproses data berbasis gambar dan teks. Mekanisme kerja CNN melibatkan ekstraksi fitur melalui lapisan *konvolusi, pooling*, dan *fully connected*. Khusus untuk data teks, CNN unggul dalam mengenali pola-pola lokal dan sekuensial [12]. Tahapan *pooling* dibedakan menjadi *max pooling* untuk mengambil nilai maksimum dan *average pooling* untuk nilai rerata [13]. Dalam lapisan *konvolusi*, model mengidentifikasi pola input menggunakan kernel yang bobotnya terus diperbarui selama proses pelatihan agar representasi fitur pada data menjadi optimal [14].

Strategi *multi-task learning* memungkinkan model untuk memproses dua tugas klasifikasi secara simultan dengan meminimalkan gabungan dua *cross-entropy loss* yang spesifik untuk label kekerasan dan sentimen. Model dievaluasi secara berkala pada setiap *epoch* menggunakan parameter akurasi, *presisi, recall*, dan F1-*score*. Pendekatan ini tidak hanya mengoptimalkan efisiensi pelatihan, tetapi juga secara signifikan memperbaiki kemampuan model dalam mengekstraksi pola dari narasi laporan yang beragam dan kompleks [15].

3. HASIL DAN PEMBAHASAN

3.1. Hasil Evaluasi Model

Kinerja model analisis sentimen dievaluasi melalui serangkaian metrik yang relevan untuk menentukan sejauh mana akurasi dan konsistensi hasil klasifikasi yang dihasilkan.

1) Metrik Evaluasi

```
========================================
A. CLASSIFICATION REPORT (PRECISION, RECALL, F1)
========================================
SENTIMEN:
              precision    recall  f1-score   support

           0       0.98      0.99      0.98       423
           1       0.99      0.97      0.98       387

    accuracy                           0.98       810
   macro avg       0.98      0.98      0.98       810
weighted avg       0.98      0.98      0.98       810

EMOSI:
              precision    recall  f1-score   support

           0       0.64      0.44      0.52       114
           1       0.44      0.49      0.46       128
           2       0.88      0.72      0.79       269
           3       0.58      0.81      0.67       118
           4       0.62      0.70      0.66       181

    accuracy                           0.65       810
   macro avg       0.63      0.63      0.62       810
weighted avg       0.67      0.65      0.65       810
```

Gambar 3. Output Metrik Evaluasi *Classification Report*.

Pada Gambar 3 ditunjukkan hasil model menunjukkan kemampuan klasifikasi yang sangat baik dengan nilai akurasi sebesar 98%. Konsistensi nilai *precision, recall*, dan f1-*score* yang berada di angka 0,98 pada kelas 0 (negatif) dan kelas 1 (positif) menunjukkan bahwa model telah berhasil mengenali pola polaritas sentimen secara seimbang dan sangat akurat. Hal ini membuktikan bahwa fitur teks yang diekstraksi mampu membedakan kedua polaritas dengan sangat jelas.

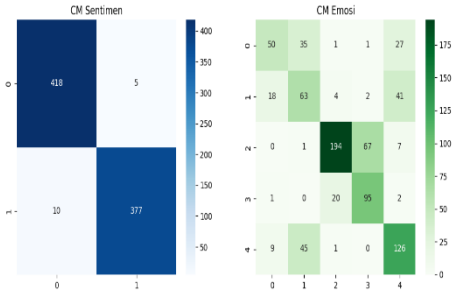
TABEL VII. PERBANDINGAN HASIL AKURASI MODEL PADA SETIAP *EPOCH*

Eksperimen	*Epoch*	Akurasi Sentimen	Akurasi Emosi	*Loss*
Epoch 1	1	95,18%	60,99%	0,9950
Epoch 2	2	99,13%	74,33%	0,6286
Epoch 3	3	99,89%	82,14%	0,4555
Epoch 4	4	99,97%	89,26%	0,2890
Epoch 5	5	99,95%	94,13%	0,1703

Berdasarkan Tabel 7, performa model *hybrid* IndoBERT–CNN mengalami peningkatan yang konsisten pada setiap *epoch* selama proses pelatihan. Pada *epoch* pertama, model memperoleh akurasi sentimen sebesar 95,18% dan akurasi emosi sebesar 60,99%, dengan nilai *loss* sebesar 0,9950. Hasil tersebut menunjukkan bahwa model telah mampu mempelajari pola dasar dari data, namun kemampuan dalam mengklasifikasikan emosi masih belum optimal.

Pada *epoch* kedua hingga *epoch* keempat, performa model terus mengalami peningkatan. Akurasi sentimen meningkat menjadi 99,13%, 99,89%, dan 99,97%, sedangkan akurasi emosi meningkat menjadi 74,33%, 82,14%, dan 89,26%. Di sisi lain, nilai *loss* terus menurun dari 0,6286 menjadi 0,2890. Peningkatan akurasi yang diikuti dengan penurunan nilai *loss* menunjukkan bahwa proses pelatihan berjalan secara efektif dan model semakin mampu mengenali pola pada data ulasan produk.

2) *Confusion Matrix*



Gambar 4. Visualisasi *Confusion Matrix*.

Berdasarkan hasil pengujian model hibrida IndoBERT-CNN menggunakan data validasi yang disajikan pada Gambar 4, ditunjukkan hasil menunjukkan kemampuan model dalam mengklasifikasikan data berdasarkan label sebenarnya (*actual* label) dan label hasil prediksi (*predicted* label). Pada klasifikasi sentimen, terlihat dimana sebagian besar data berada pada diagonal utama, yang menunjukkan bahwa model mampu mengklasifikasikan sentimen dengan tingkat ketepatan yang sangat tinggi. Sebanyak 418 data sentimen negatif berhasil diprediksi dengan benar dan 377 data sentimen positif juga berhasil diklasifikasikan secara tepat. Kesalahan klasifikasi relatif kecil, yaitu hanya 5 data negatif yang diprediksi sebagai positif dan 10 data positif yang diprediksi sebagai negatif. Pola ini menunjukkan bahwa model memiliki kemampuan diskriminasi yang sangat baik dalam membedakan dua kelas sentimen sehingga menghasilkan nilai akurasi yang tinggi.

Selanjutnya, pada matriks evaluasi Emosi Metrik *Hamming Loss* & ROC AUC

```
======================================
B. CORE METRICS & HAMMING LOSS
======================================
Akurasi Sentimen: 98.15%
Akurasi Emosi: 65.19%
Multi-task Hamming Loss: 0.1833

======================================
C. ROC AUC SCORE
======================================
ROC AUC Sentimen: 0.9980
ROC AUC Emosi: 0.9102
```

Gambar 5. Metrik *Hamming Loss* & ROC AUC.

Berdasarkan Gambar 5, hasil model mencapai akurasi sebesar 98,15% pada klasifikasi sentimen dan 65,19% pada klasifikasi emosi. Perbedaan capaian akurasi ini kembali menegaskan bahwa pola sentimen bersifat lebih polar dan mudah dikenali, sementara pola emosi bersifat lebih kompleks dan *multiclass*. Nilai *multi-task hamming loss* sebesar 0,1833 menunjukkan tingkat kesalahan klasifikasi label yang relatif rendah dalam sistem *multi-task*. Nilai ini mengindikasikan bahwa secara keseluruhan, model cukup stabil dalam memprediksi dua label (sentimen dan emosi) secara simultan, meskipun terdapat ruang untuk peningkatan pada prediktor emosi.

ROC AUC Sentimen sebesar 0,9980: Menunjukkan performa yang hampir sempurna (*excellent*) dalam memisahkan polaritas sentimen positif dan negatif. ROC AUC Emosi sebesar 0,9102: Menunjukkan kemampuan yang sangat baik (*very good*) bagi model untuk membedakan berbagai kelas emosi. Meskipun akurasi emosi berada di angka 65,19%, nilai ROC AUC yang tinggi (0,9102) membuktikan bahwa model sebenarnya memiliki kapasitas diskriminasi yang kuat, namun performa akurasinya terkendala oleh ambang batas (*threshold*) klasifikasi yang belum optimal atau ketimpangan distribusi data emosi pada dataset.

4) Hasil *Training*

```
Loading weights: 100%|██████████| 199/199 [00:00<00:00, 9619.08it/s]
EPOCH 1 | Loss: 0.9950 | Sent Acc: 95.18% | Emot Acc: 60.99%
EPOCH 2 | Loss: 0.6286 | Sent Acc: 99.13% | Emot Acc: 74.33%
EPOCH 3 | Loss: 0.4555 | Sent Acc: 99.89% | Emot Acc: 82.14%
EPOCH 4 | Loss: 0.2890 | Sent Acc: 99.97% | Emot Acc: 89.26%
EPOCH 5 | Loss: 0.1703 | Sent Acc: 99.95% | Emot Acc: 94.13%
[INFO] Training selesai. File 'best_model.pth' dan 'training_results.png' tersimpan.
```

Gambar 6. Hasil *Training*

Berdasarkan Gambar 6 ditunjukkan hasil pencapaian nilai *loss* sebesar 0,1703 menunjukkan bahwa model telah berhasil mengoptimalkan bobot parameter secara sangat efektif dalam meminimalkan fungsi *error*. Tingkat akurasi sentimen sebesar 99,95% merupakan hasil yang sangat memuaskan, di mana model hampir mencapai titik klasifikasi sempurna dalam membedakan polaritas sentimen pada ulasan produk.

Sementara itu, akurasi emosi sebesar 94,13% menunjukkan kemampuan model yang sangat tangguh dalam menangani kompleksitas klasifikasi emosi yang bersifat *multi-class* (mencakup *anger, fear, happy, love, sadness*). Konsistensi performa ini memberikan bukti empiris bahwa arsitektur *hybrid* IndoBERT-CNN yang diusulkan mampu bekerja secara sinergis dalam menangkap fitur semantik dan pola emosional teks dengan tingkat ketepatan yang sangat tinggi pada dataset penelitian ini.

Hasil penelitian yang diperoleh menunjukkan bahwa pendekatan *hybrid multi-task learning* berbasis IndoBERT dan *convolutional neural network* (CNN) mampu melakukan klasifikasi sentimen dan emosi secara simultan pada ulasan produk *e-commerce* berbahasa Indonesia. Model yang diusulkan memperoleh akurasi klasifikasi sentimen sebesar 98,15% dan akurasi klasifikasi emosi sebesar 65,19%. Untuk mengetahui posisi penelitian ini dalam perkembangan penelitian sebelumnya, dilakukan perbandingan dengan beberapa penelitian terdahulu yang

menggunakan pendekatan serupa pada tugas analisis sentimen maupun klasifikasi emosi.

penelitian ini memiliki kesamaan dengan penelitian Latifa maysara (2026) yang memanfaatkan IndoBERT untuk analisis sentimen. Penelitian tersebut memperoleh akurasi sebesar 98,52%, sedangkan penelitian ini memperoleh akurasi sentimen sebesar 98,15%. Perbedaan tersebut dapat dipengaruhi oleh arsitektur model, skenario eksperimen, serta tujuan penelitian. Penelitian Latifa Maysara berfokus pada analisis sentimen (*single-task*), sedangkan penelitian ini menerapkan *multi-task learning* sehingga model harus mempelajari dua tugas secara bersamaan, yaitu klasifikasi sentimen dan klasifikasi emosi.

Pada aspek klasifikasi emosi, penelitian ini memiliki keterkaitan dengan penelitian Chowanda et al. (2024) yang mengombinasikan BERT, CNN, dan BiLSTM untuk pengenalan emosi pada ulasan produk Indonesia. Penelitian tersebut memperoleh akurasi sebesar 69,26%, sedangkan penelitian ini memperoleh akurasi 65,19%. Selain itu, penelitian Raharko dan Yamasari (2024) yang menggunakan metode *support vector machine* (SVM) memperoleh akurasi 66,30%, sementara penelitian Ariyanto et al. (2024) memperoleh akurasi sekitar 62% menggunakan pendekatan *contextual embeddings*. Hasil tersebut menunjukkan bahwa performa klasifikasi emosi pada penelitian ini masih berada pada kisaran hasil penelitian terdahulu, meskipun menggunakan pendekatan yang berbeda.

Perbedaan utama penelitian ini dibandingkan penelitian-penelitian sebelumnya terletak pada penggunaan arsitektur hybrid IndoBERT–CNN dalam skema *multi-task learning*. Sebagian besar penelitian terdahulu hanya berfokus pada satu tugas klasifikasi, yaitu analisis sentimen atau klasifikasi emosi secara terpisah. Sebaliknya, penelitian ini mengintegrasikan kedua tugas tersebut ke dalam satu model dengan *shared feature extractor* berbasis IndoBERT dan ekstraksi fitur lokal menggunakan CNN melalui mekanisme *feature fusion*. Pendekatan ini memungkinkan model memanfaatkan representasi fitur yang sama untuk dua tugas klasifikasi secara simultan, sehingga menghasilkan akurasi sentimen yang tinggi serta kemampuan klasifikasi emosi yang kompetitif.

4. KESIMPULAN DAN SARAN

4.1 Kesimpulan

Berdasarkan hasil penelitian, implementasi, serta pengujian yang telah dilakukan mengenai penerapan arsitektur *multi-task hybrid* IndoBERT-CNN untuk klasifikasi sentimen dan emosi pada ulasan produk *e-commerce* berbahasa Indonesia menggunakan PRDECT-ID Dataset, maka diperoleh beberapa kesimpulan sebagai berikut :

1. Rangkaian enam tahapan *text preprocessing* yang dirancang secara terstruktur meliputi *Case Folding, Cleaning, Trim Space, Tokenizing, Normalization*, dan *Stopword Removal* terbukti secara empiris mampu menyaring data teks ulasan konsumen dari gangguan (*noise*). Proses ini berhasil meningkatkan keseragaman fitur bahasa sebelum ditransformasikan menjadi representasi tensor numerik oleh model bahasa.
2. Integrasi arsitektur hibrida secara *multi-task learning* berhasil mengekstraksi fitur semantik secara optimal. Lapisan IndoBERT bertugas menangkap hubungan konteks semantik global jangka panjang (*global context relations*), sementara lapisan *Convolutional Neural Network* (CNN) bertindak sebagai pengekstrak pola frasa unik lokal (*local feature extractor*) yang mencerminkan karakteristik ulasan produk *e-commerce*.
3. Berdasarkan hasil pengujian komparatif dari lima skenario model yang diujikan, *Epoch* ke 5 yang menggunakan arsitektur hibrida berbasis IndoBERT (*indobert-base-p2*) dengan pengaturan *hyperparameter learning rate* terkecil yaitu $2\,x\,10^{-5}$ keluar sebagai model dengan performa terbaik (*superior* model). Model ini sukses mendominasi akurasi klasifikasi tugas ganda secara simultan dengan perolehan akhir pada *epoch* kelima sebesar 99,95% untuk klasifikasi Sentimen dan sebesar 94,13% untuk deteksi kelas Emosi.

4.2 Saran

Untuk pengembangan penelitian klasifikasi teks *multi-task* selanjutnya, disarankan beberapa hal berikut :

1. Penanganan Ketidakseimbangan Data (*Imbalanced Data*) pada target klasifikasi, Untuk penelitian selanjutnya, disarankan untuk menerapkan teknik penyeimbangan data sintetis seperti SMOTE (*Synthetic Minority Over-sampling Technique*) atau pendekatan oversampling dan undersampling yang adaptif, khususnya pada tugas klasifikasi emosi *multi-class*. Hal ini krusial karena distribusi data asli lapangan pada dataset PRDECT-ID sangat timpang antara kelas mayoritas (*Anger* dan *Fear*) dengan kelas minoritas (*Happy, Love, Sadness*), yang secara matematis menjadi batasan utama (*limitation*) rendahnya nilai rata-rata Macro ROC AUC pada model hibrida saat ini.
2. Optimasi *Hyperparameter*: Menerapkan metode pencarian otomatis seperti *Grid Search* atau *Bayesian Optimization* untuk menemukan kombinasi *batch size* dan jumlah *layer* CNN yang lebih bervariasi.
3. Penelitian selanjutnya disarankan menggunakan dataset yang lebih beragam dan seimbang agar kemampuan generalisasi model dapat ditingkatkan. Selain itu, pengembangan arsitektur model melalui optimasi hiperparameter, penambahan mekanisme *attention*, atau eksplorasi model *deep learning* lainnya dapat dilakukan untuk meningkatkan performa klasifikasi emosi atau mengeksplorasi model berbasis *large language* model (LLM) untuk meningkatkan performa klasifikasi, khususnya pada tugas klasifikasi emosi yang masih memiliki akurasi lebih rendah dibandingkan klasifikasi sentimen.

Daftar Pustaka

[1] Tukino, "Penerapan Algoritma Convolutional Neural Network Untuk Klasifikasi Sentimen Pada Layanan e-Commerce," Jan. 2025. [Online]. Available: http://journal.aptikomkepri.org/index.php/JDDAT44JURNALDESAINDANANALISISTEKNOLOGI

[2] R. Aulia, A. Ruslani, and M. D. Fauzan, "Pengembangan Model Transformer untuk Analisis Sentimen dan Ringkasan Ulasan E-commerce Berbahasa Indonesia," vol. 4, no. 1, 2025.

[3] A. Maskhuri and T. Hardiani, "Analisis Sentimen Pengguna pada Aplikasi Tokopedia Menggunakan Algoritma Convolutional Neural Network," *Technology and Science (BITS)*, vol. 6, no. 4, 2025, doi: 10.47065/bits.v6i4.6923.

[4] A. Sabil Deva Pratama and N. Rijati, "Pengenalan Emosi Ulasan Pelanggan E-Commerce Menggunakan Deep Learning Berbasis Transformer Emotion Recognition of E-Commerce Customer Reviews Using Transformer-Based Deep Learning."

[5] F. A. Khatami, B. Irawan, S. Si, and C. Setianingsih, "ANALISIS SENTIMEN TERHADAP REVIEW APLIKASI LAYANAN E-COMMERCE MENGGUNAKAN METODE CONVOLUTIONAL NEURAL NETWORK SENTIMENT ANALYSIS OF E-COMMERCE APPLICATION REVIEWS USING THE CONVOLUTIONAL NEURAL NETWORK METHOD."

[6] S. Prasetyo and T. Dewayanto, "PENERAPAN MACHINE LEARNING, DEEP LEARNING, DAN DATA MINING DALAM DETEKSI KECURANGAN LAPORAN KEUANGAN-A SYSTEMATIC LITERATURE REVIEW," *DIPONEGORO JOURNAL OF ACCOUNTING*, vol. 13, no. 3, pp. 1–12, 2024, [Online]. Available: http://ejournal-s1.undip.ac.id/index.php/accounting

[7] D. Apriliani, S. Febbi Handayani, I. Triadi Saputra, T. Informatika, and P. Harapan Bersama, "Implementasi Natural Language Processing (NLP) Dalam Pengembangan Aplikasi Chatbot Pada SMK YPE Nusantara Slawi Implementation of Natural Language Processing (NLP) in the Development of a Chatbot Application at SMK YPE Nusantara Slawi," 2023.

[8] A. Y. Timur, "SISTEM REKOMENDASI LAGU INDONESIA MENGGUNAKAN METODE CONTENT-BASED FILTERING DAN COSINE SIMILARITY," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 1, Jan. 2025, doi: 10.23960/jitet.v13i1.5949.

[9] N. Herlinawati, I. Ariyati, and A. Yuli Asih, "ANALISIS SENTIMEN ZOOM CLOUD MEETINGS DI PLAY STORE MENGGUNAKAN METODE INDOBERT," Feb. 2026.

[10] R. Adriansyah, R. Setya Perdana, and P. P. Adikara, "ANALISIS PENGARUH MODEL

EMBEDDING PADA KUALITAS RETRIEVAL DOKUMEN: STUDI KASUS PERATURAN REKTOR UNIVERSITAS BRAWIJAYA," 2026. [Online]. Available: http://j-ptiik.ub.ac.id

[11] T. Ramadhan, "Performance Comparison of IndoBERT and Bi-LSTM Models for Sentiment Analysis of Shopee App Users in Indonesia," 2026. [Online]. Available: http://jurnal.polibatam.ac.id/index.php/JAIC

[12] Y. Prastya, "HYBRID CNN-SVM UNTUK ANALISIS SENTIMEN KLUB TIM NASIONAL," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 3S1, Oct. 2025, doi: 10.23960/jitet.v13i3S1.7764.

[13] Y. Pratama, E. Rasywir, F. Fachruddin, D. Kisbianty, and B. Irawan, "Eksperimen Layer Pooling menggunakan Standar Deviasi untuk Klasifikasi Dataset Citra Wajah dengan Metode CNN," *Building of Informatics, Technology and Science (BITS)*, vol. 5, no. 1, Jun. 2023, doi: 10.47065/bits.v5i1.3604.

[14] G. A. Pratama, E. Y. Puspaningrum, and H. Maulana, "CONVOLUTIONAL NEURAL NETWORK DAN FASTER REGION CONVOLUTIONAL NEURAL NETWORK UNTUK KLASIFIKASI KUALITAS BIJI KOPI ARABIKA," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 3, Aug. 2024, doi: 10.23960/jitet.v12i3.4887.

[15] M. Jundi, R. Rahmaddeni, P. Utami, S. Perdana Arifin, and L. Sinaga, "Klasifikasi Pengaduan Kekerasan Berbasis Bidirectional Encoder Representations from Transformers (BERT) pada Kementrian Pemberdayaan Perempuan dan Perlindungan Anak Kementrian (PPPA)," *Jurnal Pustaka AI (Pusat Akses Kajian Teknologi Artificial Intelligence)*, vol. 5, no. 2, pp. 258–264, Aug. 2025, doi: 10.55382/jurnalpustakaai.v5i2.1109.