

## **PENERAPAN METODE CNN UNTUK ANALISIS SENTIMEN PENGGUNA APLIKASI DOLA**

### **IMPLEMENTATION OF THE CNN METHOD FOR SENTIMENT ANALYSIS OF DOLA APPLICATION USERS**

**Zendra Utama<sup>1</sup>, Pahrizal<sup>2</sup>**

Teknik Informatika, Universitas Muhammadiyah Bengkulu<sup>1,2</sup>  
[zendrautama36@gmail.com](mailto:zendrautama36@gmail.com)<sup>1</sup>

#### **ABSTRACT**

*This study aims to analyze sentiment in Dola application user reviews on the Google Play Store using the Convolutional Neural Network (CNN) method. The background of this research is the increasing number of user reviews that can be utilized to determine user satisfaction with the application. The research data were collected through a crawling process using Python, resulting in 3000 user reviews. The research stages included data pre-processing such as cleaning, case folding, tokenization, stopword removal, slangword normalization, and stemming, followed by sentiment labeling into positive, neutral, and negative categories based on user ratings. The dataset was divided into 70% training data and 30% testing data. The results showed that the CNN model was able to classify sentiments with an accuracy rate of 82%. Based on the evaluation results, the CNN method is considered effective in identifying sentiment patterns in Dola application user reviews and can be used as evaluation material for application development.*

**Keywords:** *Sentiment Analysis, Convolutional Neural Network, CNN, Google Play Store, Dola Application*

#### **ABSTRAK**

Penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna aplikasi Dola pada Google Play Store menggunakan metode Convolutional Neural Network (CNN). Latar belakang penelitian ini adalah meningkatnya jumlah ulasan pengguna yang dapat dimanfaatkan untuk mengetahui tingkat kepuasan terhadap aplikasi. Data penelitian diperoleh melalui proses crawling menggunakan Python dengan total 3000 ulasan pengguna. Tahapan penelitian meliputi pre-processing data berupa cleaning, case folding, tokenization, stopword removal, slangword, dan stemming, kemudian dilanjutkan dengan pelabelan sentimen positif, netral, dan negatif berdasarkan rating pengguna. Dataset dibagi menjadi 70% data latih dan 30% data uji. Hasil penelitian menunjukkan bahwa model CNN mampu mengklasifikasikan sentimen dengan tingkat akurasi sebesar 82%. Berdasarkan hasil evaluasi, metode CNN dinilai efektif dalam mengenali pola sentimen pada ulasan pengguna aplikasi Dola dan dapat digunakan sebagai bahan evaluasi pengembangan aplikasi.

**Kata Kunci:** Analisis Sentimen, Convolutional Neural Network, CNN, Google Play Store,

#### **PENDAHULUAN**

Perkembangan teknologi digital dan meningkatnya penggunaan aplikasi berbasis mobile telah mendorong munculnya berbagai platform layanan yang memudahkan aktivitas pengguna sehari-hari. Salah satu aplikasi yang berkembang adalah aplikasi Dola, yang digunakan oleh banyak pengguna untuk berbagai kebutuhan komunikasi dan layanan digital (Qudsi *et al.*, 2021). Seiring meningkatnya jumlah pengguna, ulasan dan komentar yang diberikan di berbagai platform menjadi sumber data yang sangat penting untuk memahami kepuasan dan

pengalaman pengguna terhadap aplikasi tersebut (Aprianti *et al.*, 2025).

Analisis sentimen merupakan salah satu pendekatan dalam bidang text mining yang digunakan untuk mengidentifikasi opini atau emosi yang terkandung dalam sebuah teks, apakah bersifat positif, negatif, atau netral (Sitio *et al.*, 2024). Dalam konteks pengembangan aplikasi, analisis sentimen dapat membantu pengembang memahami persepsi pengguna sehingga dapat meningkatkan kualitas layanan dan fitur yang disediakan (Maulana *et al.*, 2025).

Beberapa metode telah digunakan dalam analisis sentimen, mulai dari

pendekatan berbasis machine learning tradisional hingga metode deep learning (Ayuningtiyas *et al.*, 2025). Salah satu metode deep learning yang efektif dalam pengolahan data teks adalah *Convolutional Neural Network* (CNN). CNN dikenal mampu mengekstraksi fitur penting dari data teks secara otomatis melalui lapisan konvolusi, sehingga dapat meningkatkan akurasi dalam klasifikasi sentiment (Dwiyansaputra *et al.*, 2025).

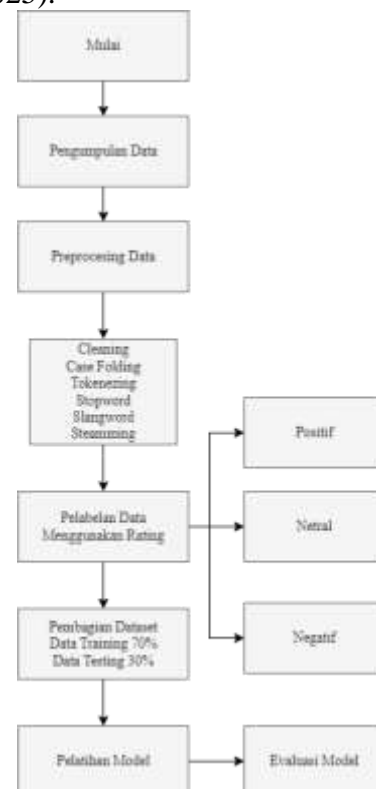
Hasil penelitian terdahulu menunjukkan bahwa metode *Convolutional Neural Network* (CNN) memiliki performa yang baik dalam analisis sentimen. Penelitian yang dilakukan oleh Sitio *et al.*, (2024). Menunjukkan bahwa CNN mampu mengklasifikasikan sentimen pada data Twitter dengan hasil yang baik dalam mengidentifikasi opini pengguna. Selain itu, penelitian Maulana *et al.*, (2025). Menunjukkan bahwa penggunaan CNN pada analisis sentimen pengguna aplikasi Indodax mampu memberikan hasil klasifikasi yang efektif dalam mengenali sentimen positif dan negatif. Penelitian Saputra *et al.*, (2024). Yang membandingkan metode RNN dan CNN dalam klasifikasi sentimen ulasan aplikasi di Google Play Store juga menunjukkan bahwa CNN memiliki performa yang lebih stabil dan akurat. Berdasarkan hasil penelitian terdahulu tersebut, metode CNN dinilai sesuai untuk digunakan dalam penelitian ini karena mampu mengekstraksi fitur teks secara otomatis dan menghasilkan klasifikasi sentimen yang efektif.

Penelitian ini bertujuan untuk menerapkan metode CNN dalam menganalisis sentimen pengguna aplikasi Dola berdasarkan ulasan yang tersedia. Dengan menggunakan pendekatan ini, diharapkan dapat diperoleh hasil klasifikasi sentimen yang akurat serta memberikan gambaran mengenai tingkat kepuasan pengguna terhadap aplikasi tersebut. Hasil penelitian ini diharapkan dapat menjadi masukan bagi pengembang aplikasi dalam melakukan evaluasi dan pengembangan

fitur di masa mendatang (Yuniarossy *et al.*, 2024).

## METODE

Penelitian ini menerapkan metode *Convolutional Neural Network* (CNN) untuk menilai pandangan pengguna tentang aplikasi Dola yang diambil dari Google Play Store (Vector & Algorithm, 2025). Pendekatan ini diharapkan dapat mengelompokkan pandangan pengguna ke dalam kategori sentimen positif, netral, dan negatif dengan tingkat akurasi yang lebih tinggi (Rinardi *et al.*, 2024). Dengan menganalisis berbagai ulasan dan komentar pengguna terhadap aplikasi Dola, sistem ini bertujuan untuk membantu pengembang aplikasi memahami respons serta tingkat kepuasan pengguna terhadap layanan dan fitur yang tersedia. Selain itu, hasil analisis sentimen ini juga dapat digunakan sebagai bahan evaluasi dan pendukung pengambilan keputusan dalam pengembangan serta peningkatan kualitas aplikasi Dola di masa mendatang (Kamal *et al.*, 2025).



**Gambar 1. Alur Penelitian**

### **Crawling Data**

Data penelitian dikumpulkan menggunakan bahasa pemrograman Python. Pengambilan data dilaksanakan dengan metode crawling yang memanfaatkan library harvest dan dijalankan di platform *Google Colaboratory* yang bisa diakses secara online. Hasil proses crawling menghasilkan sebanyak 3000 data ulasan.

### **Pre-Processing Data**

Salah satu tahap krusial dalam penelitian ini adalah *pre-processing* data, yang dilakukan untuk mempersiapkan data sebelum diolah oleh model *Convolutional Neural Network* (CNN). Fokus dari tahap ini adalah untuk memperbaiki kualitas data dengan menghapus gangguan dan elemen yang tidak diperlukan agar data menjadi lebih bersih serta terstruktur untuk dianalisis. Langkah-langkah *pre-processing* meliputi *text cleaning*, *case folding*, *tokenization*, *stopword removal*, dan *stemming*. Setelah seluruh tahapan tersebut dilaksanakan, data teks selanjutnya diubah menjadi format vektor numerik agar bisa digunakan sebagai input dalam model CNN untuk menganalisis sentimen (Martha *et al.*, 2026).

### **Pelabelan Data**

Pelabelan data merupakan Langkah krusial dalam analisis sentimen karena berperan dalam menentukan jenis emosi dari setiap komentar pengguna aplikasi Dola (M & Hardiansyah, 2026). Dalam penelitian ini, proses pelabelan dilakukan berdasarkan rating yang diberikan pengguna pada Google Play Store, yang dianggap merefleksikan pandangan pengguna tentang aplikasi tersebut. Klasifikasi sentimen dibagi menjadi tiga kategori, yaitu positif, netral, dan negatif. Rating 4 dan 5 dikategorikan sebagai sentimen positif, rating 3 sebagai netral, serta rating 1 dan 2 sebagai negatif. Metode pelabelan yang berdasarkan rating ini memungkinkan proses pelabelan yang lebih konsisten dan objektif sesuai dengan

karakteristik data dari ulasan aplikasi Dola. Label yang diperoleh kemudian digunakan dalam tahap pelatihan dan pengujian model CNN.

### **Pembagian Dataset**

Pembagian dataset dilakukan untuk memisahkan informasi menjadi bagian data pelatihan dan data pengujian, Dengan tujuan untuk mendukung proses pembelajaran mesin. Data pelatihan dimanfaatkan untuk melatih model agar dapat mengenal pola serta keterkaitan dalam informasi, sementara data pengujian berguna untuk mengevaluasi sejauh mana kemampuan model dalam memprediksi data yang belum pernah dilihat sebelumnya. Dalam penelitian ini, dataset dibagi menjadi dua segmen dengan proporsi 70% untuk data latih dan 30% untuk data uji demi memperoleh hasil model yang optimal serta kemampuan generalisasi yang baik.

### **Klasifikasi Model**

Dalam penelitian ini, Sebuah *model Convolutional Neural Network* (CNN) dikembangkan dengan beberapa lapisan penting, yaitu lapisan embedding, lapisan konvolusi, lapisan pooling, lapisan fully connected, serta lapisan output. Masing-masing lapisan berfungsi dalam mengambil fitur serta mengidentifikasi pola hubungan antara kata-kata dalam ulasan pengguna aplikasi Dola. CNN dipilih karena kemampuannya yang efisien dalam mengelola data teks yang tidak terstruktur dan berukuran besar. Metode CNN digunakan sebagai pendekatan utama dalam menganalisis sentimen dari ulasan pengguna aplikasi Dola yang diambil dari Google Play Store.

### **Evaluasi Model**

Pada fase terakhir dari penelitian ini, dilakukan penilaian terhadap model *Convolutional Neural Network* (CNN) dengan menggunakan pendekatan *confusion matrix* untuk menilai efektivitas model dalam melakukan klasifikasi sentimen ulasan pengguna aplikasi Dola

(Saputra et al., 2024). Penilaian dilakukan dengan memanfaatkan berbagai metrik seperti, *Akurasi*, *presisi*, *recall*, dan *F1-score*. Akurasi berfungsi untuk mengevaluasi seberapa tepat prediksi model terhadap keseluruhan data uji, sedangkan presisi menggambarkan seberapa tepat model akurat dalam memprediksi data dalam kategori sentimen tertentu. Selain itu, *recall* digunakan untuk mengukur kemampuan model dalam mengidentifikasi semua data yang termasuk dalam kategori sentimen yang benar, sementara *F1-score* dipakai sebagai nilai rata-rata harmonis antara *precision* dan *recall* untuk menilai sejauh mana kinerja model seimbang. Penilaian ini dipakai untuk mengukur seberapa baik kinerja model CNN dalam klasifikasi sentimen dari ulasan pengguna aplikasi Dola secara akurat dan menyeluruh. Berikut adalah formula penghitungan dari metode confusion matrix.

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

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

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

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

Dengan menggunakan metrik yang telah disebutkan, peneliti bisa mengukur seberapa baik kinerja model *Convolutional Neural Network* (CNN) dalam mengidentifikasi sentimen dari ulasan pengguna aplikasi Dola, serta mengevaluasi seberapa efektif model dalam secara akurat membedakan dengan tepat antara sentimen negatif, netral, dan positif secara akurat. Hasil dari evaluasi ini menjadi tolak ukur utama untuk menilai kinerja model CNN dalam menganalisis pola teks yang terdapat dalam ulasan dari

pengguna yang diambil dari Google Play Store.

## HASIL DAN PEMBAHASAN

Dalam penelitian ini, teknik pengumpulan data dilaksanakan dengan teknik *crawling* menggunakan bahasa pemrograman Python pada platform Google Colaboratory. Data yang dikumpulkan berasal dari komentar pengguna aplikasi Dola yang tersedia di Google Play Store. Proses *crawling* ini berhasil mendapatkan 3000 data ulasan dari pengguna yang mencakup berbagai perspektif mengenai kualitas layanan, fitur aplikasi, serta pengalaman dalam menggunakan aplikasi Dola.

Sebelum melakukan proses klasifikasi sentimen, data harus melalui tahap *pre-processing terlebih dahulu*. Langkah-langkah ini mencakup *cleaning*, *case folding*, *tokenization*, *stopword removal*, dan *stemming*. Tujuan dari proses ini adalah untuk menghilangkan gangguan seperti seperti simbol, angka, karakter khusus, serta kata yang tidak relevan sehingga menghasilkan teks yang lebih teratur dan siap untuk digunakan dalam proses klasifikasi menggunakan metode *Convolutional Neural Network* (CNN). Setelah menyelesaikan tahap *pre-processing* dilakukan, jumlah data yang dapat dimanfaatkan dalam penelitian menjadi sebanyak 2578 data ulasan.

Selanjutnya, data tersebut dibagi menjadi dua segmen, yang terdiri dari 70% data untuk pelatihan (*training data*) dan 30% data pengujian (*testing data*). Data pelatihan digunakan untuk membangun model CNN dalam mengidentifikasi pola bahasa yang sesuai dengan sentimen, sementara data pengujian berfungsi untuk menilai seberapa akurat model dapat melakukan prediksi pada data yang belum pernah dipelajari sebelumnya.

Model *Convolutional Neural Network* (CNN) yang diterapkan pada kajian ini terdiri daripada lapisan penting, yaitu *embedding layer*, *convolution layer*, *pooling layer*, *fully connected layer*, dan

*output layer*. Setiap lapisan memiliki fungsi dalam mengekstraksi fitur dan mengenali hubungan antar kata dalam teks ulasan pengguna aplikasi Dola. Proses pelatihan model dilakukan menggunakan library TensorFlow dan Keras dengan dukungan GPU untuk mempercepat proses komputasi.

Hasil dari pengujian mengindikasikan bahwa model CNN dapat mengategorikan sentimen ulasan dari pengguna aplikasi Dola ke dalam tiga kategori, yakni sentimen positif, netral, dan negatif. Proses evaluasi model dilaksanakan dengan memanfaatkan metode *confusion matrix* yang menggunakan metrik evaluasi berupa *accuracy*, *precision*, *recall*, dan *F1-score*. Melihat hasil evaluasi tersebut, model CNN memperlihatkan kinerja yang memuaskan dalam mendeteksi pola sentimen pada data teks ulasan dari pengguna aplikasi Dola.

Secara umum, ulasan dengan sentimen positif banyak berisi apresiasi terhadap kemudahan penggunaan aplikasi, fitur yang membantu, serta kualitas layanan yang diberikan. Sementara itu, ulasan negatif umumnya berisi kritik terkait gangguan sistem, error aplikasi, maupun kendala penggunaan fitur tertentu. Adapun ulasan netral cenderung berupa komentar informatif tanpa menunjukkan emosi atau opini yang dominan.

Berdasarkan hasil penelitian yang diperoleh, metode *Convolutional Neural Network* (CNN) terbukti efektif digunakan dalam analisis sentimen berbasis teks, khususnya pada ulasan pengguna aplikasi Dola. Kemampuan CNN dalam mengekstraksi fitur secara otomatis dari data teks mampu membantu proses klasifikasi sentimen menjadi lebih akurat. Hasil studi ini diharapkan bisa berfungsi sebagai dasar penilaian untuk para pengembang aplikasi Dola dalam memperbaiki mutu kualitas layanan dan pengalaman pengguna di waktu yang akan datang.

## Pengumpulan Data

Data yang dipakai dalam studi ini di ambil dari komentar pengguna aplikasi Dola yang dapat diakses di Google Play Store. Untuk mengumpulkan data, metode *crawling* diterapkan dengan memanfaatkan bahasa pemrograman Python dan *library* pendukung yang diopersikan pada *platform Google Colaboratory*. Proses Pengambilan data melibatkan ekstraksi ulasan pengguna serta penilaian rating yang diberikan kepada aplikasi Dola.

Pengumpulan data dilakukan selama periode penelitian pada tahun 2025 berhasil mengumpulkan total 3000 data ulasan dari pengguna. Data yang berhasil dikumpulkan mencakup berbagai komentar dan reaksi pengguna mengenai kualitas layanan, fitur dalam aplikasi, tingkat kemudahan penggunaan, serta pengalaman pengguna saat menggunakan aplikasi Dola. Dataset ini selanjutnya dimanfaatkan sebagai sumber utama dalam proses analisis sentimen dengan metode *Convolutional Neural Network* (CNN).

**Tabel 1. Hasil Pengumpulan Data**

tanggal	komentar	rating
5/12/2026 12:34	mantap	5
5/12/2026 12:33	lumayan si tapi kalau jawab matematika nya agak ga jelas tapi masih bagusss tapi ai lumayan anðŸ˜†ðŸ˜†ðŸ˜†	3
5/12/2026 12:31	bola ini mah Fatin orang gua	5
5/12/2026 12:31	ada jawaban yang salah	2
5/12/2026 12:31	please banget updet dong ada fitur bisa chatan sama chat yang dola yang lama maksud ku kan aku tadi chatan sama dola yang udah dari lama sama aku terus eh malah kepencet topik baru yah ilang deh semua nya tambahin dong	5

**Pre-Processing Data**

Proses ini dilaksanakan guna memperoleh informasi yang lebih rapi dan terorganisir dari data asli yang didapat dari tanggapan pengguna aplikasi Dola di Google Play Store. Tahap *pre-processing* merupakan bagian yang sangat penting dalam analisis sentimen karena kualitas data yang baik akan sangat memengaruhi tingkat akurasi model dalam melakukan klasifikasi sentimen. Pada penelitian ini, proses *pre-processing* dilakukan sebelum data digunakan pada model *Convolutional Neural Network* (CNN). Adapun tahapan dalam proses *pre-processing* adalah sebagai berikut.

- a. Tahap *cleaning* dilakukan untuk membersihkan data ulasan pengguna aplikasi Dola dari berbagai elemen yang tidak diperlukan dalam proses analisis sentimen. Pada tahap ini dilakukan penghapusan data duplikat, karakter khusus, simbol, angka, tanda baca, serta teks yang mengandung noise sehingga data menjadi lebih bersih dan mudah diproses oleh model *Convolutional Neural Network* (CNN). Selain itu, atribut yang tidak relevan dalam proses klasifikasi seperti URL, emoji, dan karakter tertentu juga dihilangkan

**Tabel 2. Hasil Cleaning**

Sebelum Cleaning	Sesudah Cleaning
lumayan si tapi kalau jawab matematika nya agak ga jelas tapi masih bagusss tapi ai lumayan an	lumayan si tapi kalau jawab matematika nya agak ga jelas tapi masih bagusss tapi ai lumayan an

- b. Tahap *case folding* dilakukan untuk meningkatkan konsistensi data teks dengan cara mengubah seluruh huruf kapital menjadi huruf kecil (*lowercase*). Proses ini bertujuan agar tidak terjadi perbedaan penulisan kata yang sebenarnya memiliki makna yang sama, sehingga dapat membantu model *Convolutional Neural Network* (CNN)

dalam mengenali pola kata secara lebih konsisten pada proses analisis sentimen.

**Tabel 3. Hasil Case folding**

Sebelum Case Folding	Sesudah Case Folding
lumayan si tapi kalau jawab matematika nya agak ga jelas tapi masih bagusss tapi ai lumayan an	lumayan si tapi kalau jawab matematika nya agak ga jelas tapi masih bagusss tapi ai lumayan an

- c. Tahap *tokenization* dilakukan untuk memecah teks ulasan pengguna aplikasi Dola menjadi bagian-bagian kecil berupa kata atau token. Proses ini bertujuan untuk mempermudah sistem dalam mengenali serta menganalisis pola kata pada setiap ulasan pengguna. Hasil dari proses *tokenization* kemudian digunakan sebagai dasar dalam pengolahan data teks pada model *Convolutional Neural Network* (CNN) untuk proses klasifikasi sentiment

**Tabel 4. Hasil Tokenezing**

Sebelum Tokenezing	Sesudah Tokenezing
lumayan si tapi kalau jawab matematika nya agak ga jelas tapi masih bagusss tapi ai lumayan an	['lumayan', 'si', 'tapi', 'kalau', 'jawab', 'matematika', 'nya', 'agak', 'ga', 'jelas', 'tapi', 'masih', 'lumayan', 'ai', 'an']

- d. Tahap *stopword removal* dilakukan untuk menyingkirkan kata-kata yang tidak berkontribusi secara signifikan dalam analisis sentimen, seperti kata hubung, kata depan, dan kata umum lainnya. Proses ini bertujuan untuk meningkatkan kualitas data serta membantu model *Convolutional Neural Network* (CNN) dalam memusatkan analisis pada kata-kata yang memiliki arti krusial dalam klasifikasi sentimen ulasan pengguna aplikasi Dola.

**Tabel 5. Stopward Removal**

Sebelum Stopward	Sesudah Stopward
lumayan si tapi kalau jawab matematika nya agak ga jelas tapi	['lumayan', 'si', 'matematika', 'nya', 'ga', 'agak ga jelas tapi']

masih bagusss 'bagusss', 'ai',  
tapi ai lumayan 'lumayan', 'an']  
anðY~†ðY~†ðY~†

- e. Tahap *slangword* dilakukan untuk mengubah kata-kata tidak baku, singkatan, atau bahasa gaul yang terdapat pada ulasan pengguna menjadi kata baku sesuai dengan kaidah bahasa Indonesia. Proses ini bertujuan agar data teks menjadi lebih konsisten dan mudah dipahami oleh model *Convolutional Neural Network* (CNN) saat menganalisis sentimen, sehingga hasil klasifikasinya bisa lebih tepat.

**Tabel 6. Slangword**

Sebelum Slangword	Sesudah Slangword
lumayan si tapi kalau jawab matematika nya agak ga jelas tapi masih bagusss tapi ai lumayan anðY~†ðY~†ðY~†	['lumayan', 'si', 'matematika', 'nya', 'tidak', 'bagusss', 'ai', 'lumayan', 'an']

- f. Tahap *stemming* adalah langkah terakhir dalam tahap *pre-processing* yang bertujuan untuk mengubah kata berimbuhan menjadi kata dasar dengan menghapus prefiks dan sufiks pada kata tersebut. Proses ini dilakukan agar kata-kata yang memiliki makna sama tetapi bentuk penulisan berbeda dapat diseragamkan, sehingga membantu model *Convolutional Neural Network* (CNN) dalam mengenali pola kata secara lebih efektif pada proses analisis sentimen.

**Tabel 7. Stemming**

Sebelum Stemming	Sesudah Stemming
lumayan si tapi kalau jawab matematika nya agak ga jelas tapi masih bagusss tapi ai lumayan anðY~†ðY~†ðY~†	['lumayan', 'si', 'matematika', 'nya', 'tidak', 'bagusss', 'ai', 'lumayan', 'an']

**Pelabelan**

Tahap pelabelan pada data dilakukan untuk membagi kometar pengguna aplikasi

Dola ke dalam tiga jenis sentiment yaitu, positif, netral, dan negatif. Sebelum proses pelabelan dilakukan, data ulasan terlebih dahulu melalui berbagai tahapan *pre-processing* seperti *cleaning*, *case folding*, *tokenization*, *stopword removal*, *slangword*, dan *stemming* agar data menjadi lebih bersih dan lebih mudah untuk dianalisis.

Pada penelitian ini, proses pelabelan dilakukan berdasarkan rating yang diberikan pengguna pada Google Play Store. Rating 4 dan 5 dikategorikan sebagai sentimen positif, rating 3 sebagai sentimen netral, sedangkan rating 1 dan 2 dikategorikan sebagai sentimen negatif. Proses pelabelan ini bertujuan untuk membantu model *Convolutional Neural Network* (CNN) dalam mengenali pola bahasa pada setiap kategori sentimen sehingga hasil klasifikasi dapat dilakukan dengan lebih akurat.

**Tabel 8. Hasil Pelabelan**

Komentar	Rating	label
mantap	5	positif
lumayan si matematika nya bagusss ai lumayan an	3	netral
Bola mah fatin orang gua	5	positif
salah	2	negatif
please banget updet fitur chatan chat dola maksud ku chatan dola udah eh pencet topik yah ilang deh nya tambahin	5	positif

**Klasifikasi**

Tahap pelatihan Model *Convolutional Neural Network* (CNN) dilaksanakan untuk mengklasifikasikan sentiment dari data kometar pengguna aplikasi Dola yang telah melewati proses sebelumnya. Tujuan dari proses klasifikasi ini Adalah untuk merancang dan menguji model CNN dalam mengelompokan kometar pengguna ke

dalam kategori sentiment positif, netral, dan negative berdasarkan teks komentar yang tersedia di Google Play Store.

Model CNN yang digunakan memiliki beberapa bagian penting, dimulai dari lapisan *embedding layer* yang bertugas untuk mengubah kata kata dalam bentuk vektor numerik. Selanjutnya digunakan lapisan konvolusi (*Conv1D*) digunakan untuk mengidentifikasi pola-pola penting yang terdapat pada teks ulasan. Setelah itu, digunakan lapisan *global max pooling diterapkan* untuk mengekstrak fitur-fitur yang paling dominan dari hasil konvolusi, dan kemudian dilanjutkan dengan lapisan *dense* yang digunakan untuk klasifikasi sentimen akhir.

Proses pelatihan untuk model dilakukan dengan menggunakan Kumpulan data yang telah dibagi menjadi dua bagian, yaitu data latih dan data uji dengan rasio 70% dan 30%. Pelatihan model direncanakan menggunakan jumlah *epoch* sebanyak 10 putaran. Berdasarkan hasil pelatihan yang ditunjukkan pada gambar, nilai akurasi model mengalami peningkatan pada setiap *epoch*. Pada *epoch* pertama, model memperoleh nilai *accuracy* sebesar 0.7032 dan *validation accuracy* sebesar 0.7306. Seiring bertambahnya proses pelatihan, performa model terus meningkat hingga pada *epoch* ke-10 model memperoleh nilai *accuracy* sebesar 0.9835 dengan *validation accuracy* sebesar 0.8236.

Hasil pengujian model menunjukkan bahwa pendekatan *Convolutional Neural Network* (CNN) berkemampuan untuk mengklasifikasikan sentimen komen pengguna aplikasi Dola dengan cukup baik. Berdasarkan penilaian akhir, model memperoleh tingkat akurasi sebesar 0.8236 atau sekitar 82%. Nilai ini jelas menunjukkan bahwa model CNN dapat dengan mampu mengenali corak sentimen dalam ulasan pengguna secara efektif berdasarkan data teks yang telah diproses terlebih dahulu.

Gambar 2 memperlihatkan tahapan pelatihan model CNN yang diimplementasikan menggunakan bahasa

pemrograman *Python* dengan library *TensorFlow* dan Keras pada platform *Google Colaboratory*. Proses ini dilakukan untuk melatih model agar mampu memahami pola bahasa dari ulasan pengguna aplikasi Dola dan melakukan klasifikasi sentimen secara otomatis.

Model: "sequential\_3"

Layer (type)	Output Shape	Param #
embedding_3 (Embedding)	(None, 100, 128)	640,000
conv1d_5 (Conv1D)	(None, 98, 128)	82,496
max_pooling1d_2 (MaxPooling1D)	(None, 48, 128)	0
conv1d_6 (Conv1D)	(None, 46, 64)	24,544
global_max_pooling1d_3 (GlobalMaxPooling1D)	(None, 64)	0
dense_6 (Dense)	(None, 64)	4,160
dropout_3 (Dropout)	(None, 64)	0
dense_7 (Dense)	(None, 3)	195

Total params: 751,043 (2.87 MB)  
 Trainable params: 751,043 (2.87 MB)  
 Non-trainable params: 0 (0.00 B)

**Gambar 2.** Model CNN yang Digunakan

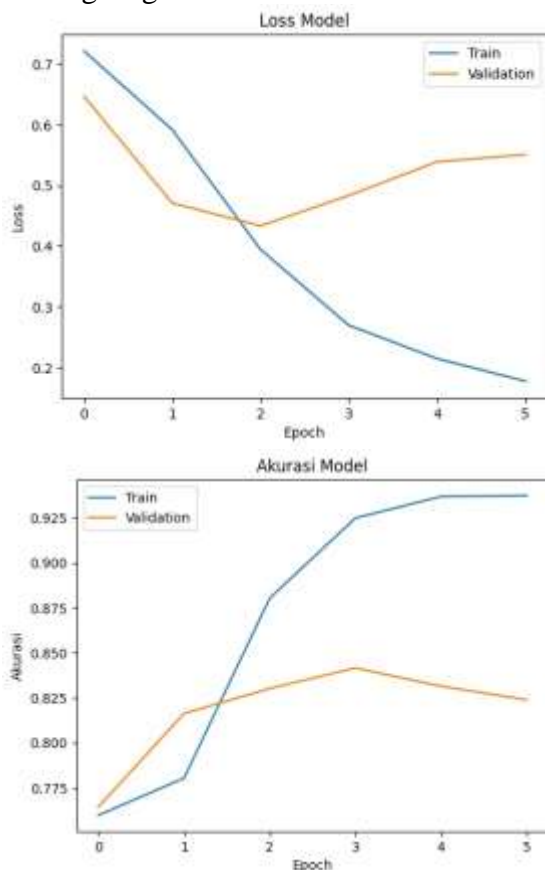
**Tabel 9.** Perbandingan Akurasi pada Setiap Epoch

Epoch	Accuracy/loss	Val accuracy/ Val Loss
Epoch 1	0.7032 / 0.7343	0.7306 / 0.6443
Epoch 2	0.8162/0.5026	0.8430/0.4426
Epoch 3	0.8914/0.3394	0.8527/0.4290
Epoch 4	0.9161 /0.2434	0.8256 /0.4697
Epoch 5	0.9471 /0.1718	0.8450 /0.4597
Epoch 6	0.9597/ 0.1311	0.8508 / 0.5001
Epoch 7	0.9724/ 0.0900	0.8353 / 0.5442
Epoch 8	0.9767 / 0.0705	0.8314 / 0.5722
Epoch 9	0.9811 / 0.0604	0.8275 / 0.6214
Epoch 10	0.9835 / 0.0554	0.8236 / 0.6461

Berdasarkan hasil pelatihan model CNN selama 10 *epoch*, kinerja model menunjukkan kemajuan yang cukup jelas. Pada *epoch* pertama, model memperoleh nilai *accuracy* sebesar 70,32% dengan *validation accuracy* sebesar 73,06%. Seiring bertambahnya epoch, nilai *accuracy* pada data pelatihan terus meningkat hingga mencapai 98,35% pada *epoch* ke-10. Sementara itu, nilai *validation accuracy* mengalami peningkatan pada awal pelatihan dan mencapai nilai tertinggi sebesar 85,27% pada epoch ke-3, kemudian mengalami penurunan secara bertahap hingga 82,36% pada epoch ke-10.

Selain itu, nilai *loss* pada data pelatihan mengalami penurunan dari 0.7343 pada epoch pertama menjadi 0.0554 pada epoch ke-10. Penurunan ini menunjukkan bahwa model CNN mampu mempelajari pola sentimen dari data ulasan dengan baik selama proses pelatihan. Namun, nilai *validation loss* cenderung meningkat setelah *epoch* ke-3, dari 0.4290 menjadi 0.6461 pada epoch ke-10. Kondisi ini menunjukkan bahwa model mulai mengalami *overfitting*, yaitu model bisa ditampilkan dengan bagus di data validasi tidak meningkat secara konsisten.

Sebagaimana ditunjukkan pada Tabel 9 yang menyajikan perbandingan nilai akurasi pada setiap *epoch*, Gambar 3 yang menampilkan proses pelatihan model CNN serta memperlihatkan grafik *accuracy* dan *loss* selama proses pelatihan model berlangsung.



Gambar 3. Grafik Akurasi dan Loss

### Evaluasi

Evaluasi terhadap model dilakukan dengan memakai metode *confusion matrix* dan laporan klasifikasi untuk menilai

seberapa baik model *Convolutional Neural Network* (CNN) dalam mengklasifikasikan sentimen dari ulasan pengguna aplikasi Dola ke dalam tiga kategori yaitu positif, netral, dan negatif. Hasil pengujian pada 516 data uji menunjukkan bahwa model ini berhasil mencapai tingkat akurasi sebesar 82%, yang menunjukkan bahwa model ini cukup baik memiliki dalam mengklasifikasikan sentimen.

Berdasarkan hasil *confusion matrix* pada Gambar 4, model menunjukkan performa terbaik pada kategori sentimen positif. Dari total 377 data sentimen positif, sebanyak 329 data berhasil diklasifikasikan dengan benar sebagai sentimen positif. Namun, masih terdapat beberapa kesalahan klasifikasi, yaitu 39 data positif diprediksi sebagai negatif dan 9 data diprediksi sebagai netral.

Dalam kategori sentimen negatif, model tersebut sukses mengidentifikasi 95 dari total 118 data dengan akurat sebagai sentimen negatif, sementara 23 data lainnya secara keliru diprediksi sebagai sentimen

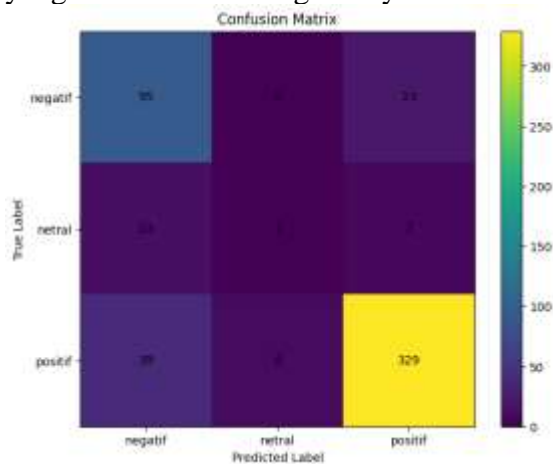
positif. Sebaliknya, untuk kategori sentimen netral, model hanya mampu mengkategorikan 1 dari 21 data dengan tepat. Hal ini menunjukkan bahwa model masih menghadapi kesulitan dalam memahami pola bahasa pada sentimen netral karena jumlah data netral yang relatif lebih sedikit jika dibandingkan dengan kategori lainnya.

Hasil *classification report* yang terlihat di Gambar 5 menunjukkan bahwa kategori sentimen yang positif memperoleh tingkat presisi sebanyak 0.92, *recall* sebesar 0.87, dan *F1-score* sebesar 0.89. angka-angka ini mengidentifikasi bahwa model CNN sangat efektif dalam mengenali sentimen positif pada tinjauan yang diberikan oleh pengguna aplikasi Dola.

Untuk pengkategorian sentimen negatif, model berhasil meraih nilai *precision* sebesar 0.65, *recall* sebesar 0.81, dan *F1-score* sebesar 0.72. Sementara itu, kategori sentimen netral mendapatkan nilai *precision* sebesar 0.10, *recall* sebesar 0.05,

dan *F1-score* sebesar 0.06. Rendahnya kinerja pada kategori netral disebabkan oleh sedikitnya jumlah data netral, yang mengakibatkan model kurang efektif dalam mengenali pola sentimen yang ada.

Secara keseluruhan, model CNN mencapai nilai *macro average F1-score* sebesar 0.56 dan *weighted average F1-score* sebesar 0.82. evaluasi ini mengindikasikan bahwa metode *Convolutional Neural Network* (CNN) sangat efisien dalam menganalisis sentimen dari komentar pengguna aplikasi Dola, terutama dalam mengidentifikasi sentimen positif dan negatif berdasarkan data teks yang diambil dari Google Play Store.



Gambar 4. Confusion Matrix

Classification Report:				
	precision	recall	f1-score	support
negatif	0.65	0.81	0.72	118
netral	0.10	0.05	0.06	21
positif	0.92	0.87	0.89	377
accuracy			0.82	516
macro avg	0.55	0.58	0.56	516
weighted avg	0.82	0.82	0.82	516

Gambar 5. Classification

### Pembahasan

Hasil dari pengujian yang sudah dilakukan menunjukkan bahwa penggunaan metode *Convolutional Neural Network* (CNN) dalam analisis sentiment dari ulasan pengguna aplikasi Dola memperoleh hasil yang baik dengan tingkat akurasi 82%. Penemuan ini menunjukkan bahwa model CNN mampu mengenali pola-pola sentimen dalam ulasan pengguna

aplikasi Dola yang diambil dari Google Play Store, terutama dalam kategori sentimen positif dan negatif.

Langkah-langkah pre-processing data yang mencakup *cleaning*, *case folding*, *tokenization*, *stopword removal*, *slangword*, dan *stemming* terbukti memainkan peran penting dalam meningkatkan mutu data sebelum diproses oleh model CNN. Tahapan ini bertujuan untuk mengurangi gangguan dalam data teks seperti simbol, karakter khusus, istilah yang tidak formal, dan kata-kata yang tidak sesuai, sehingga model dapat mempelajari karakteristik penting dari teks ulasan dengan lebih baik. Dengan data yang memiliki kualitas lebih tinggi, proses pengelompokan sentimen dapat dilakukan dengan lebih efektif.

Hasil pelatihan model menunjukkan bahwa nilai *accuracy* pada data latih mengalami peningkatan yang signifikan pada setiap *epoch*, mulai dari 70,32% pada *epoch* pertama hingga mencapai 98,35% pada *epoch* ke-10. Sementara itu, nilai *validation accuracy* berada pada kisaran 82% hingga 85%. Keadaan tersebut mengindikasikan bahwa model CNN dapat memahami pola data dengan baik, meskipun terdapat indikasi *overfitting* ketika nilai *training accuracy* meningkat jauh lebih tinggi dibandingkan *validation accuracy*. Hal ini juga terlihat dari nilai *validation loss* yang mengalami fluktuasi pada beberapa *epoch* selama proses pelatihan.

Berdasarkan evaluasi dengan menggunakan *confusion matrix* dan *classification report*, model CNN menunjukkan kinerja terbaik di kategori sentimen positif. Model berhasil mencapai nilai *precision* 0.92, *recall* 0.87, dan *F1-score* 0.89 untuk kategori sentimen positif. Sementara itu, untuk kategori sentimen negatif, model meraih nilai *precision* sebesar 0.65, *recall* sebesar 0.81, dan *F1-score* sebesar 0.72. Hasil itu menunjukkan bahwa model cukup efektif dalam mengidentifikasi sentimen negatif,

meskipun masih terdapat beberapa kesalahan dalam klasifikasi.

Kategori sentimen netral menunjukkan kinerja paling rendah dibandingkan dengan kategori yang lain. Hal ini terjadi karena jumlah data sentimen netral yang tergolong sedikit jika dibandingkan dengan data sentimen positif dan negatif, sehingga model mengalami kesulitan dalam memahami pola bahasa pada kategori ini. Ketidakseimbangan jumlah data di setiap kategori sentimen merupakan salah satu faktor yang memengaruhi kinerja model dalam proses klasifikasi

Secara umum, temuan penelitian ini mengindikasikan bahwa teknik *Convolutional Neural Network* (CNN) efisien diterapkan dalam analisis sentimen dari ulasan pengguna aplikasi Dola yang berbasis teks. Kemampuan CNN untuk secara otomatis mengekstrak fitur dari data teks mendukung model dalam mengidentifikasi pola bahasa dan melakukan klasifikasi sentimen dengan tingkat akurasi yang baik. Studi ini juga menyajikan gambaran tentang pandangan pengguna terhadap aplikasi Dola berdasarkan komentar yang ditinggalkan di Google Play Store

Meski begitu, penelitian ini tetap memiliki beberapa batasan, terutama pada jumlah data sentimen netral yang cenderung sedikit dan belum diterapkannya metode optimisasi model atau embedding pra-latih yang lebih rumit. Karena itu, studi selanjutnya dapat dieksplorasi dengan menerapkan metode lain seperti *Long Short-Term Memory* (LSTM), *Bidirectional LSTM*, atau *Bidirectional Encoder Representations from Transformers* (BERT) untuk meningkatkan hasil klasifikasi sentimen dengan lebih baik

## SIMPULAN

Studi ini sukses menerapkan *metode Convolutional Neural Network* (CNN) untuk menganalisis sentimen ulasan pengguna aplikasi Dola yang diambil dari Google Play Store. Model CNN dibangun

dengan beberapa lapisan kunci, yaitu Embedding Layer untuk mengkonversi kata menjadi representasi vektor, Convolutional Layer dan Pooling Layer untuk mengambil fitur penting dari teks ulasan, serta Dense Layer dan Output Layer untuk menghasilkan klasifikasi sentimen dalam kategori positif, netral, dan negatif

Hasil pengujian menunjukkan bahwa model CNN dapat mengklasifikasikan sentimen dengan baik dan mencapai tingkat akurasi 82%. Hasil evaluasi dengan menggunakan *confusion matrix*, *precision*, *recall*, dan *F1-score* mengindikasikan bahwa model menunjukkan performa terbaik dalam kategori sentimen positif, sedangkan kinerja pada kategori sentimen netral masih terbilang rendah akibat jumlah data yang lebih sedikit dibandingkan dengan kategori lainnya

Tahapan *pre-processing* seperti *cleaning*, *case folding*, *tokenization*, *stopword removal*, *slangword*, dan *stemming* terbukti membantu meningkatkan kualitas data sehingga proses klasifikasi sentimen dapat dilakukan secara lebih optimal. Selain itu, penggunaan bahasa pemrograman Python dengan library TensorFlow dan Keras mempermudah proses pembangunan dan pelatihan model CNN dalam analisis sentimen berbasis teks.

Secara keseluruhan, penelitian menunjukkan bahwa pendekatan *Convolutional Neural Network* (CNN) efektif dalam menganalisis sentimen dari ulasan pengguna aplikasi Dola. Temuan dari penelitian ini diharapkan dapat mendukung pengembang aplikasi dalam mengerti pandangan pengguna serta menjadi acuan untuk meningkatkan kualitas fitur dan layanan aplikasi di waktu yang akan datang. Penelitian selanjutnya bisa diperluas dengan meningkatkan jumlah dataset, mengoptimalkan parameter model, atau menerapkan metode lain seperti LSTM dan BERT untuk meningkatkan akurasi dan keseimbangan dalam klasifikasi sentimen.

**DAFTAR PUSTAKA**

- Aprianti, Y., Hananto, A. L., & Hilabi, S. S. (2025). *Klasifikasi Sentimen Komentar Pengguna pada Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes*. 101–110. <https://doi.org/10.47002/metik.v9i1.1023>
- Ayuningtiyas, P., Tania, K. D., & Sari, W. K. (2025). *Sentiment-Based Knowledge Discovery pada Aplikasi iPusnas Menggunakan Metode Machine Learning dan Deep Learning*. 9(5), 2486–2497.
- Dwiyanaputra, R., Murpratiwi, S. I., & Aranta, A. (2025). *ANALISIS SENTIMEN PADA PENGGUNA APLIKASI X TERHADAP PEMILIHAN UMUM PRESIDEN 2024 MENGGUNAKAN METODE CONVOLUTIONAL NEURAL NETWORK (CNN)*. 9(1), 635–642.
- Kamal, S., Rachman, F. I., & Wahyuni, T. (2025). *Analisis Sentimen Text Dengan Metode CNN Study Kasus Tempat Wisata Makassar*. 7(1), 48–57.
- M, F. A. J. H., & Hardiansyah, B. (2026). *ANALISIS SENTIMEN TERHADAP ULASAN ONLINE UNTUK KENDARAAN LISTRIK MENGGUNAKAN DEEP LEARNING DI PLATFORM X*. 10(1), 1458–1465.
- Martha, A. A., Eviyanti, A., Fitriani, A. S., Ratna, I., Astutik, I., Sidoarjo, U. M., Sidoarjo, K., Timur, J., Sentimen, A., & Scraping, W. (2026). *WEB SCRAPING DAN NATURAL LANGUAGE PROCESSING MENGGUNAKAN CNN UNTUK ANALISIS SENTIMEN LINTAS PLATFORM DIGITAL*. 10(1), 1621–1627.
- Maulana, M. T., Muflikhah, L., & Fatyanosa, T. N. (2025). *Analisis Sentimen Pengguna Indodax Menggunakan FastText dan Convolutional Neural Network (CNN)*. 9(6).
- Qudsi, D. H., Lubis, J. H., Syaliman, K. U., & Najwa, N. F. (2021). *ANALISIS SENTIMEN PADA DATA SARAN MAHASISWA TERHADAP KINERJA DEPARTEMEN DI PERGURUAN TINGGI MENGGUNAKAN CONVOLUTIONAL NEURAL NETWORK*. *Jurnal Teknologi Informasi Dan Ilmu Komputer (JTIK)*, 8(5), 1067–1076. <https://doi.org/10.25126/jtiik.202184842>
- Rinardi, A., Sri, C., Aditya, K., & Chandranegara, D. R. (2024). *Analisis Sentimen Ulasan Fitur Music Aplikasi Instagram Pada Google Play Store Menggunakan Metode Convolutional Neural Network (CNN)*. 6(4), 415–424.
- Saputra, S. N., Setiaji, G. G., Teja, M., & Cipta, A. (2024). *Perbandingan Kinerja RNN dan CNN Dalam Klasifikasi Sentimen Ulasan Pengguna Aplikasi di Play Store*. 6(1), 349–362. <https://doi.org/10.47065/josyc.v6i1.6408>
- Sitio, G. Y., Rumapea, S. A., & Lumbanraja, D. P. (2024). *Analisis Sentimen Pemindahan Ibu Kota Negara Di Media Sosial Twitter Menggunakan metode Convolutional Neural Network (CNN)*. 3(2), 97–104.
- Vector, S., & Algorithm, M. (2025). *Klasifikasi Sentimen Layanan pada Aplikasi by . U menggunakan Algoritma Support Vector Machine Classification of Service Sentiments on the by . U Application using the*. 14, 1967–1976.
- Yuniarossy, B. A., Hindrayani, K. M., & Terza, A. (2024). *ANALISIS SENTIMEN TERHADAP ISU FEMINISME DI TWITTER MENGGUNAKAN MODEL CONVOLUTIONAL NEURAL NETWORK*. 5(1), 477–491.