

## ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI BY.U PADA GOOGLE PLAY STORE MENGGUNAKAN ALGORITMA CNN

### SENTIMENT ANALYSIS OF USER REVIEWS OF THE BY.U APPLICATION ON GOOGLE PLAY STORE USING THE CNN ALGORITHM

Altra Yudha Mawlana<sup>1</sup>, Erwin Dwika Putra<sup>2</sup>

Teknik Informatika, Universitas Muhammadiyah Bengkulu<sup>1,2</sup>

[yuda19mi@gmail.com](mailto:yuda19mi@gmail.com)<sup>1</sup>

#### ABSTRACT

The increasing use of the by.U digital service application is accompanied by a growing number of user reviews on the Google Play Store, which reflect user satisfaction and dissatisfaction. The large volume of unstructured text data makes manual analysis less effective, thus requiring an automated approach through sentiment analysis. This study aims to classify user review sentiments into positive, neutral, and negative categories using the Convolutional Neural Network (CNN) method. A total of 3,000 reviews were collected, of which 2,984 reviews were processed through preprocessing before sentiment labeling based on user ratings. The dataset was then divided into 80% training data and 20% testing data. This study conducted several experimental scenarios to analyze the effect of preprocessing techniques and model configurations on classification performance. The results show that the CNN model with stopword removal achieved the best accuracy of 76%. Meanwhile, the overall evaluation of the model resulted in an accuracy of 73%, with a precision of 75%, recall of 73%, and F1-score of 73%, indicating a fairly good classification performance. However, the model still has limitations in accurately identifying neutral sentiment.

**Keywords:** Convolutional Neural Network (CNN), sentiment analysis, user reviews, text classification, Google Play Store

#### ABSTRAK

Peningkatan penggunaan aplikasi layanan digital by.U diikuti oleh bertambahnya jumlah ulasan pengguna pada Google Play Store yang mencerminkan tingkat kepuasan dan ketidakpuasan pengguna. Banyaknya data teks yang tidak terstruktur menyebabkan analisis secara manual menjadi kurang efektif, sehingga diperlukan pendekatan otomatis melalui analisis sentimen. Penelitian ini bertujuan untuk mengklasifikasikan sentimen ulasan pengguna ke dalam kategori positif, netral, dan negatif menggunakan metode Convolutional Neural Network (CNN). Sebanyak 3.000 ulasan berhasil dikumpulkan, dan 2.984 ulasan diproses melalui tahapan preprocessing sebelum dilakukan pelabelan sentimen berdasarkan rating pengguna. Dataset kemudian dibagi dengan rasio 80% data latih dan 20% data uji. Penelitian ini melakukan beberapa skenario pengujian untuk menganalisis pengaruh preprocessing dan konfigurasi model terhadap kinerja klasifikasi. Hasil penelitian menunjukkan bahwa model CNN dengan penerapan stopword removal menghasilkan akurasi terbaik sebesar 76%. Sementara itu, hasil evaluasi keseluruhan model menunjukkan akurasi sebesar 73% dengan nilai precision sebesar 75%, recall sebesar 73%, dan F1-score sebesar 73%, yang mengindikasikan performa klasifikasi yang cukup baik. Namun, model masih memiliki keterbatasan dalam mengidentifikasi sentimen netral secara optimal.

**Kata kunci:** Convolutional Neural Network (CNN), ulasan pengguna, klasifikasi teks, Google Play Store

#### PENDAHULUAN

Penggunaan aplikasi layanan digital seperti by.U terus meningkat seiring dengan kebutuhan masyarakat terhadap layanan telekomunikasi berbasis digital. Sejalan dengan peningkatan jumlah pengguna, aplikasi by.U juga menerima banyak ulasan di Google Play Store yang disertai dengan rating dan komentar teks. Ulasan tersebut merupakan bentuk *electronic word of mouth* (e-WOM) yang

mencerminkan pengalaman serta tingkat kepuasan pengguna terhadap layanan yang diberikan[1]. Rating yang diberikan pengguna berfungsi sebagai indikator awal dalam menunjukkan kecenderungan penilaian pengguna, baik yang merepresentasikan kepuasan maupun ketidakpuasan terhadap aplikasi by.U[2].

Namun, jumlah ulasan yang sangat besar dan beragam menyulitkan pengembang untuk mengetahui

kecenderungan pendapat pengguna secara menyeluruh apabila dilakukan secara manual. Oleh karena itu, analisis sentimen diperlukan sebagai pendekatan otomatis untuk mengelompokkan ulasan pengguna berdasarkan rating yang diberikan. Pada penelitian ini, rating digunakan sebagai dasar penentuan kelas sentimen, di mana rating 5 dan 4 dianggap sebagai sentimen positif, rating 3 dianggap sebagai sentimen netral, dan rating 2 dan 1 dianggap sebagai sentimen negatif. Pelabelan sentimen dapat dilakukan secara konsisten dan objektif dengan metode ini. Selanjutnya, teknik pemrosesan bahasa alami (NLP) dan algoritma *Convolutional Neural Network* (CNN) digunakan untuk menganalisis ulasan teks dan mengidentifikasi pola bahasa yang menggambarkan perasaan pengguna terhadap aplikasi by.U.

*Convolutional Neural Network* (CNN) adalah salah satu arsitektur deep learning yang banyak digunakan dalam analisis sentimen. CNN pertama kali dikembangkan untuk mengenali pola visual pada data citra melalui arsitektur LeNet-5[3]. Seiring dengan perkembangan penelitian di bidang *Natural Language Processing* (NLP), CNN juga banyak diterapkan pada pemrosesan teks, termasuk tugas klasifikasi sentimen. CNN memiliki kemampuan dalam mengekstraksi pola-pola lokal pada teks, seperti kata atau frasa penting, yang berperan dalam merepresentasikan kecenderungan opini pengguna[4]. CNN sangat penting untuk memahami umpan balik pengguna pada aplikasi by. U di Google Play Store, yang memiliki banyak data tidak terstruktur. Studi ini menggunakan CNN untuk membagi ulasan pengguna ke dalam kategori sentimen positif, netral, dan negatif. Ini dilakukan dengan membagi data yang diberi label dengan penilaian dan menggunakan CNN untuk melakukan klasifikasi. Ini dilakukan agar model dapat dengan tepat memprediksi kategori sentimen untuk umpan balik baru[5].

Dalam penelitian ini, Prosedur klasifikasi terdiri atas dua langkah utama:

langkah pelatihan untuk menciptakan model dengan memanfaatkan kumpulan data yang sudah diberikan label berdasarkan penilaian, dan langkah pengujian untuk mengevaluasi kemampuan model dalam memperkirakan kategori sentimen pada data yang tidak dikenal sebelumnya. Dalam pelatihan model *neural network* seperti CNN, penentuan jumlah *epoch* menjadi salah satu faktor penting karena memengaruhi proses pembelajaran model terhadap pola data[6]. Selain itu, untuk meningkatkan efisiensi komputasi, proses pelatihan dibagi ke dalam sejumlah *batch* berukuran lebih kecil sehingga pembaruan bobot model dapat dilakukan secara bertahap dan lebih optimal.

Penelitian oleh Nurfauziyah et al. (2025) memanfaatkan algoritma CNN untuk mempelajari perasaan pengguna Aplikasi X tentang pemilihan umum presiden Indonesia 2024. Dataset yang digunakan sebanyak 10.250 tweet yang melalui tahapan preprocessing serta pelabelan sentimen menggunakan metode VADER. Data kemudian diklasifikasikan ke dalam tiga kelas sentimen, maksudnya, ada tiga jenis: positif, netral, dan negatif. Hasil penelitian menunjukkan bahwa kombinasi CNN dengan pembobotan TF-IDF memiliki akurasi sebesar 84%, dengan skor akurasi, recall, dan F1 masing-masing sebesar 83% [7]. Penelitian ini menunjukkan bahwa CNN dapat mengolah teks media sosial yang tidak terstruktur.

Penelitian tambahan yang dilakukan oleh Ginni Yema Sitio et al. menggunakan metode CNN untuk mengamati perasaan orang tentang pemindahan Ibu Kota Indonesia di media sosial Twitter. Emosi memiliki sifat netral, positif, dan negatif. Model CNN dengan 100 epoch dan ukuran batch 4 memiliki tingkat akurasi sebesar 94,94%, menurut hasil penelitian. Nilai ketepatan, recall, dan F1-nya masing-masing mencapai 67%, 62%, dan 64% [8]. Hasil ini menunjukkan bahwa CNN memiliki kemampuan untuk mengkategorikan perasaan dari data teks media sosial.

Selain itu, Sukma Nindi Listyarini dan Dimas Aryo Anggoro (2021) menggunakan CNN untuk mengamati bagaimana orang berperasaan tentang pelaksanaan Pilkada 2020 di tengah pandemi Covid-19. Penelitian ini mengumpulkan lima ratus tweet dan membaginya ke dalam dua kelas sentimen, positif dan negatif, dengan menggunakan API Twitter. Sebelum pelatihan model CNN dengan variasi jumlah lapisan dan epoch, tahapan preprocessing termasuk membersihkan, memotong stopword, tokenisasi, dan stemming. Hasil penelitian menunjukkan bahwa model CNN dengan empat lapisan convolutional dan seratus epoch memiliki akurasi sebesar 90%. Hasil ini menunjukkan bahwa menambah lapisan dan epoch dapat meningkatkan kinerja klasifikasi[9].

Studi tambahan yang dilakukan oleh Ade Zakharia, Herliyani Hasanah, dan Agustina Srirahayu (2025) meneliti penggunaan AI dalam seni digital melalui CNN. Dataset, yang mencakup 1.000 tweet berbahasa Indonesia, dibagi menjadi dua kategori sentimen: negatif dan positif. Model CNN menggunakan embedding layer, Conv1D, dropout, global max pooling, dan dense layer. Selain itu, model dilatih selama sepuluh epoch menggunakan optimizer Adam. Hasil penelitian menunjukkan akurasi sebesar 71,29%, dengan performa model lebih baik dalam mengklasifikasikan sentimen positif dibandingkan negatif[10]. Penelitian ini menunjukkan bahwa meskipun CNN cukup efektif, pengaturan parameter dan jumlah data sangat memengaruhi hasil klasifikasi serta potensi overfitting. Variasi jumlah dataset, jumlah layer, epoch, serta teknik preprocessing dan pembobotan kata menjadi faktor penting yang memengaruhi tingkat akurasi model. Oleh karena itu, penggunaan metode CNN dalam penelitian ini dinilai relevan dan memiliki dasar empiris yang kuat.

## METODE

Penelitian ini berfokus pada cara mengkategorikan sentimen ulasan pengguna aplikasi by.U, yaitu menemukan kecenderungan opini positif, netral, atau negatif. Metode ini menggunakan algoritma CNN yang mengkategorikan sentimen berdasarkan rating dan teks ulasan yang dikumpulkan dari Google Play Store. Rating 4 dan 5 dianggap sebagai sentimen yang positif, rating 3 dianggap sebagai sentimen netral, dan rating 1 dan 2 dianggap sebagai sentimen yang negatif[11]. Selanjutnya, data melalui tahap preprocessing, yang mencakup perbaikan teks, penyusunan case, tokenisasi, penghapusan kata tidak relevan (juga dikenal sebagai penghapusan stopword), dan stemming. Agar dapat digunakan sebagai input model, proses sequencing dan padding digunakan untuk mengubah data yang telah diproses ke bentuk numerik. Model CNN dibangun dan dilatih menggunakan bahasa pemrograman Python dan library TensorFlow dan Keras pada platform Google Colaboratory yang memanfaatkan kemampuan GPU. Dataset dibagi menjadi dua kategori: data pelatihan dan data pengujian, masing-masing dengan proporsi 80% dan 20%. Langkah terakhir dari penelitian ini adalah menilai model menggunakan matrix confusion serta metrik akurasi, presisi, recall, dan skor F1 untuk menilai kinerja klasifikasi secara keseluruhan.



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**, sebagaimana ditunjukkan pada Tabel 1.

**Tabel 1. Example Hasil Crawling Data**

Tanggal	Komentar	Rating
06/12/2025	Dulu suka banget pake by.U karna murah kenceng, sekarang harga MAHAL masa aktif bentar bikin pelanggan pindah operator.	3
09/12/2025	Udah bisaaa, makasih admin by.U bantuanya terkhusus dan ramah banget.	5
10/12/2025	Aplikasi by.U cukup membantu buat beli paket, tampilannya juga simpel.	5
10/12/2025	Kenapa sih lama-lama by.U sinyal jadi susah, operator tolong di perbaiki dong, saya beli kuota tuh pake duit, ini kenapa setiap nelpon langsung hilang sinyal.	1
10/12/2025	Aplikasi ini sangat bagus buat pengguna kartu by.U, karena benar-benar memudahkan para pengguna kartu by.U	5

### Pre-Processing Data

Salah satu tahap krusial dalam studi ini adalah pre-processing, yang dilakukan untuk mempersiapkan data sebelum diolah oleh model CNN. Tahap ini bertujuan untuk meningkatkan kualitas data dengan menghilangkan kebisingan dan komponen yang tidak diperlukan. Ini akan membuat data lebih terorganisir dan siap untuk analisis. Pre-processing terdiri dari beberapa langkah, termasuk pembersihan teks, pengolahan case, penghapusan stopword, stemming, dan tokenisasi. Setelah itu, seluruh hasil dari tahapan ini diubah menjadi vektor numerik, yang dapat

dipakai sebagai input untuk model CNN saat melakukan analisis sentimen.

### Pelabelan Data

Proses pelabelan data sangat penting untuk analisis sentimen karena memungkinkan untuk mengidentifikasi kategori emosi dari setiap ulasan pengguna. Dalam penelitian ini, penentuan sentimen dilakukan berdasarkan skor pengguna di Google Play Store, yang dianggap menunjukkan pendapat pengguna tentang aplikasi by.U. rating emosi dikategorikan ke dalam tiga kategori: positif, netral, dan negatif. rating 4 dan 5 dianggap sebagai emosi positif, rating 3 dianggap netral, dan rating 1 dan 2 dianggap sebagai emosi negatif. Metode pelabelan berbasis rating ini memungkinkan proses pelabelan yang konsisten dan objektif sesuai dengan karakteristik data ulasan aplikasi. Keputusan dari pelabelan ini seterusnya dimanfaatkan sebagai label kategori dalam tahap latihan dan pengujian model CNN[12].

### Pembagian Dataset

Pembagian dataset merupakan langkah untuk memisahkan data menjadi dua kelompok yang lebih kecil, yaitu data pelatihan dan data uji, yang bertujuan mendukung proses pembelajaran mesin. Data pelatihan berfungsi untuk membantu model mengenali pola serta hubungan yang ada dalam data, sehingga prediksinya menjadi lebih akurat. Namun, seberapa efektif model dalam memprediksi data yang belum pernah dilihat sebelumnya dievaluasi melalui pengujian yang dilakukan terpisah dari proses pelatihan[13]. Untuk mencapai kinerja model yang optimal dan kemampuan generalisasi yang baik, penelitian ini membagi dataset menjadi dua segmen. 80% dari dataset digunakan untuk pelatihan dan 20% sisanya untuk pengujian

### Membangun Model

Dalam penelitian ini, model *Convolutional Neural Network* (CNN)

terdiri dari sejumlah lapisan penting. Model ini mencakup lapisan embedding, lapisan konvolusi, lapisan pengambilan sampel, lapisan yang sepenuhnya terhubung, dan lapisan hasil. Setiap lapisan memiliki tugas untuk mengekstraksi fitur dan mengidentifikasi pola hubungan antar kata dalam teks ulasan pengguna. CNN dapat mengklasifikasikan teks dengan baik, terutama data tidak terstruktur dan berukuran besar[14]. Hasilnya, metode utama penelitian ini adalah CNN untuk melakukan analisis sentimen ulasan pengguna untuk aplikasi by.U, yang tersedia di Google Play Store.

#### Pelatihan dan Pengujian Model

Untuk menguji model CNN, data latih dari tahap preprocessing dan pembagian dataset digunakan. Selama proses pelatihan, algoritma optimasi Adam digunakan untuk secara adaptif memperbarui bobot model berdasarkan nilai gradien karena model memiliki sentimen negatif, netral, dan positif. Selain itu, karena model memiliki fungsi kerugian kategori *cross-entropy*, algoritma ini juga digunakan untuk memperbaiki nilai gradien[15]. Data uji digunakan untuk menguji kemampuan model untuk memprediksi perasaan pengguna terhadap aplikasi setelah pelatihan selesai pada interval waktu yang by.U yang belum pernah terjadi sebelumnya.

#### Evolusi Model

Evaluasi model dilakukan untuk menilai kemampuan algoritma untuk mengklasifikasikan sentimen ulasan. Tahap evaluasi ini bertujuan untuk mengukur kemampuan algoritma untuk melakukan klasifikasi secara akurat. Evaluasi ini dilakukan dengan menggunakan beberapa metrik yang biasa digunakan dalam analisis sentimen, seperti sebagai berikut[16]:

a. *Accuracy* merupakan metrik evaluasi yang dipakai untuk menunjukkan seberapa tepat model klasifikasi; nilai ini diperoleh dengan membandingkan jumlah prediksi yang benar dengan

keseluruhan data yang diuji. Nilai akurasi didapat dari perbandingan antara jumlah data yang benar-benar diklasifikasikan dengan total data, yang dapat dilihat pada rumus (1):

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

- TP: jumlah data positif yang diprediksi benar
  - TN: jumlah data negatif yang diprediksi benar
  - FP: jumlah data negatif yang diprediksi sebagai positif
  - FN: jumlah data positif yang diprediksi sebagai negative
- b. *Precision* adalah metode penilaian yang digunakan untuk menilai akurasi model dalam meramalkan kelas positif. Nilai precision menunjukkan rasio prediksi positif yang akurat dibandingkan dengan total data prediksi positif, seperti tercantum dalam rumus (2):

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

- c. Kemampuan model untuk menemukan dan mengklasifikasikan data positif diukur dengan metrik evaluasi yang disebut recall. Nilai recall diperoleh dengan membandingkan jumlah prediksi positif yang benar dengan jumlah data positif yang sebenarnya, yang dapat dilihat pada rumus (3):

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

- d. *F1-Score* adalah metode penilaian yang diterapkan untuk menyeimbangkan akurasi dan integritas model saat melakukan klasifikasi. Ini menyatukan nilai precision dan recall dalam bentuk rata-rata harmonis, sebagaimana ditunjukkan dalam rumus (4):

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

Peneliti dapat menggunakan metrik evaluasi ini untuk menilai tingkat

keefektifan model CNN dalam mengklasifikasikan sentimen ulasan pengguna aplikasi By.U. Mereka juga dapat melihat sejauh mana model mampu membedakan secara akurat sentimen positif, netral, dan negatif.

## HASIL DAN PEMBAHASAN

Data ulasan pengguna untuk aplikasi by.U yang diperoleh dari Google Play Store digunakan dalam penelitian ini. Sebanyak 3.000 ulasan pengguna dikumpulkan melalui teknik *scraping*. Ulasan tersebut mencerminkan berbagai tanggapan pengguna terhadap layanan, fitur, kualitas jaringan, harga paket, serta pengalaman penggunaan aplikasi.

Setelah melalui tahap preprocessing yang meliputi *cleaning*, *case folding*, *tokenizing*, *stopword removal*, dan *padding*, jumlah data yang dapat digunakan dalam proses analisis menjadi 2.984 ulasan. Tahap ini dilakukan untuk menghilangkan noise seperti simbol, angka, dan karakter yang tidak relevan agar data lebih siap diproses oleh model.

Dataset kemudian dibagi menjadi 80% untuk pelatihan dan 20% untuk pengujian. Tujuan dari bagian ini adalah untuk melatih model CNN dan menguji kemampuan model untuk diterapkan pada data baru.

Hasil pelatihan dan pengujian menunjukkan bahwa model CNN dapat dengan akurat mengkategorikan sentimen ulasan ke dalam tiga kategori: negatif, netral, dan positif. Selain itu, hasil ini menunjukkan bahwa ulasan pengguna dapat dianalisis secara otomatis di Google Play Store untuk mengetahui persepsi pengguna terhadap aplikasi by.U.

Secara umum, variasi isi ulasan menunjukkan perbedaan pola bahasa antara sentimen positif, negatif, dan netral. Ulasan negatif biasanya berisi kritik terkait gangguan layanan atau aplikasi, sedangkan ulasan positif lebih banyak berisi penghargaan atas kemudahan penggunaan dan biaya paket. Hal ini menjadi dasar penting untuk diskusi tentang seberapa

efektif model mengenali pola bahasa berdasarkan sentimen.

### Pre-processing data

Pada tahap persiapan, data ulasan pengguna dari aplikasi by.U yang dikumpulkan dari Google Play Store dibersihkan dan diubah menjadi format yang mampu diproses oleh model CNN. Dalam analisis sentimen, ini sangat penting karena data teks ulasan sering mengandung gangguan seperti tanda baca berlebihan, simbol, angka, dan gaya penulisan yang berbeda. Akibatnya, preprocessing dilakukan untuk meningkatkan kualitas data, yang memungkinkan analisis sentimen yang lebih baik. Berikut ini adalah penjelasan tentang prosedur pre-processing[17]:

- Cleaning* dilakukan dengan menghapus pola, tanda baca, angka, dan simbol khusus dari teks ulasan. Tahap ini bertujuan untuk mengurangi suara agar data teks lebih bersih dan siap untuk diproses pada tahap analisis berikutnya. Tabel 2 menunjukkan contoh hasil pembersihan.

**Tabel 2. Hasil Cleaning**

Text Asli	Sesudah Cleaning
Aplikasi By.U cukup membantu buat beli paket, tampilannya juga simpel.	Aplikasi ByU cukup membantu buat beli paket tampilannya juga simpel

- Semua huruf dalam teks ulasan diubah menjadi huruf kecil selama proses *case folding*. Tahap ini bertujuan untuk memastikan format teks sama sehingga proses analisis berikutnya tidak terpengaruh oleh perbedaan penggunaan huruf kapital. Tabel 3 menunjukkan contoh hasil *folding case*.

**Tabel 3. Hasil Case Folding**

Sesudah Cleaning	Sesudah Case Folding
Aplikasi ByU cukup membantu buat beli paket	aplikasi byu cukup membantu buat beli paket

tampilannya juga simpel	paket tampilannya juga simpel
-------------------------	-------------------------------

- c. Kata-kata umum seperti "ada", "adalah", "agar", dan lainnya dihilangkan untuk menghilangkan *stopwords*. Tujuan dari langkah ini adalah untuk mengurangi noise pada data sambil berkonsentrasi pada analisis kata-kata yang memiliki makna penting. Dengan demikian, kinerja model dalam mengklasifikasikan sentimen dapat ditingkatkan. Tabel 4 menunjukkan contoh hasil *stopwords*.

**Tabel 4. Hasil Stopwords**

Sesudah Case Folding	Sesudah Stopwords
aplikasi byu cukup membantu buat beli paket tampilannya juga simpel	aplikasi byu cukup membantu beli paket tampilannya simpel

- d. *Stemming* adalah proses mengembalikan kata ke bentuk aslinya. Tujuan dari langkah ini adalah untuk menyatukan berbagai variasi kata sehingga dianggap sebagai entitas yang sama oleh model. Salah satu contohnya adalah perubahan dari kata "mendukung" menjadi "dukung". Oleh karena itu, redundansi data dapat dikurangi dan proses analisis dapat dilakukan dengan lebih efisien. Tabel 5 menunjukkan beberapa hasil *stemming*.

**Tabel 5. Hasil Stemming**

Sesudah Cleaning	Sesudah Stemming
aplikasi cukup membantu paket tampilannya simpel	aplikasi byu cukup bantu beli paket tampilannya simpel

- e. Teks ulasan dibagi menjadi bagian kata yang lebih kecil untuk melakukan proses tokenisasi. Proses pemrosesan bahasa alami sangat penting karena memungkinkan setiap kata direpresentasikan dalam bentuk token yang selanjutnya dikonversi menjadi vektor numerik sebagai input bagi model CNN. Contoh hasil *tokenisasi* seperti Tabel 6.

**Tabel 6. Hasil Tokenisasi**

Sesudah Stemming	Sesudah Tokenisasi
aplikasi byu cukup bantu beli paket tampilannya simpel	['aplikasi', 'byu', 'cukup', 'bantu', 'beli', 'paket', 'tampil', 'simpel']

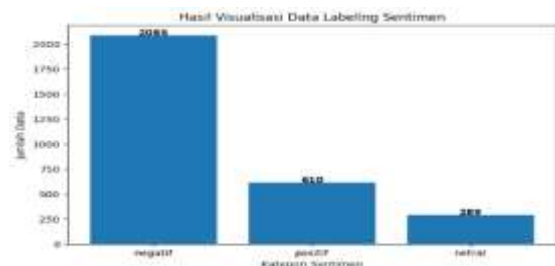
- f. Untuk memastikan panjang total data, yaitu seratus token, setiap ulasan dibungkus. Karena model CNN membutuhkan input yang seragam agar dapat diproses secara konsisten selama tahap pelatihan dan pengujian, proses ini diperlukan.

**Labeling Data**

Hasil pelabelan sentimen dari 2.984 ulasan pengguna untuk aplikasi by.U yang dikumpulkan dari Google Play Store menunjukkan distribusi sentimen yang berbeda. Proses pelabelan sentimen pada penelitian ini ditentukan berdasarkan rating yang diberikan oleh pengguna, di mana rating tinggi dikategorikan sebagai sentimen positif, rating rendah sebagai sentimen negatif, dan rating menengah sebagai sentimen netral.

**Table 7. Example Hasil Labeling Data**

Komentar	Rating	Label
dulu suka banget pake byu karna murah kenceng sekarang harga mahal masa aktif bentar bikin langgan pindah operator	3	Netral
udah bisa makasih admin byu bantuanya terkhusus dan ramah banget	5	Positif
suka banget pake kartu byu internetnya murah	5	Positif
kenapa sih lama lama byu sinyal jadi susah operator tolong di perbaiki saya beli kuota tuh pake duit ini kenapa setiap telfon langsung hilang sinyal	1	Negatif
aplikasi ini sangat bagus untuk pengguna kartu byu karna memudahkan para pengguna kartu byu	5	Positif



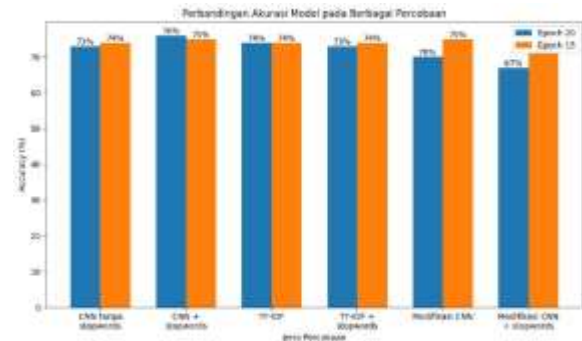
**Gambar 2. Hasil Visualisasi Data Labeling**

### Pelatihan dan Pengujian Model

Model CNN dilatih dengan 80% data sebagai data latih dan 20% sebagai data uji. Ukuran batchnya adalah 32. Untuk melihat bagaimana jumlah epoch memengaruhi kinerja model, penelitian ini menggunakan dua skenario pelatihan: 15 epoch dan 20 epoch. Epoch adalah satu siklus penuh proses pelatihan di mana seluruh data pelatihan diproses oleh model. Jumlah epoch yang lebih banyak memungkinkan model untuk mempelajari pola dan karakteristik data lebih dalam. Namun, penggunaan epoch yang terlalu banyak dapat menyebabkan *overfitting*, yaitu keadaan di mana model terlalu mengadaptasi data pelatihan sehingga tidak berperformansi baik pada data pengujian. Sebaliknya, jumlah epoch yang berlebihan dapat mengakibatkan *underfitting*, di mana model tidak mampu mengenali pola data dengan baik. Karena itu, variasi dalam jumlah epoch dalam penelitian ini bertujuan untuk mencari keseimbangan terbaik antara kemampuan pembelajaran model dan kemampuannya untuk generalisasi terhadap data baru. Data uji hanya dimanfaatkan untuk mengevaluasi kemampuan model ketika diterapkan pada data baru. Mereka tidak dimanfaatkan selama proses pelatihan

**Tabel 8. Data Hasil Akurasi Pada Enam Percobaan**

Jenis Percobaan		Accuracy Epoch 20	Accuracy Epoch 15
Model CNN tanpa stopwords pada proses preprocessing		73%	74%
Dengan tambahan stopwords pada proses preprocessing		76%	75%
Menggunakan TF-IDF		74%	74%
Gabungan TF-IDF dan stopwords		73%	74%
Modifikasi arsitektur CNN		70%	75%
Modifikasi arsitektur CNN + stopwords		67%	71%



**Gambar 3. Perbandingan Akurasi Model CNN pada Berbagai Percobaan**

Dalam penelitian ini, percobaan dilakukan untuk menilai sentimen dari ulasan pengguna aplikasi by.U dengan menggunakan metode CNN yang memiliki pengaturan dan parameter pelatihan model yang bervariasi. Uji coba itu dilakukan untuk mengetahui dampak pengaturan model pada hasil klasifikasi sentimen. Skenario percobaan yang dilakukan adalah sebagai berikut:

- Percobaan 1: Pada percobaan pertama, model CNN dilatih tanpa menerapkan proses *stopword removal* pada tahap *pre-processing*. Hasil penelitian menunjukkan bahwa model memperoleh akurasi sebesar 73% pada era ke-20 dan 74% pada era ke-15.
- Percobaan 2: Pada percobaan kedua, proses *stopword removal* diterapkan pada tahap *pre-processing*. Hasil pengujian menunjukkan peningkatan kinerja model. Akurasi epoch ke-20 sebesar 76% dan epoch ke-15 sebesar 75%.
- Percobaan 3: Percobaan ketiga menerapkan metode Term Frequency–Inverse Document Frequency (TF-IDF) untuk merepresentasikan karakteristik teks. Hasil uji menunjukkan bahwa model tidak memperlihatkan perbedaan signifikan dalam kinerja antara epoch ke-20 dan ke-15, dengan tingkat akurasi sebesar 74% untuk keduanya.
- Percobaan 4: Pada percobaan keempat, digunakan kombinasi representasi fitur TF-IDF dan proses *stopword removal* pada tahap *pre-processing*. Hasil pengujian menunjukkan bahwa model tidak menunjukkan peningkatan yang

- signifikan dalam kinerja dibandingkan percobaan sebelumnya; sebaliknya, ia memperoleh akurasi sebesar 73% pada epoch ke-20 dan 74% pada epoch ke-15.
- e. Percobaan 5: Pada percobaan kelima, dilakukan modifikasi terhadap arsitektur model CNN untuk melihat pengaruh perubahan struktur jaringan terhadap kinerja klasifikasi. Hasil tes menunjukkan bahwa model memiliki akurasi sebesar 70% pada epoch ke-20 dan 75% pada epoch ke-15. Hasil ini menunjukkan bahwa modifikasi arsitektur belum meningkatkan performa secara konsisten dibandingkan dengan konfigurasi sebelumnya.
  - f. Pecobaan 6: Pada percobaan keenam, dilakukan modifikasi arsitektur CNN yang dikombinasikan dengan penerapan *stopword removal* pada tahap *pre-processing*. Hasil penelitian menunjukkan bahwa model memiliki tingkat akurasi terendah pada epoch ke-20 dan ke-15, masing-masing sebesar 67% dan 71%. Hal ini menunjukkan bahwa kombinasi modifikasi arsitektur dan penghapusan *stopword* belum mampu meningkatkan kinerja model secara optimal.

Hasil pengujian menunjukkan bahwa Penerapan penghapusan *stopword* pada arsitektur CNN standar memberikan kinerja terbaik dibandingkan variasi yang lain. Dalam percobaan kedua, model CNN yang menggunakan *stopwords* mencapai akurasi tertinggi 76% pada epoch ke-20 dan 75% pada epoch ke-15. Sementara itu, penggunaan TF-IDF maupun kombinasi TF-IDF dan *stopwords* tidak memberikan peningkatan performa yang signifikan dibandingkan konfigurasi dasar. Selain itu, modifikasi arsitektur CNN juga belum mampu meningkatkan hasil klasifikasi secara konsisten, bahkan pada beberapa percobaan menunjukkan penurunan akurasi.

Selain akurasi, studi ini menerapkan metrik presisi, recall, dan F1-score untuk masing-masing kelas sentimen. Hasil uji menunjukkan bahwa model CNN

memperoleh akurasi sebesar 73%, presisi sebesar 78%, recall sebesar 87%, dan F1-score sebesar 82%. Untuk kelas sentimen netral, tingkat akurasi mencapai 17%, recall berada di angka 20%, dan F1-score sebesar 82%. Sementara itu, dengan rata-rata nilai precision 75%, recall 73%, dan F1-score 73%, kelas positif memperoleh precision 88%, recall 57%, dan F1-score 69%, yang mengindikasikan bahwa meskipun prediksi positif cukup akurat, masih terdapat sejumlah data positif yang belum terdeteksi dengan baik secara efektif.

**Tabel 9. Nilai Precision, Recall, dan F1-Score Evaluasi Model**

Jenis Klasifikasi	Precision	Recall	F1-Score
Negatif	78%	87%	82%
Netral	17%	20%	18%
Positif	88%	57%	69%

## SIMPULAN

Penelitian ini menemukan bahwa metode *Convolutional Neural Network* (CNN) dapat membagi ulasan pengguna aplikasi by.U di Google Play Store ke dalam tiga kategori: positif, netral, dan negatif. Data yang digunakan terdiri dari 2.984 ulasan yang telah melewati tahap preprocessing dan dibagi menjadi data latih dan data uji dengan rasio 80:20. Berdasarkan hasil beberapa skenario pengujian, model terbaik diperoleh pada penerapan *stopword removal* dengan akurasi tertinggi sebesar 76% pada epoch ke-20. Sementara itu, Hasil evaluasi dengan *confusion matrix* menunjukkan akurasi sebesar 73% dengan nilai akurasi rata-rata berat sebesar 75%, *recall* sebesar 73%, dan *skor F1* sebesar 73%. Hasil ini menunjukkan bahwa model melakukan klasifikasi sentimen dengan cukup baik. Hasil penelitian juga menunjukkan bahwa penerapan *stopword removal* memberikan pengaruh positif terhadap peningkatan akurasi model, sedangkan penggunaan TF-IDF dan modifikasi arsitektur CNN tidak memberikan peningkatan performa yang signifikan. Namun, akibat karakteristik data yang sering kali ambigu, kelas tersebut

menunjukkan nilai precision, recall, dan F1-score yang rendah, yang menandakan bahwa model masih menghadapi kesulitan dalam mengidentifikasi sentimen netral. Sebagai hasilnya, penelitian tambahan perlu memperbanyak jumlah dataset, melakukan penyeimbangan kategori, dan merancang model yang lebih rumit seperti Long Short-Term Memory (LSTM) atau Bidirectional Encoder Representations from Transformers (BERT) guna meningkatkan stabilitas dan akurasi hasil klasifikasi

#### DAFTAR PUSTAKA

- [1] S. Fransiska and A. I. Gufroni, "Sentiment Analysis Provider by . U on Google Play Store Reviews with TF-IDF and Support Vector Machine ( SVM ) Method," vol. 7, no. 2, pp. 203–212, 2020.
- [2] S. Vector and M. Algorithm, "Klasifikasi Sentimen Layanan pada Aplikasi by . U menggunakan Algoritma Support Vector Machine Classification of Service Sentiments on the by . U Application using the," vol. 14, pp. 1967–1976, 2025.
- [3] S. Tinggi, M. Klimatologi, J. M. No, and T. Tinggi, "Klasifikasi Cuaca Berbasis Citra dengan Model CNN LeNet-5 yang Dimodifikasi," no. 204, pp. 401–410, 2022.
- [4] S. Oktavia, E. Putri, A. A. Arifiyanti, A. Rezha, and E. Najaf, "Convolutional Neural Network Approach for Aspect-Based Sentiment Analysis of Tourism Reviews," vol. 8, no. 1, 2025, doi: 10.32877/bt.v8i1.2582.
- [5] S. N. Saputra, G. G. Setiaji, M. Teja, and A. Cipta, "Perbandingan Kinerja RNN dan CNN Dalam Klasifikasi Sentimen Ulasan Pengguna Aplikasi di Play Store," vol. 6, no. 1, pp. 349–362, 2024, doi: 10.47065/josyc.v6i1.6408.
- [6] A. Rinardi, C. Sri, K. Aditya, and D. R. Chandranegara, "Analisis Sentimen Ulasan Fitur Music Aplikasi Instagram Pada Google Play Store Menggunakan Metode Convolutional Neural Network ( CNN )," vol. 6, no. 4, pp. 415–424, 2024.
- [7] R. Dwiyanaputra, S. I. Murpratiwi, and A. Aranta, "ANALISIS SENTIMEN PADA PENGGUNA APLIKASI X TERHADAP PEMILIHAN UMUM PRESIDEN 2024 MENGGUNAKAN METODE CONVOLUTIONAL NEURAL NETWORK ( CNN )," vol. 9, no. 1, pp. 635–642, 2025.
- [8] G. Y. Sitio, S. A. Rumapea, and D. P. Lumbanraja, "Analisis Sentimen Pemindahan Ibu Kota Negara Di Media Sosial Twitter Menggunakan metode Convolutional Neural Network ( CNN )," vol. 3, no. 2, pp. 97–104, 2024.
- [9] S. N. Listyarini and D. A. Anggoro, "Analisis Sentimen Pilkada di Tengah Pandemi Covid-19 Menggunakan Convolution Neural Network ( CNN ) Sentiment Analysis Regional Elections During Pandemic Covid-19 Using Convolution Neural Network ( CNN )," vol. 1, no. 7, pp. 261–268, 2021.
- [10] A. Zakharia *et al.*, "ANALISIS SENTIMEN PUBLIK DI TWITTER TENTANG DIGITAL," vol. 13, no. 3, pp. 536–545.
- [11] M. T. Maulana, L. Muflikhah, and T. N. Fatyanosa, "Analisis Sentimen Pengguna Indodax Menggunakan FastText dan Convolutional Neural Network ( CNN )," vol. 9, no. 6, 2025.
- [12] C. R. Anissa, K. D. Tania, W. K. Sari, S. Informasi, and U. Sriwijaya, "Sentiment Analysis on Google Play Store Reviews to Measure User Perception of the Gojek Application Using CNN," vol. 9, no. 6, pp. 3322–3328, 2025.
- [13] A. Liawati *et al.*, "ANALISIS SENTIMEN KOMENTAR POLITIK DI MEDIA SOSIAL X

- DENGAN PENDEKATAAN DEEP LEARNING,” vol. 7, no. 6, pp. 3557–3563, 2023.
- [14] F. A. J. H. M and B. Hardiansyah, “ANALISIS SENTIMEN TERHADAP ULASAN ONLINE UNTUK KENDARAAN LISTRIK MENGGUNAKAN DEEP LEARNING DI PLATFORM X,” vol. 10, no. 1, pp. 1458–1465, 2026.
- [15] H. A. Wijaya *et al.*, “Penggunaan algoritma cnn dengan adam optimizer dalam aplikasi web untuk prediksi jenis kelamin bunga,” vol. 9, no. 3, pp. 3825–3832, 2025.
- [16] Y. Akbar, A. N. Ihsan, S. Tinggi, I. Komputer, and C. Karya, “Analisis sentimen,” vol. 6, 2023.
- [17] K. Setiawan, K. J. Apriyanto, S. Tinggi, I. Komputer, and C. Karya, “DEVELOPING A NAIVE BAYES-BASED SENTIMENT ANALYSIS MODEL FOR,” vol. 8, pp. 1111–1121, 2025.