

## Penerapan Model LSTM dan CNN Untuk Klasifikasi Sentimen Pada Ulasan Aplikasi Roblox

Ipin Sugiyarto<sup>1</sup>, Umi Faddillah<sup>2</sup>, Lady Agustin Fitriana<sup>3\*</sup>

<sup>1</sup>Program Studi Sistem Informasi, Universitas Nusa Mandiri

<sup>2,3</sup>Program Studi Sistem Informasi, Universitas Bina Sarana Informatika

\*Lady.lag@bsi.ac.id

### Abstrak

Dalam era digital yang berkembang pesat, platform permainan daring seperti Roblox telah menjadi ruang sosial interaktif kompleks, di mana pengguna berinteraksi, berkreasi, dan membangun pengalaman virtual bersama. Penelitian ini bertujuan membandingkan performa dua arsitektur deep learning, yaitu Convolutional Neural Network (CNN) dan Long Short-Term Memory (LSTM), dalam mengklasifikasikan sentimen ulasan pengguna Roblox dari Google Play Store ke dalam kategori positif, negatif, dan netral. Dataset terdiri dari 5.000 ulasan berbahasa Indonesia yang dihimpun melalui web scraping dengan pustaka `google_play_scraper`. Pra-pemrosesan mencakup cleaning, case folding, tokenization, normalisasi kata tidak baku, penghapusan stopword, stemming, dan pelabelan sentimen berbasis leksikon. Data dikonversi ke representasi numerik via Tokenizer dan padding, lalu dibagi 80:20 untuk training dan testing. CNN mencapai akurasi 89%, precision 0,88, recall 0,81, dan F1-score 0,83 lebih unggul dibanding LSTM (akurasi 86,60%, precision 0,80, recall 0,82, F1-score 0,81). CNN efektif mengekstrak pola spasial teks, sementara LSTM menangkap hubungan temporal antar kata. Penelitian ini menegaskan keunggulan CNN untuk analisis sentimen teks pendek berbahasa Indonesia, menyediakan benchmark model deep learning untuk ulasan aplikasi gaming, serta implikasi praktis bagi pengembang Roblox dalam memahami umpan balik pengguna guna perbaikan fitur.

Kata kunci : Analisis Sentimen, Convolutional Neural Network (CNN), Deep Learning, Long Short-Term Memory (LSTM), Roblox

### Abstract

*In the rapidly evolving digital era, online gaming platforms like Roblox have transformed into complex interactive social spaces where users interact, create, and co-build virtual experiences. This study aims to compare the performance of two deep learning architectures Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) in classifying user reviews of the Roblox app from the Google Play Store into positive, negative, and neutral sentiment categories. The dataset comprises 5,000 Indonesian-language reviews collected via web scraping using the `google_play_scraper` library. Preprocessing involved text cleaning, case folding, tokenization, normalization of informal words, stopword removal, stemming, and lexicon-based sentiment labeling. Data were converted to numerical representations using Tokenizer and padding, then split into 80% training and 20% testing subsets. CNN achieved superior performance with 89% accuracy, 0.88 precision, 0.81 recall, and 0.83 F1-score, outperforming LSTM (86.60% accuracy, 0.80 precision, 0.82 recall, 0.81 F1-score). CNN effectively extracts spatial patterns in text, while LSTM captures temporal word dependencies. This research affirms CNN's superiority for short Indonesian text sentiment analysis, provides a deep learning benchmark for gaming app reviews, and offers practical implications for Roblox developers to interpret user feedback for feature enhancements.*

Keywords : Sentiment Analysis, Convolutional Neural Network (CNN), Deep Learning, Long Short-Term Memory (LSTM), Roblox.

### 1. Pendahuluan

Dalam era digital yang mengalami kemajuan pesat dewasa ini, platform permainan daring

tidak lagi sekadar menjadi media hiburan, tetapi juga telah menjadi medium sosial yang membentuk cara individu berinteraksi dan

berpengalaman di dunia maya. Salah satu platform yang memperoleh perhatian besar di kancah global adalah Roblox, yang per Februari 2025 mencatatkan jumlah pengguna aktif harian melebihi 85,3 juta di seluruh dunia<sup>[1]</sup>. Basis pengguna yang sangat luas, terutama berasal dari kalangan anak-anak hingga remaja, menjadikan Roblox sebagai ekosistem digital yang dinamis serta terus mengalami inovasi. Seiring dengan pertumbuhan jumlah penggunanya, ulasan terhadap aplikasi Roblox di Google Play Store juga meningkat secara signifikan. Kondisi ini membuka peluang bagi para peneliti untuk menelaah lebih jauh persepsi serta tingkat kepuasan pengguna melalui pendekatan analisis sentimen, yang dapat mengungkap dinamika emosional dan pengalaman pengguna secara lebih mendalam. Melakukan analisis sentimen terhadap ulasan pengguna memiliki urgensi yang tinggi karena dapat memberikan pemahaman komprehensif mengenai reaksi emosional pengguna terhadap berbagai aspek aplikasi, mulai dari fitur, performa teknis, hingga isu keamanan. Menurut Liu (2020), analisis sentimen berperan penting dalam mendeteksi pola persepsi publik secara otomatis melalui representasi teks, sehingga mampu membantu pengembang memahami kebutuhan serta ekspektasi pengguna<sup>[2]</sup>. Beberapa penelitian terdahulu yang memanfaatkan algoritma machine learning

konvensional seperti *Support Vector Machine* (SVM) dan *Naive Bayes* mengungkapkan bahwasanya sebagian besar ulasan pengguna terhadap Roblox cenderung bernada positif, dengan tingkat akurasi klasifikasi mencapai sekitar 90%<sup>[3]</sup>. Meskipun demikian, pendekatan berbasis *machine learning* tradisional masih memiliki keterbatasan, terutama dalam menangkap konteks semantik dan hubungan antar kata pada teks ulasan yang panjang atau bersifat naratif.

Sejalan dengan perkembangan pesat teknologi kecerdasan buatan, penerapan metode deep learning semakin meluas dalam bidang analisis sentimen. Hal ini disebabkan oleh kemampuannya yang unggul dalam mempelajari serta merepresentasikan karakteristik teks secara lebih mendalam, termasuk hubungan kontekstual dan pola kompleks yang sulit ditangkap oleh pendekatan tradisional. Dengan memanfaatkan struktur jaringan saraf berlapis, deep learning mampu mengekstraksi makna yang lebih kaya dari data teks, sehingga menghasilkan interpretasi sentimen yang lebih akurat dan adaptif terhadap variasi bahasa alami. Dua pendekatan arsitektur yang banyak digunakan untuk mendukung proses ini meliputi *Long Short-Term Memory* (LSTM) dan *Convolutional Neural Network* (CNN). Model LSTM terbukti efektif dalam memahami hubungan berurutan antar kata dan

mempertahankan konteks pada teks yang panjang, karena arsitekturnya dirancang untuk mengatasi permasalahan vanishing gradient pada jaringan saraf tradisional<sup>[4]</sup>. Sebaliknya, CNN unggul dalam mengenali pola-pola lokal serta frasa penting yang muncul di dalam teks, sehingga sangat cocok untuk melakukan klasifikasi sentimen terhadap ulasan yang lebih singkat dan langsung mengekspresikan opini inti <sup>[5]</sup>.

Berbagai penelitian membuktikan bahwasanya CNN dapat melampaui performa LSTM ketika data teks bersifat ringkas dan struktur bahasanya lebih eksplisit. Selain itu, sejumlah studi juga memperkenalkan arsitektur hybrid CNN-LSTM yang menggabungkan keunggulan keduanya, dengan hasil menunjukkan peningkatan stabilitas serta akurasi dalam proses klasifikasi dibandingkan penggunaan model tunggal <sup>[6][7]</sup>. Studi lainnya menegaskan bahwasanya CNN memiliki keunggulan dalam efisiensi waktu pemrosesan, sedangkan LSTM lebih peka terhadap urutan kata yang berpengaruh pada nuansa sentimen di dalam teks <sup>[8][9]</sup>.

Berdasarkan uraian sebelumnya, tujuannya dari penelitian ini guna menganalisis sentimen ulasan pengguna aplikasi Roblox dari Google Play Store menggunakan dua metode deep learning, yaitu *Long Short-Term Memory (LSTM)* maupun *Convolutional Neural Network (CNN)*. Penelitian

juga membandingkan kinerja kedua model melalui metrik evaluasi di antaranya akurasi, *precision*, *recall*, dan F1-score untuk menentukan model paling efektif dalam mengklasifikasikan ulasan menjadi positif, negatif, atau netral. Temuan ini diharapkan dapat mendukung pengembangan sistem analisis ulasan serta memberikan wawasan bagi pengembang Roblox dalam menambah tingkatan kualitas layanan dan pengalaman pengguna.

## **2. Tinjauan Pustaka**

### **2.1. Penelitian Terkait**

Penelitian mengenai pengembangan *website company profile* dan sistem informasi berbasis web telah banyak dilakukan, baik pada konteks akademik maupun bisnis. Penelitian mengenai *framework CodeIgniter* sebagai media informasi publik sudah dilakukan <sup>[11]</sup>, namun belum mengimplementasikan sistem manajemen konten (*content management system*) yang memungkinkan pengelolaan data secara dinamis.

- Penelitian tentang *company profile* dengan objek perusahaan berskala nasional yang berfokus pada perancangan *website company profile* menggunakan *framework PHP dan Bootstrap* dilakukan pada 2022. Hasilnya menunjukkan bahwa *website* yang dirancang secara profesional mampu meningkatkan kredibilitas perusahaan dan memperluas jangkauan informasi <sup>[12]</sup>.

Penelitian tersebut menekankan pentingnya desain antarmuka dan struktur navigasi, namun belum mengintegrasikan sistem manajemen konten (CMS) yang memungkinkan pembaruan data secara mandiri oleh pengguna non-teknis.

- Penelitian lain mengembangkan *website company profile* berbasis *framework Laravel* dengan pendekatan *object-oriented programming* <sup>[13]</sup>. Sistem yang dihasilkan memiliki tampilan modern dan fitur portofolio yang dinamis. Namun, penelitian tersebut berfokus pada keamanan dan autentikasi pengguna tanpa memberikan perhatian khusus terhadap aspek branding perusahaan di ruang digital. Akibatnya, fungsi *website* lebih kuat pada sisi teknis daripada membangun identitas visual perusahaan secara menyeluruh.
- Sementara itu, penelitian lainnya merancang *website company profile* untuk lembaga pendidikan menggunakan *framework CodeIgniter* dengan metode *waterfall* <sup>[14]</sup>. *Website* tersebut berhasil menyajikan informasi akademik, profil lembaga, dan data tenaga pengajar. Meski demikian, sistem masih bersifat statis tanpa fitur pengelolaan konten secara mandiri, serta belum mengoptimalkan aspek desain visual sebagai bagian dari strategi *branding*.

Berdasarkan penelitian diatas, penerapan *framework CodeIgniter* untuk pengembangan *website company profile* perlu menekankan integrasi antara fungsi sistem dan strategi branding digital. Penelitian-penelitian sebelumnya berfokus pada penyajian data dan fitur teknis, sementara penelitian ini menghadirkan pembaruan berupa pengembangan sistem berbasis *CodeIgniter* dengan penambahan fitur CMS dan desain visual yang disesuaikan dengan identitas perusahaan startup DCC. Pendekatan ini tidak hanya memperhatikan aspek teknologi, tetapi juga mengutamakan pengalaman pengguna dan citra merek yang konsisten di dunia digital

## 2.2. Landasan Teori

### 1. Analisis Sentimen

Analisis sentimen sebagai bagian penting dari *Natural Language Processing (NLP)* yang berfokus pada pengenalan dan pengelompokan opini atau emosi dalam teks ke dalam tiga kategori utama: positif, negatif, maupun netral <sup>[10]</sup>. Pendekatan ini telah menjadi metode yang banyak diterapkan dalam berbagai konteks, mulai dari penilaian terhadap ulasan produk dan jasa, analisis opini politik, hingga pengamatan terhadap reaksi publik terhadap layanan digital <sup>[11]</sup>. Secara umum, proses analisis sentimen melibatkan beberapa tahapan utama yang saling berkaitan, yakni pra-pemrosesan data

teks (*text preprocessing*), ekstraksi fitur (*feature extraction*), klasifikasi sentimen (*sentiment classification*), serta evaluasi hasil (*evaluation*) untuk mengukur performa model yang digunakan.

## 2. Deep Learning

*Deep learning* merupakan subbidang dari *machine learning* yang memanfaatkan struktur jaringan saraf berlapis guna memperoleh pembelajaran secara otomatis dari data dan mengekstraksi representasi data tanpa memerlukan banyak campur tangan manual [12]. Dalam konteks analisis teks, pendekatan ini memiliki kemampuan untuk memahami hubungan semantik, konteks, serta struktur bahasa yang lebih kompleks dibandingkan dengan metode pembelajaran tradisional [13]. Beberapa model deep learning yang umum digunakan dalam pemrosesan data teks meliputi *Convolutional Neural Network (CNN)*, *Recurrent Neural Network (RNN)* maupun *Long Short-Term Memory (LSTM)*. Ketiga model tersebut dapat menyesuaikan bobot pembelajaran secara dinamis berdasarkan urutan kata maupun pola spasial yang terdapat dalam teks, sehingga berkontribusi terhadap peningkatan akurasi dan keandalan hasil klasifikasi sentimen [14].

## 3. Long Short-Term Memory (LSTM)

*Long Short-Term Memory (LSTM)* adalah evolusi dari arsitektur *Recurrent Neural Network*

(*RNN*), dirancang guna menangani tantangan *vanishing gradient* yang kerap muncul pada jaringan saraf tradisional. Model ini memiliki struktur internal yang lebih kompleks, terdiri atas sel memori (*memory cell*) serta tiga gerbang pengendali utama, yaitu input gate, output gate dan forget gate. Kombinasi ketiga komponen tersebut memungkinkan LSTM untuk mempertahankan serta memanfaatkan informasi penting dalam jangka waktu yang panjang, sehingga mampu memahami konteks urutan kata maupun keterkaitan antar bagian teks dengan lebih efektif [15]. Dalam konteks analisis sentimen, LSTM terbukti sangat sesuai untuk menangani teks dengan panjang variatif atau narasi yang memiliki hubungan antarkalimat yang kompleks, karena kemampuannya dalam memproses informasi berurutan secara kontekstual dan mendalam [16].

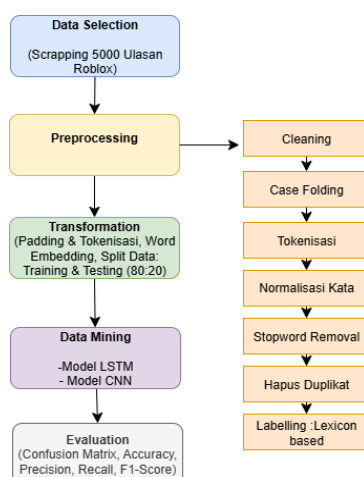
## 4. Convolutional Neural Network (CNN)

Meskipun CNN awalnya dikembangkan untuk pengolahan citra, arsitekturnya juga terbukti efektif dalam klasifikasi teks. CNN bekerja dengan menerapkan filter konvolusi ke matriks representasi kata (*word embedding*) untuk mendeteksi pola lokal seperti frasa penting yang mengindikasikan sentimen [17]. Kelebihan CNN terletak pada efisiensi komputasi dan kemampuannya dalam menangkap fitur spasial secara lokal [18]. Dalam beberapa penelitian, CNN menunjukkan performa tinggi dalam klasifikasi

sentimen, terutama untuk teks pendek atau ulasan langsung<sup>[19]</sup>

### 3. Metode Penelitian

Penelitian ini mengadaptasi metode *Knowledge Discovery in Databases (KDD)* sebagai kerangka kerja konseptual untuk menggambarkan proses sistematis dalam mengekstraksi dan menemukan pengetahuan dari kumpulan ulasan pengguna aplikasi Roblox yang didapatkan melalui Google Play Store. Pendekatan ini digunakan agar memastikan setiap tahap analisis berjalan secara terstruktur dan berurutan, mulai dari pengumpulan hingga interpretasi hasil. Adapun tahapan penelitian dalam penerapan metode KDD pada studi ini dijabarkan pada gambar 1 sebagai berikut:



Gambar 1. Tahapan Penelitian

#### 3.1. Data Selection

Tahap *data selection* dilakukan untuk mengumpulkan ulasan pengguna aplikasi Roblox yang didapatkan dari Google Play Store.

Proses pengambilan data dilaksanakan secara otomatis mengadopsi pustaka *google\_play\_scraper* dengan bahasa Python. Pada penelitian ini, data yang diambil berjumlah 5.000 baris ulasan pengguna dengan parameter language 'id' (Bahasa Indonesia) dan metode penyortiran menggunakan *Sort.NEWEST*, sehingga ulasan yang diperoleh merupakan ulasan terbaru. Setiap data ulasan yang diambil terdiri dari beberapa atribut, yaitu *Review ID*, *Username*, *Rating*, *Review Text*, dan *Date*.

#### 3.2. Preprocessing

Pada tahap preprocessing, data ulasan dibersihkan dan dipersiapkan agar siap digunakan dalam analisis. Proses ini mencakup *cleaning* untuk menghapus karakter tak relevan, *case folding* untuk menyeragamkan huruf menjadi kecil, serta tokenisasi untuk memecah teks menjadi kata-kata. Selanjutnya dilakukan normalisasi kata agar kata tidak baku disesuaikan dengan bentuk standar, diikuti *stopword removal* untuk menghapus kata umum yang tidak berpengaruh pada makna sentimen. Data duplikat dihapus melalui proses *duplicate removal*, kemudian dilakukan pelabelan (*labeling*) menggunakan metode *lexicon-based* untuk menentukan kategori sentimen positif, netral atau negatif pada setiap ulasan.

#### 3.3. Tranformation

Tahap *transformation* dilakukan untuk mengubah data teks hasil *preprocessing* menjadi bentuk numerik yang dapat diproses oleh model deep learning. Proses ini meliputi padding dan tokenisasi, yaitu mengubah kata menjadi urutan angka dan menyeragamkan panjang urutan agar sesuai dengan kebutuhan model. Data yang sudah diproses selanjutnya dibagi menjadi *data training* (80%) maupun *testing* (20%) agar model dapat dilatih serta dievaluasi secara seimbang dan representatif.

### 3.4. Data Mining

Tahapan *data mining* berperan sebagai inti dari keseluruhan proses analisis, di mana pendekatan deep learning dimanfaatkan untuk mengidentifikasi dan mengelompokkan sentimen pada ulasan pengguna Roblox. Penelitian ini menerapkan dua model utama. Model LSTM dibangun dengan lapisan *Embedding*, *LSTM layer*, dan *Dense layer* untuk mempelajari urutan kata serta konteks kalimat. Sementara itu, model CNN mencakup *Embedding layer*, *Convolutional layer*, *Pooling layer*, dan *Fully Connected layer* yang berfungsi mengekstraksi fitur spasial dari teks.

### 3.5. Evaluation

Tahap *evaluation* dilakukan untuk menilai performa model LSTM dan CNN dalam mengklasifikasi sentimen ulasan aplikasi Roblox.

Evaluasi dilakukan dengan membandingkan hasil prediksi model terhadap label sebenarnya pada data testing. Evaluasi kinerja model dilaksanakan dengan mengadopsi metrik seperti akurasi (*accuracy*), presisi (*precision*), *recall*, maupun *F1-score*, yang berfungsi menilai kemampuan model dalam mengidentifikasi sentimen secara akurat dan konsisten

## 4. Hasil dan Pembahasan

### 4.1. Hasil Penelitian

#### 1. Data Selection

Pada tahap data selection, ulasan pengguna Roblox diambil dari Google Play Store dengan mempergunakan pustaka *google\_play\_scraper* berbasis Python, dengan hasil proses *scraping* ditampilkan melalui gambar 2 berisi kode dan data yang berhasil dikumpulkan. .

```
from google_play_scraper import reviews, Sort
app_id = 'com.roblox.client'
def get_reviews(app_id, lang='id', count=5000, sort=Sort.NEWEST, filter_score_
try:
    result, continuation_token = reviews(
        app_id,
        lang=lang,
        country='id',
        sort=sort,
        count=count,
        filter_score_with=filter_score_with,
        filter_device_with=filter_device_with,
        continuation_token=continuation_token
    )
```

Gambar 2. Proses *Scrapping*

Sebanyak 5.000 ulasan berhasil dikumpulkan dengan parameter bahasa Indonesia (*lang='id'*) dan metode penyortiran *Sort.NEWEST* untuk memperoleh ulasan terbaru yang ditampilkan pada gambar 3.

Rating	Review Text	Date
5	seru banget, plccc game nya ga bikin cepet bos...	2025-08-05 12:10:00
5	game nya bnyak dan juga seru	2025-08-05 12:09:58
5	keren info Mabar pfts	2025-08-05 12:09:58
5	GAME YANG SANGAT SERU DAN LUAR BIASA. INI GAME...	2025-08-05 12:09:38
4	gamenya seru , tapi saya masih merasakan bug y...	2025-08-05 12:09:38

Gambar 3. Hasil Scrapping

## 2. Preprocessing Data

Tahap *preprocessing* dilakukan untuk mempersiapkan data ulasan agar siap diolah pada tahap pemodelan. Proses ini bertujuan membersihkan dan menstandarkan teks ulasan yang diperoleh dari hasil *scrapping*. Pada tahap ini dilakukan serangkaian langkah pengolahan teks yang meliputi *cleaning*, *case folding*, *tokenisasi*, normalisasi kata, *stopword removal*, penghapusan duplikat, serta pelabelan (*labeling*).

### a. Cleaning dan Case Folding

Tahap pra-pemrosesan meliputi *cleaning* untuk menghapus *noise* (seperti URL, simbol, dan angka) serta *case folding* untuk menyeragamkan teks menjadi huruf kecil guna menjamin konsistensi data sebelum tahap klasifikasi.

cleaning	case_folding
seru banget plccc game nya ga bikin cepet bosennn	seru banget plccc game nya ga bikin cepet bosennn
game nya bnyak dan juga seru	game nya bnyak dan juga seru
keren info Mabar pfts	keren info mabar pfts
GAME YANG SANGAT SERU DAN LUAR BIASA INI GAME ...	game yang sangat seru dan luar biasa ini game ...
gamenya seru tapi saya masih merasakan bug ya...	gamenya seru tapi saya masih merasakan bug ya...

Gambar 4. Hasil *Cleaning* dan *Case Folding*

### b. Normalisasi Kata dan Tokenisasi

Normalisasi kata dilakukan untuk mengubah kata tidak baku menjadi baku menggunakan kamus

dari GitHub melalui fungsi *replace\_taboo\_words()*, misalnya “bgt” menjadi “banget” dan “gk” menjadi “nggak”. Teks kemudian ditokenisasi dengan *split()* agar siap diproses pada tahap *word embedding* serta pemodelan LSTM dan CNN.

normalisasi	tokenize
seru banget plccc game ya tidak bikin cepet bo...	[seru, banget, plccc, game, ya, tidak, bikin, ...]
game ya banyak dan juga seru	[game, ya, banyak, dan, juga, seru]
keren info mabar pfts	[keren, info, mabar, pfts]

Gambar 5. Hasil Normalisasi dan Tokenisasi

### c. Stopword Removal dan Stemming Data

Tahap *stopword removal* menghapus kata umum yang kurang bermakna, seperti “dan”, “yang”, dan “di”. Selanjutnya, *stemming* dengan Sastrawi mengubah kata ke bentuk dasar, misalnya “bermain” menjadi “main” dan “makanan” menjadi “makan”.

stopword removal	stemming_data
[seru, banget, plccc, game, ya, bikin, cepat, ...]	seru banget plccc game ya bikin cepat bosennn
[game, ya, seru]	game ya seru
[keren, info, mabar, pfts]	keren info mabar pfts

Gambar 6. Hasil Stopword Removal dan Stemming Data

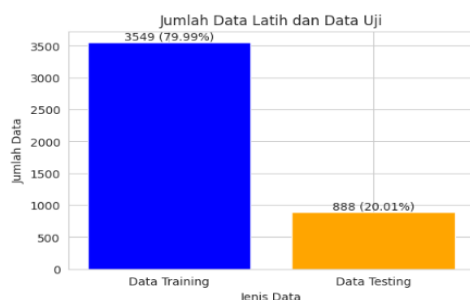
### d. Labelling

Pelabelan sentimen dilakukan dengan pendekatan *lexicon-based* menggunakan InSet Lexicon. Nilai skor sentimen dihitung dengan rumus:

### Sentiment Score

$$= positive_{count} - negative_{count}$$

Jika hasilnya  $> 0$ , ulasan diberi label Positif; jika  $< 0$ , diberi label Negatif; dan jika  $= 0$ , diberi label Netral.



Gambar 7. Data Training dan Testing

### 3. Transformation

Tahap *transformation* mengonversi teks menjadi data numerik dengan panjang 100 token per ulasan, sedangkan label sentimen diubah ke bentuk numerik. Data selanjutnya dibagi menjadi 80% data latih (3.549 sampel) dan 20% data uji (888 sampel).

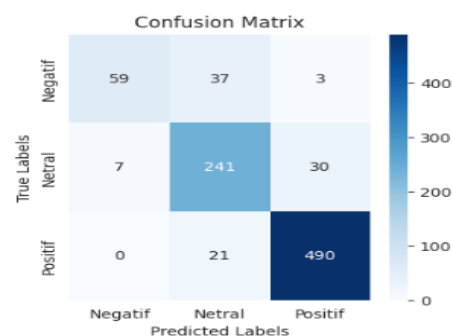
### 4. Data Mining

Tahap *data mining* menerapkan model deep learning untuk klasifikasi sentimen menggunakan algoritma CNN dan LSTM pada data yang telah melalui tahap *preprocessing* dan *transformation*.

#### a. Conventional Neural Network (CNN) Model

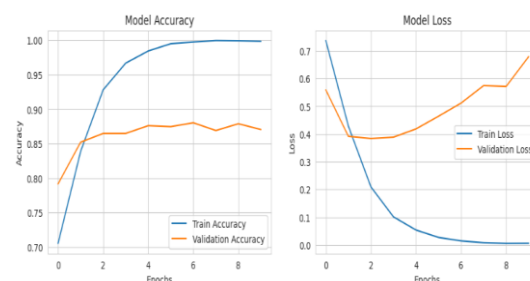
Berdasarkan hasil pengujian yang ditunjukkan pada gambar 8, model memperoleh akurasi sebesar 0.89 (89%), dengan *precision*, *recall*, dan *F1-score* berturut-turut untuk kelas Negatif (0.89, 0.60, 0.72), Netral (0.81, 0.87, 0.84), dan Positif

(0.94, 0.96, 0.95). Hasil ini menunjukkan bahwa CNN paling akurat dalam mendeteksi sentimen positif, sedangkan performa pada kelas negatif lebih rendah akibat jumlah data yang lebih sedikit dan distribusi yang tidak seimbang.



Gambar 8. Confusion Matrix CNN

Gambar 9 memperlihatkan akurasi dan loss model CNN, di mana akurasi data latih meningkat hingga mendekati 100% dan akurasi validasi stabil pada kisaran 86–88%. Penurunan loss pelatihan yang diikuti kenaikan loss validasi setelah epoch ke-3 menunjukkan terjadinya overfitting.

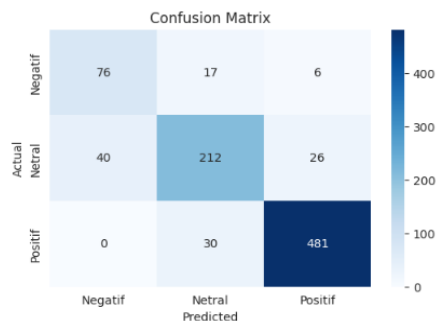


Gambar 9. Grafik Accuracy dan Model Loss

#### b. Long Short-Term Memory (LSTM)

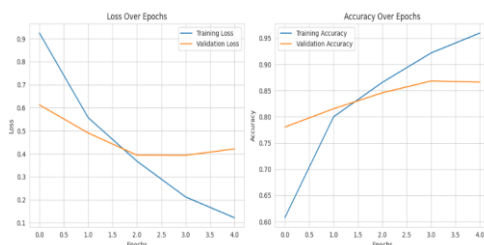
Model LSTM mencapai akurasi 86,60%, dengan kinerja terbaik pada kelas positif (*precision*, *recall*, dan *f1-score* 0,94). Pada gambar 10, kelas netral memperoleh *f1-score* 0,79 dan kelas negatif 0,71.

Nilai macro average sebesar 0,81 dan weighted average 0,87 menunjukkan model cukup baik, meski performa kelas netral dan negatif masih perlu ditingkatkan.



Gambar 10. Confusion Matrix LSTM

Gambar 11 memperlihatkan hasil pelatihan LSTM, di mana loss pelatihan menurun tajam dan loss validasi stabil di kisaran 0,4, menunjukkan tidak terjadi overfitting yang signifikan. Akurasi pelatihan meningkat hingga 0,95 dan akurasi validasi mencapai 0,86, menandakan kemampuan generalisasi model yang baik.



Gambar 11. Grafik Accuracy dan Model Loss LSTM

## 5. Evaluation

Tahap evaluasi membandingkan kinerja CNN dan LSTM dalam analisis sentimen ulasan Roblox. CNN memperoleh akurasi 89%, lebih tinggi dibandingkan LSTM sebesar 86,60%. CNN menunjukkan performa terbaik pada kelas positif

dengan f1-score 0,95 serta lebih stabil pada kelas negatif. Sementara itu, LSTM memiliki akurasi lebih rendah, tetapi unggul dalam kestabilan hasil dan pemahaman konteks urutan kata, khususnya pada kelas netral.

Tabel 1. Perbandingan Kinerja Model CNN dan LSTM

Model	Accuracy (%)	Precision	Recall	F1-Score
CNN	89.00	0.88	0.81	0.83
LSTM	86.60	0.80	0.82	0.81

Hasil ini menunjukkan bahwa CNN lebih unggul dalam mengklasifikasikan sentimen pada ulasan aplikasi Roblox karena mampu mengekstraksi fitur spasial dari teks dengan lebih efektif dibandingkan LSTM yang berfokus pada urutan temporal.

## 4.2. Pembahasan

Hasil penelitian menunjukkan bahwa tahapan pengolahan data, mulai dari pengambilan ulasan hingga evaluasi model, berperan penting dalam analisis sentimen ulasan aplikasi Roblox. Proses *scraping* berhasil mengumpulkan 5.000 ulasan berbahasa Indonesia yang relevan, sedangkan tahap preprocessing seperti cleaning, normalisasi, tokenisasi, stopwords removal, dan stemming mampu meningkatkan kualitas data. Pelabelan sentimen berbasis leksikon memberikan dasar klasifikasi yang sistematis, meskipun masih terdapat ketidakseimbangan distribusi kelas, khususnya pada sentimen negatif. Pada tahap pemodelan, CNN dan LSTM

menunjukkan karakteristik kinerja yang berbeda. CNN memperoleh akurasi tertinggi sebesar 89% dan performa sangat baik pada kelas positif, menandakan kemampuannya dalam mengekstraksi fitur teks secara efektif, meskipun terdapat indikasi overfitting. Sementara itu, LSTM mencapai akurasi 86,60% dengan performa yang lebih stabil dan kemampuan generalisasi yang baik, terutama dalam memahami konteks urutan kata pada kelas netral. Secara keseluruhan, CNN lebih unggul dalam hal akurasi, sedangkan LSTM menawarkan kestabilan dan pemahaman konteks yang lebih baik, sehingga pemilihan model bergantung pada kebutuhan analisis sentimen yang diinginkan

## 5. Kesimpulan

Penelitian ini bertujuan untuk membandingkan kinerja model CNN dan LSTM dalam melakukan klasifikasi sentimen pada ulasan pengguna aplikasi Roblox. Berdasarkan hasil analisis, kedua model mampu mengklasifikasikan sentimen dengan baik, namun CNN menunjukkan performa lebih unggul dibandingkan LSTM dengan tingkat akurasi dan F1-score yang lebih tinggi. Hal ini menunjukkan bahwa CNN lebih efektif dalam mengekstraksi fitur spasial teks dibandingkan LSTM yang berfokus pada dependensi urutan kata. Kontribusi penelitian ini terletak pada penerapan pendekatan lexicon-based sebagai metode pelabelan awal yang diintegrasikan

dengan model deep learning untuk meningkatkan efisiensi proses analisis sentimen. Keterbatasan penelitian ini terletak pada ukuran data yang masih terbatas dan belum diterapkannya teknik optimasi parameter yang lebih kompleks, sehingga penelitian selanjutnya disarankan untuk menggunakan dataset yang lebih besar serta melakukan tuning hyperparameter agar hasil klasifikasi dapat lebih optimal.

## 6. Daftar Pustaka

- [1] J. Lee, E. Im, I. Yeo, and G. Gim, "A Study on the Strategy of SWOT Extraction in the Metavers Platform Review Data: Using NLP Techniques," *Stud. Comput. Intell.*, vol. 1074, pp. 173–188, 2022, doi: 10.1007/978-3-031-19604-1\_13.
- [2] B. Liu, *Sentiment Analysis Mining Opinions, Sentiments, and Emotions*. Chicago: Cambridge University Press, 2020.
- [3] F. S. Pamungkas and I. Kharisudin, "Analisis Sentimen dengan SVM, NAIVE BAYES dan KNN untuk Studi Tanggapan Masyarakat Indonesia Terhadap Pandemi Covid-19 pada Media Sosial Twitter," *Prosiding Seminar Nasional Matematika*, vol. 4, pp. 1–7, 2021.
- [4] R. G. Purnasiwi, Kusriani, and M. Hanafi, "Analisis Sentimen Pada Review Produk Skincare Menggunakan Word Embedding dan Metode Long Short-Term Memory (LSTM)," *Innov. J. Soc. Sci. Res.*, vol. 3, no. 2, pp. 11433–11448, 2023.
- [5] W. Lu, Y. Duan, and Y. Song, "Self-Attention-Based Convolutional Neural Networks for Sentence Classification," in *2020 IEEE 6th International Conference on Computer and Communications, ICC3 2020*, 2020, pp. 2065–2069, doi: 10.1109/ICC351575.2020.9345092.

- [6] A. Lighthart, C. Catal, and B. Tekinerdogan, "Systematic reviews in sentiment analysis: a tertiary study," *Artif. Intell. Rev.*, vol. 54, no. 7, pp. 4997–5053, 2021, doi: 10.1007/s10462-021-09973-3.
- [7] S. F. Pane, J. Ramdan, A. G. Putrada, M. N. Fauzan, R. M. Awangga, and N. Alamsyah, "A Hybrid CNN-LSTM Model with Word-Emoji Embedding for Improving the Twitter Sentiment Analysis on Indonesia's PPKM Policy," *Proceeding - 6th Int. Conf. Inf. Technol. Inf. Syst. Electr. Eng. Appl. Data Sci. Artif. Intell. Technol. Environ. Sustain. ICITISEE 2022*, pp. 51–56, 2022, doi: 10.1109/ICITISEE57756.2022.10057720.
- [8] S. Imron, E. I. Setiawan, J. Santoso, and M. H. Purnomo, "Aspect Based Sentiment Analysis Marketplace Product Reviews Using BERT, LSTM, and CNN," *Jurnal RESTI*, vol. 7, no. 3, academia.edu, pp. 586–591, 2023, doi: 10.29207/resti.v7i3.4751.
- [9] W. Li, L. Zhu, Y. Shi, K. Guo, and E. Cambria, "User reviews: Sentiment analysis using lexicon integrated two-channel CNN-LSTM family models," *Appl. Soft Comput. J.*, vol. 94, 2020, doi: 10.1016/j.asoc.2020.106435.
- [10] Lady Agustine Fitrana, S. Linawati, N. Herlinawati, R. Sa'adah, and S. Seimahuria, "Analisis Sentimen Pengguna Twitter Terhadap Brand Indosat Menggunakan Metode Naïve Bayes Classifier," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 3, pp. 4291–4297, 2024, doi: 10.36040/jati.v8i3.9866.
- [11] A. Muzaki and A. Witanti, "Sentiment Analysis of the Community in the Twitter To the 2020 Election in Pandemic Covid-19 By Method Naive Bayes Classifier," *J. Tek. Inform.*, vol. 2, no. 2, pp. 101–107, 2021, doi: 10.20884/1.jutif.2021.2.2.51.
- [12] K. M. Hasib, U. Naseem, A. J. Keya, S. Maitra, K. Mithu, and M. G. R. Alam, "Systematic Literature Review on Sentiment Analysis in Airline Industry," *SN Comput. Sci.*, vol. 6, no. 1, 2025, doi: 10.1007/s42979-024-03567-w.
- [13] W. Jia, W. Wang, and Z. Zhang, "From simple digital twin to complex digital twin Part I: A novel modeling method for multi-scale and multi-scenario digital twin," *Adv. Eng. Informatics*, vol. 53, 2022, doi: 10.1016/j.aei.2022.101706.
- [14] A. H. Ombabi, W. Ouarda, and A. M. Alimi, "Deep learning CNN-LSTM framework for Arabic sentiment analysis using textual information shared in social networks," *Soc. Netw. Anal. Min.*, vol. 10, no. 1, 2020, doi: 10.1007/s13278-020-00668-1.
- [15] N. Hossain, M. R. Bhuiyan, Z. N. Tumpa, and S. A. Hossain, "Sentiment Analysis of Restaurant Reviews using Combined CNN-LSTM," *2020 11th Int. Conf. Comput. Commun. Netw. Technol. ICCCNT 2020*, 2020, doi: 10.1109/ICCCNT49239.2020.9225328.
- [16] W. Lu *et al.*, "Application of Entity-BERT model based on neuroscience and brain-like cognition in electronic medical record entity recognition," *Front. Neurosci.*, vol. 17, 2023, doi: 10.3389/fnins.2023.1259652.
- [17] K. K. Mohbey, G. Meena, S. Kumar, and K. Lokesh, "A CNN-LSTM-Based Hybrid Deep Learning Approach for Sentiment Analysis on Monkeypox Tweets," *New Gener. Comput.*, vol. 42, no. 1, pp. 89–107, 2024, doi: 10.1007/s00354-023-00227-0.
- [18] M. M. F. Fahima and Prasanna Sumathipala, "Sentiment Analysis and its Application in the Education Sector: A Comprehensive Review," *Elem. J. Educ. Res.*, vol. 3, no. 1, pp. 105–131, 2025, doi: 10.61166/elm.v3i1.89.
- [19] L. Yang, Y. Li, J. Wang, and R. S. Sherratt, "Sentiment Analysis for E-Commerce Product Reviews in Chinese Based on Sentiment Lexicon and Deep Learning," *IEEE Access*, vol. 8, pp. 23522–23530, 2020, doi: 10.1109/ACCESS.2020.2969854..