

Analisis Sentimen Komentar YouTube pada Program *Clash of Champions* Ruangguru Menggunakan *Deep Learning* Berbasis LSTM

Iska Rahmawati^{1,*}, Nur Aini²

^{1,2}Universitas Amikom Yogyakarta, Yogyakarta, Indonesia

Informasi Artikel

Sejarah Artikel:
Submit : 16 Mei 2025
Revisi : 26 Juni 2025
Diterima : 29 Juni 2025
Diterbitkan: 30 Juni 2025

Kata Kunci

Analisis sentimen, LSTM, YouTube, komentar pengguna, Ruangguru, *Clash of Champions*

Correspondence

E-mail: biskarahma@students.amikom.ac.id*

A B S T R A K

Pertumbuhan platform digital seperti YouTube telah membuka ruang interaksi yang masif antar pengguna, khususnya melalui kolom komentar pada konten edukatif. Penelitian ini bertujuan untuk menganalisis sentimen publik terhadap program "*Clash of Champions*" dari Ruangguru dengan memanfaatkan komentar pengguna di YouTube dan menggunakan algoritma Long Short-Term Memory (LSTM). Sebanyak 43.664 komentar dikumpulkan selama 46 hari menggunakan YouTube API, kemudian diproses melalui tahap *data preprocessing* (*text cleaning, tokenizing, stemming*), pelabelan sentimen secara manual, transformasi fitur menggunakan TF-IDF, dan pelatihan model LSTM yang telah *define-tuning* dengan data relevan berbahasa Indonesia informal. Hasil klasifikasi menunjukkan bahwa 72,6% komentar bersentimen positif dan 27,4% bersentimen negatif. Evaluasi model menggunakan metrik *accuracy, precision, recall, dan F1-score* dengan nilai mencapai 90%, menunjukkan bahwa model memiliki kemampuan klasifikasi yang stabil dan andal. Temuan ini menunjukkan efektivitas LSTM dalam menangkap pola sentimen pada data teks yang bersifat informal dan lokal, serta memberikan kontribusi dengan menyajikan pendekatan berbasis *deep learning* untuk analisis opini publik terhadap konten edukatif digital di Indonesia.

Abstract

The growth of digital platforms such as YouTube has opened massive interactive spaces among users, particularly through comment sections on educational content. This study aims to analyze public sentiment toward Ruangguru's "*Clash of Champions*" program by utilizing user comments on YouTube and applying the Long Short-Term Memory (LSTM) algorithm. A total of 43,664 comments were collected over 46 days using the YouTube API, then processed through several stages: *data preprocessing* (*text cleaning, tokenizing, stemming*), manual sentiment labeling, feature transformation using TF-IDF, and LSTM model training that was fine-tuned with informal Indonesian language data. The classification results show that 72.6% of the comments carry positive sentiment, while 27.4% are negative. Model evaluation using *accuracy, precision, recall, and F1-score* yielded values of 90%, indicating a stable and reliable classification performance. These findings demonstrate the effectiveness of LSTM in capturing sentiment patterns within informal and local language texts and contribute by presenting a deep learning-based approach to analyze public opinion toward educational digital content in Indonesia.

This is an open access article under the CC-BY-SA license



1. Pendahuluan

Di era digital yang semakin berkembang, platform berbagi video seperti YouTube telah mengalami transformasi signifikan menjadi media pembelajaran yang efektif. Salah satu program edukatif yang memanfaatkan platform ini adalah *Clash of Champions* oleh Ruangguru, yang menggabungkan elemen kompetisi dan pembelajaran interaktif sebagai strategi untuk meningkatkan keterlibatan siswa dalam kegiatan belajar daring. Tingginya interaksi pengguna dalam bentuk komentar pada video-video

program ini menunjukkan adanya partisipasi aktif, namun persepsi publik terhadap konten tersebut belum banyak dianalisis secara komprehensif. Hal ini penting untuk dipahami, terutama dalam konteks efektivitas penyampaian materi pendidikan melalui media sosial. Penelitian Widodo & Safitri [1] menegaskan bahwa *platform* digital seperti YouTube dapat dimanfaatkan secara efektif dalam konteks pembelajaran daring, khususnya melalui pendekatan *deep learning* untuk memahami persepsi pengguna.

Long Short-Term Memory (LSTM) unggul dibanding metode tradisional seperti *Naïve Bayes* dan *Support Vector Machine (SVM)* karena kemampuannya dalam menangani dependensi jangka panjang dalam data teks, serta mempertahankan konteks dalam urutan kata. Hal ini menjadikannya sangat efektif untuk memodelkan bahasa alami, terutama pada data tidak terstruktur dan informal seperti komentar media sosial [2]. Huda [3] mencatat bahwa *LSTM* menunjukkan akurasi tinggi dalam tugas analisis sentimen dan dinilai lebih adaptif dibandingkan metode klasik. Handayani, Pratiwi, & Dairoh [4] memperkuat temuan ini melalui eksperimen pada ulasan *Twitter* berbahasa informal. Dalam lanskap *Natural Language Processing (NLP)* saat ini, *LSTM* masih menjadi fondasi penting dalam pengembangan arsitektur *deep learning* karena kemampuannya menyimpan memori jangka panjang, meskipun model berbasis *Transformer* seperti *BERT* dan *GPT* mulai mendominasi berbagai tugas pemrosesan bahasa alami. Namun, untuk skenario analisis berskala sedang yang melibatkan bahasa informal dan keterbatasan sumber daya komputasi, *LSTM* tetap menjadi pilihan yang efisien dan kompetitif [5]. Di sisi lain, meskipun analisis sentimen telah banyak diaplikasikan di platform seperti *Twitter* dan *Google Play Store*, penggunaannya secara langsung pada komentar *YouTube* dalam konteks edukatif – seperti program *Clash of Champions* dari Ruangguru belum banyak dieksplorasi [6]. Suryani & Kurniawan [7] misalnya, menunjukkan efektivitas kombinasi *LSTM* dengan *Word2Vec* dalam menangkap sentimen komentar pada konten edukasi digital. Hal ini menunjukkan adanya celah penelitian yang relevan untuk dikaji lebih dalam, khususnya dalam memahami opini publik terhadap konten pendidikan digital yang berbasis hiburan.

Memahami sentimen publik terhadap program edukasi digital sangat penting baik dari perspektif sosial maupun akademik. Dari sisi sosial, persepsi positif pengguna dapat mendorong *user engagement* yang lebih tinggi, meningkatkan partisipasi, dan memperluas akses pembelajaran yang inklusif melalui media digital. Sementara itu, secara akademik, analisis sentimen publik menjadi indikator awal terhadap efektivitas penyampaian materi, keberterimaan model pembelajaran, dan kualitas interaksi dalam ekosistem edukasi daring. Sentimen negatif yang tidak ditangani dengan tepat dapat berdampak pada rendahnya *user retention* serta menurunnya kepercayaan terhadap platform pendidikan itu sendiri. Oleh karena itu, menganalisis opini publik berbasis komentar di media sosial seperti *YouTube* tidak hanya memberikan umpan balik *real-time*, tetapi juga membantu pengembang konten dan peneliti pendidikan dalam menyempurnakan strategi penyampaian dan pendekatan pedagogis berbasis teknologi. Kesenjangan ini memperlihatkan bahwa belum ada penelitian yang secara eksplisit menganalisis persepsi publik terhadap program edukasi berbasis video menggunakan pendekatan algoritmik berbasis *deep learning* di platform *YouTube*. Di sinilah letak *novelty* dari penelitian ini: mengintegrasikan analisis sentimen berbasis *LSTM* terhadap komentar *YouTube* sebagai indikator penerimaan publik terhadap program *Clash of Champions*, yang belum banyak dieksplorasi pada kajian terdahulu.

Oleh karena itu, penelitian ini bertujuan untuk menjawab pertanyaan utama: Bagaimana sentimen publik terhadap program *Clash of Champions* yang ditayangkan Ruangguru di *YouTube*, berdasarkan klasifikasi sentimen positif dan negatif menggunakan metode *Long Short-Term Memory (LSTM)*? Dengan menjawab pertanyaan ini, hasil penelitian diharapkan dapat memberikan evaluasi yang objektif terhadap efektivitas konten edukatif yang disampaikan, sekaligus memberikan masukan untuk pengembangan program pendidikan digital di masa depan. Program *Clash of Champions* merupakan bagian dari strategi Ruangguru untuk mengemas pembelajaran dalam bentuk kompetisi antar siswa

dari berbagai daerah di Indonesia, dan disajikan secara interaktif melalui media visual. Selama masa penayangan yang berlangsung lebih dari satu bulan, program ini dirilis dalam 11 episode dan berhasil menarik lebih dari 2,2 juta penayangan, dengan puluhan ribu komentar yang menunjukkan tingginya partisipasi pengguna.

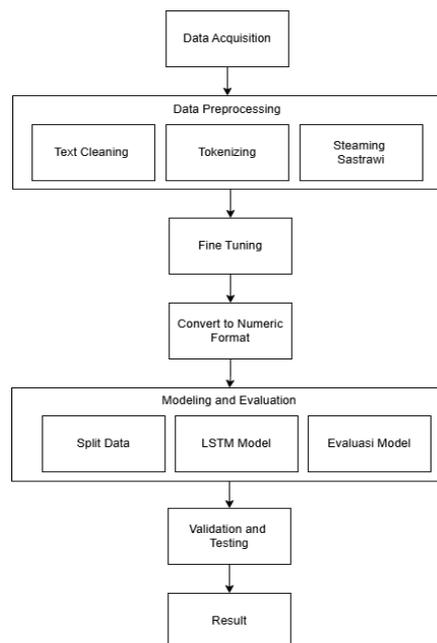
Komentar yang dianalisis dalam penelitian ini dikumpulkan menggunakan *YouTube API* selama 46 hari, dengan total 43.664 komentar dari 11 video yang termasuk dalam kategori edukasi. Komentar yang diambil dibatasi pada bahasa Indonesia informal, sesuai dengan gaya khas media sosial. Data yang digunakan mencakup komentar yang dipublikasikan selama periode Juli hingga Agustus 2024, dan mencerminkan opini publik secara *real-time* terhadap setiap episode yang ditayangkan. Data komentar yang dianalisis adalah komentar yang mengandung ekspresi persepsi, opini, kritik, dan pujian yang relevan terhadap jalannya program [8]. Selain klasifikasi sentimen, penelitian ini juga mengevaluasi performa algoritma *LSTM* melalui sejumlah metrik pengukuran seperti *accuracy*, *precision*, *recall*, dan *F1-score*, guna memastikan konsistensi dan ketepatan model dalam mengenali pola sentimen pada data teks tidak terstruktur dan bersifat lokal.

2. Metodologi Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimental berbasis data tekstual dari media sosial. Langkah-langkah penelitian dimulai dari akuisisi data, pra-pemrosesan, pelabelan sentimen, pelatihan dan pengujian model *Long Short-Term Memory (LSTM)*, hingga evaluasi performa model. Setiap tahap disusun secara sistematis untuk memastikan validitas dan reliabilitas hasil penelitian. Sejumlah penelitian terdahulu juga menggunakan pendekatan serupa dalam pemrosesan komentar media sosial, seperti pemanfaatan arsitektur *LSTM* oleh Pratama [9] dan Abd Karim [10]. Selain itu, model *Bidirectional LSTM (BiLSTM)* juga telah digunakan secara luas untuk menangkap hubungan sekuensial yang lebih kompleks [11].

2.1. Alur Penelitian

Alur penelitian disusun untuk menggambarkan tahapan sistematis dalam proses analisis sentimen terhadap video Clash of Champions pada kanal YouTube Ruangguru dengan menggunakan *algoritma Long Short-Term Memory (LSTM)*. Rangkaian proses dalam penelitian ini dirancang agar hasil analisis dapat direplikasi secara ilmiah oleh peneliti lain. Gambar 1 menunjukkan urutan dari setiap tahap yang dilaksanakan, dimulai dari pengumpulan data hingga analisis akhir.



Gambar 1. Diagram Alir Metode Penelitian

2.2. Data Acquisition

Proses diawali dengan autentikasi melalui *Google Developer Console* untuk mendapatkan *API key* sebagai otorisasi. Video target diidentifikasi menggunakan video ID, yang menjadi parameter dalam permintaan data melalui HTTP ke endpoint API. Data yang diperoleh dalam format JSON mencakup informasi penting seperti teks komentar, nama pengguna, jumlah likes, serta waktu publikasi komentar. Selanjutnya, data tersebut dikonversi menjadi format CSV menggunakan skrip Python agar memudahkan proses analisis lanjutan. Proses ini tidak hanya efisien dalam menangani volume data besar, tetapi juga memungkinkan replikasi secara ilmiah oleh peneliti lain tanpa intervensi manual.

2.3. Data Preprocessing

Pra-pemrosesan dilakukan untuk memastikan bahwa data komentar berada dalam kondisi optimal sebelum digunakan dalam pelatihan model. Tahapan ini mencakup mengurangi *noise*, menyederhanakan struktur kalimat, dan mempersiapkan data agar dapat diproses oleh *algoritma Long Short-Term Memory (LSTM)*. Langkah awal adalah *text cleaning*, yang mencakup *case folding* untuk mengubah seluruh karakter menjadi huruf kecil, dan *stopword removal* menggunakan pustaka *Sastrawi*. Penghapusan kata-kata umum yang tidak membawa makna signifikan ini membantu mengurangi kompleksitas data serta meminimalkan risiko *overfitting*.

Selanjutnya, dilakukan *tokenization* menggunakan *Keras Tokenizer*, yang memecah kalimat menjadi unit kata (token) dan mengubahnya menjadi urutan angka sesuai indeks kata. Proses ini memanfaatkan fungsi *fit_on_texts()* dan *texts_to_sequences()* untuk menghasilkan *input numerik* yang kompatibel dengan model LSTM. Terakhir, tahap *stemming* dilakukan menggunakan *Sastrawi Stemmer* untuk mengubah kata berimbuhan menjadi bentuk dasar, seperti “mengajarkan” menjadi “ajar”. Dengan representasi teks yang lebih sederhana dan konsisten, model dapat mengenali pola sentimen secara lebih efisien dan akurat.

2.4. Fine Tuning

Tahap ini bertujuan untuk meningkatkan akurasi pemodelan dengan memanfaatkan kemampuan *Large Language Model (LLM) Gemma2B* dalam memahami komentar YouTube yang berbahasa informal. *Gemma2B* tidak digunakan sebagai model klasifikasi utama, melainkan dimanfaatkan untuk **augmentasi data** dalam proses pelabelan. Model ini difine-tuning menggunakan *dataset* komentar yang mengandung bahasa tidak baku, slang, dan ekspresi khas media sosial. Setelah proses *fine-tuning* selama beberapa epoch, *Gemma2B* digunakan untuk memberikan label awal terhadap sebagian komentar yang belum diberi anotasi manual, guna mempercepat proses pelabelan dalam skema semi-supervised. Dengan cara ini, model *Gemma2B* berfungsi sebagai alat bantu dalam memperkaya dataset latihan yang digunakan oleh model utama, yaitu *Long Short-Term Memory (LSTM)*. Setelah data diperkuat dan dilabeli hasil prediksi *Gemma2B* (dengan validasi manual pada sebagian sampel), model LSTM dilatih untuk melakukan klasifikasi sentimen. Studi Tatulus & Wulandhari [12] menunjukkan bahwa pendekatan *deep learning* juga dapat diterapkan secara efektif pada isu politik besar, seperti Pemilu Indonesia 2024, sehingga memperkuat relevansi penggunaan metode ini pada domain sosial dan pendidikan.

2.5. Modeling and Evaluation

Model *LSTM* dilatih menggunakan data latihan selama beberapa *epoch* dengan pembaruan bobot secara bertahap untuk mengenali pola-pola dalam data teks. Proses pelatihan bertujuan agar model mampu memahami hubungan antar-kata dalam komentar dan menghasilkan klasifikasi sentimen yang akurat. Arsitektur *LSTM* yang digunakan memanfaatkan *memory cell* dan tiga gerbang utama: *forget gate*, *input gate*, dan *output gate*, yang berperan dalam menyaring dan meneruskan informasi secara selektif. Studi komparatif menunjukkan bahwa *BiLSTM* mampu memberikan hasil yang lebih baik

dalam memahami konteks dibandingkan *LSTM* standar [13], sedangkan pendekatan *hybrid CNN-LSTM* dapat memperkuat representasi fitur [14].

Data dibagi menjadi dua bagian, yaitu untuk pelatihan dan untuk pengujian. Data uji tidak dilibatkan dalam proses pelatihan dan hanya digunakan untuk menilai kemampuan model dalam menggeneralisasi terhadap data baru. Evaluasi model dilakukan dengan membandingkan hasil prediksi terhadap label asli pada data uji menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Metrik ini memberikan gambaran objektif mengenai performa model dalam klasifikasi sentimen serta mendeteksi kemungkinan *overfitting*. Validasi dilakukan menggunakan *validation set* dari data latih, dan diterapkan teknik *early stopping* untuk menghentikan pelatihan saat kinerja pada data validasi tidak lagi meningkat. Setelah validasi, model diuji pada data uji final untuk menilai performa nyata. Model yang baik menunjukkan kinerja stabil pada data uji dan tidak hanya unggul pada data latih, sehingga dapat diterapkan secara efektif dalam skenario analisis sentimen di dunia nyata.

2.6. Validation and Testing

Selain data uji, digunakan juga *validation set* dari sebagian data latih untuk memantau performa model selama pelatihan. Teknik *early stopping* diterapkan untuk menghentikan pelatihan ketika kinerja pada data validasi tidak lagi meningkat, menjaga keseimbangan antara pembelajaran dan generalisasi. Uji akhir dilakukan menggunakan data uji yang sepenuhnya terpisah, dan hasil prediksi dibandingkan dengan label aktual untuk memastikan bahwa model siap digunakan dalam konteks dunia nyata.

3. Hasil dan Pembahasan

Bagian ini menyajikan hasil penelitian yang diperoleh dari proses analisis sentimen terhadap komentar pengguna *YouTube* terkait program *Clash of Champions* yang diselenggarakan oleh Ruangguru. Setiap temuan dijelaskan secara singkat, didukung oleh visualisasi, tabel, maupun grafik yang relevan, tanpa mengulang isi yang sama secara berlebihan. Pembahasan ini difokuskan pada identifikasi pola sentimen, distribusi kata, serta performa model dalam mengklasifikasikan komentar. Temuan penelitian ini sejalan dengan hasil Giovanni et al. [15] yang menganalisis sentimen publik terhadap aplikasi Ruangguru menggunakan data *Twitter*.

3.1. Data Acquisition

Data dikumpulkan dari 11 episode video *Clash of Champions* yang ditayangkan di kanal *YouTube* resmi Ruangguru. Pengambilan data dilakukan secara otomatis menggunakan **YouTube Data API**, yang memungkinkan ekstraksi komentar secara terstruktur. Total data yang terkumpul berjumlah **43.664 baris**, di mana setiap baris merepresentasikan satu komentar pengguna lengkap dengan informasi seperti nama pengguna, waktu publikasi, dan ID video.

Informasi yang diperoleh meliputi beberapa atribut yang berisi konten teks komentar yang mencerminkan opini, reaksi, atau tanggapan pengguna terhadap video tersebut. Kolom ini menjadi dasar untuk analisis sentimen dalam penelitian ini. Contoh struktur data yang dikumpulkan disajikan pada Tabel 1.

Tabel 1. Tabel Data Acquisition

<i>authorDisplayName</i>	<i>textDisplay</i>	<i>video_id</i>	<i>publishedAt</i>
@Ruangguru	Yuk klaim diskon di http://ruangguru.com sekarang! 😊	dY_3acmFFRQ	2024-07-03T06:39:10Z
@ONIC_talita	coc udah mau selesai 😞 😞 😞	dY_3acmFFRQ	2024-08-17T06:25:55Z
@rh.amharr21	tim yang <i>rewatch</i> sebelum last eps	dY_3acmFFRQ	2024-08-

	besok 🗓️		16T13:57:49Z
@aisyahrahel	KIRAINN MAXWEL GK LULUS HAMPIR AJA NANGES 🤔 🤔	IDAIikm3T_o	2024-08- 16T13:16:22Z
@user-us2sr6bo7l	"baru nonton dan ternyata seru hahaha sebenarnya bisa lebih baik"	IDAIikm3T_o	2024-08- 16T11:58:53Z

3.2. Data Preprocessing

Hasil dari ketiga proses ini dirangkum dalam Tabel 2, yang menunjukkan transformasi teks asli menjadi bentuk yang lebih bersih dan siap digunakan dalam pelatihan model. Kolom *clean_text* mewakili hasil *text cleaning*, *tokenized_text* menunjukkan hasil tokenisasi, dan *stemmed_text* merupakan hasil *stemming* dari token yang telah dibersihkan.

Tabel 2. Tabel Hasil Pra-pemrosesan Komentar YouTube

<i>textDisplay</i>	<i>clean_text</i>	<i>tokenized_text</i>	<i>stemmed_text</i>
Yuk, konsultasi dan klaim diskonnya sekarang d...	yuk konsultasi dan klaim diskonnya sekarang di	[380, 6020, 9, 6021, 4751, 287, 6]	['yuk', 'konsultasi', 'dan', 'klaim', 'diskon', 'sekarang', 'di']
coc udah mau selesai 🤔 🤔	coc udah mau selesai	[135, 38, 71, 359]	['coc', 'udah', 'mau', 'selesai']
tim yang <i>rewatch</i> sebelum last eps besok 🗓️	tim yang <i>rewatch</i> sebelum <i>last</i> eps besok	[19, 7, 1266, 939, 2252, 48, 557]	['tim', 'yang', 'rewatch', 'belum', 'last', 'eps', 'besok']
KIRAINN MAXWEL GK LULUS HAMPIR AJA NANGES 🤔 🤔	kirainn maxwel gk lulus hampir aja nanges	[13000, 326, 96, 517, 1073, 31, 6022]	['kirainn', 'maxwel', 'gk', 'lulus', 'hampir', 'aja', 'nanges']

3.3. Fine Tuning

Tahap ini berfokus pada proses **pelabelan sentimen** terhadap komentar YouTube, dengan klasifikasi menjadi dua kategori, yaitu **positif** dan **negatif**. Pelabelan dilakukan setelah proses *fine-tuning* menggunakan model **Gemma2B**, yang telah disesuaikan dengan karakteristik bahasa informal pada platform YouTube. Tujuannya adalah untuk meningkatkan sensitivitas model LSTM dalam mengenali pola-pola ekspresi yang tidak baku namun umum ditemukan pada komentar digital.

Data hasil pelabelan disimpan dalam **50 file CSV**, dengan dua kolom utama: *sentiment_label*, yang merepresentasikan jenis sentimen, dan *sentiment_score*, yang menunjukkan skor numeriknya. Label *positive* diberi nilai **1**, dan label *negative* diberi nilai **-1**. Ringkasan hasil pelabelan ditampilkan dalam Tabel 3.

Tabel 3. Tabel Hasil *Fine Tuning*

<i>sentiment_label</i>	<i>sentiment_score</i>
<i>Negative</i>	-1
<i>Positive</i>	1
<i>Positive</i>	1
<i>Negative</i>	-1
<i>Positive</i>	1

3.4. Modeling and Evaluation

Model dilatih dengan pembagian **80% untuk data latih** dan **20% untuk data uji**, dengan pengaturan `random_state = 42` untuk menjaga konsistensi hasil. Data latih digunakan untuk proses pembelajaran model agar mengenali pola-pola sentimen dalam komentar, sedangkan data uji digunakan untuk mengukur kemampuan model dalam menggeneralisasi informasi terhadap data yang belum pernah dikenali sebelumnya.

Model *Long Short-Term Memory (LSTM)* dibangun dengan dua lapisan utama dan dilatih selama **12 epoch**. Tujuan utama dari pelatihan ini adalah untuk menangkap pola sekuensial dan relasi antar-kata yang muncul dalam komentar YouTube. Selama pelatihan, bobot model diperbarui secara bertahap agar akurasi semakin meningkat pada tiap iterasi. Arsitektur LSTM dirancang dengan unit *memory cell* dan gerbang-gerbang utama untuk mempertahankan informasi penting serta menyaring yang tidak relevan. Struktur arsitektur model digambarkan pada Gambar 2.

```

Model: "sequential_1"
-----
Layer (type)                Output Shape          Param #
-----
embedding_1 (Embedding)     (None, 100, 200)     4000000
lstm_1 (LSTM)                (None, 100, 64)      67840
dropout_1 (Dropout)         (None, 100, 64)      0
lstm_2 (LSTM)                (None, 100, 64)      33024
dropout_2 (Dropout)         (None, 100, 64)      0
global_max_pooling1d_1 (Glo (None, 64)            0
balMaxPooling1D)
batch_normalization_1 (Batc (None, 64)            256
hNormalization)
dense_2 (Dense)              (None, 64)            4160
dense_3 (Dense)              (None, 1)              65
-----
Total params: 4,105,345
Trainable params: 4,105,217
Non-trainable params: 128
-----
Epoch 1/100
    
```

Gambar 2. Struktur Model LSTM

Evaluasi dilakukan dengan menggunakan data uji untuk membandingkan hasil prediksi terhadap label sebenarnya. Ringkasan hasil evaluasi dapat dilihat pada Tabel 4. Hasil akurasi keseluruhan sebesar **90%** menunjukkan bahwa model bekerja dengan performa yang konsisten dan seimbang antara kedua kelas.

Tabel 4. Tabel Hasil Evaluasi Kinerja Model LSTM

Label	Precision	Recall	F1-Score	Support
Negatif	0.90	0.89	0.89	6328
Positif	0.89	0.91	0.90	6355
Accuracy	-	-	0.90	12683
Macro Avg	0.90	0.90	0.90	12683
Weighted Avg	0.90	0.90	0.90	12683

3.5. Validation and Testing

Pada tahap ini model diuji lebih lanjut untuk mengukur performanya dalam situasi yang lebih nyata. Data validasi digunakan untuk memastikan bahwa model tidak mengalami *overfitting* dan *underfitting*, sementara hasil pengujian akhir pada data uji menggambarkan kinerja model secara keseluruhan yang digambarkan pada Gambar 3. Dalam konteks *big data*, NiKomang [16] juga menunjukkan bahwa teknologi analisis sentimen dapat diaplikasikan pada opini publik tokoh nasional. Studi menggunakan metode klasik seperti *Support Vector Machine (SVM)* [17] maupun topik spesifik

seperti konsumsi makanan organik [18] juga membuktikan pentingnya analisis sentimen dalam memahami perilaku masyarakat digital.

```
Teks: 'aku tidak suka sama maxwel'
1/1 [=====] - 0s 40ms/step
Prediksi sentimen: Negatif

Teks: 'aku g suka sama maxwel'
1/1 [=====] - 0s 18ms/step
Prediksi sentimen: Negatif

Teks: 'Film ini luar biasa bagus'
1/1 [=====] - 0s 17ms/step
Prediksi sentimen: Positif

Teks: 'Saya tidak suka dengan endingnya'
1/1 [=====] - 0s 17ms/step
Prediksi sentimen: Negatif

Teks: 'Filmnya cukup menghibur tapi tidak istimewa'
1/1 [=====] - 0s 20ms/step
Prediksi sentimen: Negatif
```

Gambar 3. Validation dan Testing dengan Data Nyata

3.6. Result

Bagian ini memaparkan hasil analisis sentimen terhadap komentar pengguna YouTube terhadap program *Clash of Champions* dari Ruangguru. Tujuannya adalah untuk memahami persepsi publik melalui visualisasi word cloud, distribusi kata, serta bigram yang dapat mengungkap pola bahasa pada komentar positif dan negatif.

Visualisasi *word cloud* digambarkan pada gambar 4. digunakan untuk menampilkan kata-kata yang paling sering muncul dalam masing-masing kategori sentimen. Pada komentar **negatif**, kata-kata seperti *gak*, *liat*, *sedih*, dan *kalah* dominan, serta munculnya nama *Maxwell* mengindikasikan adanya ketidakpuasan atau kontroversi seputar peserta tersebut. Sebaliknya, pada komentar **positif**, kata-kata seperti *keren*, *seru*, *nonton*, dan *semangat* menunjukkan apresiasi terhadap program.



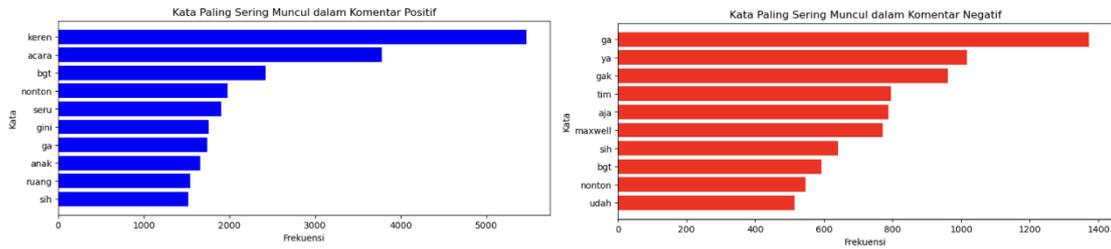
Gambar 4. Wordcloud Komentar Positif dan Negatif

Analisis distribusi frekuensi kata digambarkan pada gambar 5 memperkuat hasil word cloud sebelumnya. Pada komentar positif, kata *keren*, *acara*, dan *seru* muncul dengan frekuensi tinggi, mencerminkan penerimaan positif dari penonton. Di sisi lain, kata *gak* dan *ya* sering muncul pada komentar negatif, disertai dengan nama *Maxwell*, yang menandakan adanya ketidakpuasan terhadap peserta tertentu atau sistem kompetisi.

Kata Paling Sering Muncul dalam Komentar Positif			Kata Paling Sering Muncul dalam Komentar Negatif		
Rank	Kata	Frekuensi	Rank	Kata	Frekuensi
0	keren	5469	0	ga	1373
1	acara	3777	1	ya	1017
2	bgt	2425	2	gak	962
3	nonton	1978	3	tim	795
4	seru	1901	4	aja	788
5	gini	1758	5	maxwell	772
6	ga	1740	6	sih	642
7	anak	1656	7	bgt	593
8	ruang	1542	8	nonton	546
9	sih	1519	9	udah	514

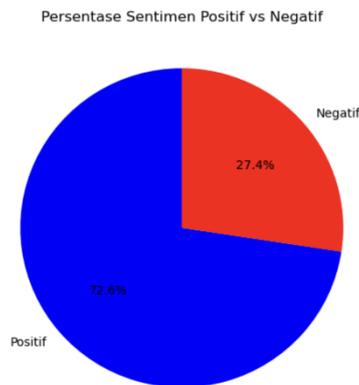
Gambar 5. Bar Chart – Kata Positif dan Negatif Terbanyak

Bigram dianalisis untuk mengidentifikasi frasa yang sering muncul secara berpasangan. Pada komentar positif, frasa seperti *keren banget*, *semangat terus*, dan *bangga tim* menggambarkan dukungan terhadap peserta, digambarkan pada gambar 6. Sementara itu, frasa negatif seperti *gak pintar*, *kenapa kalah*, dan *gini aja* menunjukkan adanya kritik terhadap sistem dan hasil kompetisi.



Gambar 6. Bigram Komentar Positif dan Negatif

Dari total **43.664 komentar**, proporsi sentimen **positif** tercatat sebesar **72,6%**, sementara **27,4%** tergolong sebagai sentimen **negative** yang digambarkan pada gambar 7. Komentar positif dominan mencerminkan dukungan terhadap format dan hiburan program, sedangkan komentar negatif banyak mengkritisi sistem kompetisi atau menyuarakan ketidakpuasan terhadap peserta seperti *Maxwell*, yang muncul di kedua kategori.



Gambar 7. Diagram Pie – Proporsi Sentimen Positif vs. Negatif

3.7. Perbandingan dengan Penelitian Sebelumnya

Hasil akurasi model *LSTM* sebesar 90% pada data komentar *YouTube* menunjukkan performa yang sangat kompetitif jika dibandingkan dengan penelitian terdahulu. Hudaya [3] dalam studi literatur menyatakan bahwa *LSTM* menunjukkan rata-rata akurasi sebesar 89% dalam berbagai studi sentimen berbasis teks, khususnya dalam konteks media sosial berbahasa Indonesia. Handayani, Pratiwi, & Dairoh [4] juga mencatat performa *LSTM* yang tinggi dalam klasifikasi sentimen terhadap ulasan *Twitter* dengan bahasa informal. Dibandingkan dengan pendekatan berbasis *Support Vector Machine (SVM)* seperti pada studi Kusumawardana [8], yang menghasilkan akurasi sebesar 83–85% pada ulasan aplikasi, pendekatan *LSTM* dalam penelitian ini menunjukkan peningkatan performa yang signifikan, khususnya dalam menangkap konteks kata dan dependensi antar *token* dalam kalimat panjang. Dengan demikian, hasil 90% akurasi menunjukkan bahwa model ini tidak hanya stabil, tetapi juga unggul dibandingkan metode konvensional dan sejalan dengan *benchmark deep learning* lain dalam studi serupa.

3.8. Analisis Kesalahan dan Keterbatasan Model

Meskipun model *LSTM* menunjukkan akurasi yang tinggi secara keseluruhan, masih terdapat beberapa kesalahan klasifikasi, terutama pada komentar yang mengandung ambiguitas, sarkasme, atau gabungan sentimen positif dan negatif. Berikut adalah beberapa contoh nyata dari komentar yang diklasifikasikan secara keliru oleh model yang digambarkan pada Table 5.

Tabel 5. Tabel Analisis Kesalahan

Komentar	Label Sebenarnya	Prediksi Model	Analisis Kesalahan
Wah keren banget acaranya, tapi nilainya sih zonk parah 😏	Negatif	Positif	Mengandung pujian di awal tetapi kritik tajam di akhir. Model lebih menangkap kata "keren" sebagai sinyal positif dan mengabaikan konteks sarkastik di belakang.
Maxwell tuh hebat, tapi sistemnya yang aneh	Negatif	Positif	Kalimat ini menyampaikan ketidakpuasan terhadap sistem, bukan terhadap peserta. Model kesulitan mengenali kontras opini karena tidak adanya pemisahan entitas dalam satu kalimat.
Hadehh... yaudahlah 😏	Negatif	Positif	Kalimat pendek, ekspresif, dan tanpa kata eksplisit sentimen. Model tidak dapat membaca nada emosional atau ironi dari emoji.

Kesalahan-kesalahan ini menunjukkan bahwa model *LSTM* masih memiliki keterbatasan dalam menangkap nuansa emosional, sarkasme, dan konflik opini dalam satu kalimat. Selain itu, komentar yang sangat pendek atau dominan berisi ekspresi non-verbal seperti *emoji* juga menjadi tantangan tersendiri. Hal ini sejalan dengan penelitian Liawati et al. [19] yang menekankan keterbatasan *LSTM* dalam menangkap dinamika opini politik yang kompleks, serta Kurniawan [20] yang menemukan kendala serupa dalam analisis kebijakan *Merdeka Belajar Kampus Merdeka*. Abd Karim [10] juga menyoroti bahwa meskipun *LSTM* fleksibel dalam banyak domain, model ini masih kesulitan menangkap konteks ketika sentimen bercampur dalam satu kalimat. Alghifari, Edi, & Firmansyah [11] kemudian membuktikan bahwa *BiLSTM* dapat menjadi alternatif yang lebih baik karena mampu memahami arah konteks dari dua sisi sekuensial.

Dengan demikian, keterbatasan ini membuka peluang penelitian lanjutan untuk memanfaatkan model yang lebih kontekstual seperti *IndoBERT* atau pendekatan multimodal yang menggabungkan teks dan *emoji* secara simultan guna meningkatkan akurasi deteksi sentimen dalam komentar media sosial.

4. Kesimpulan

Hasil analisis terhadap 43.664 komentar pengguna YouTube menunjukkan bahwa mayoritas respons publik terhadap program *Clash of Champions* dari Ruangguru bersifat positif 72,6%, dan 27,4% menunjukkan sentimen negatif. Model *Long Short-Term Memory (LSTM)* berhasil mengklasifikasikan komentar dengan akurasi tinggi, mencapai 90% untuk seluruh metrik *accuracy*, *precision*, *recall*, dan *F1-score*, yang menandakan efektivitas *LSTM* dalam menangani teks informal di media sosial. Penelitian ini memberikan kontribusi ilmiah dengan menerapkan NLP berbasis deep learning untuk menganalisis persepsi publik dalam konteks pendidikan digital Indonesia khususnya pada platform YouTube yang belum banyak diteliti. Temuan ini dapat menjadi dasar evaluasi konten edukatif dan strategi engagement dalam platform pembelajaran digital. Meski demikian, penelitian ini memiliki keterbatasan penggunaan model tunggal tanpa perbandingan arsitektur lain, belum disertakannya analisis demografis, serta potensi bias dalam pelabelan. Oleh karena itu, penelitian selanjutnya disarankan untuk mengeksplorasi model berbasis transformer seperti *IndoBERT* dan memperkaya analisis dengan data kontekstual dan multimodal untuk menghasilkan wawasan publik yang lebih holistik.

Daftar Pustaka

- [1] B. Widodo and D. A. Safitri, "Sentiment Analysis on Digital Learning Platforms Using Deep Learning Approaches: A Case Study of Indonesian YouTube Channels," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 11, no. 3, pp. 221–230, 2023.
- [2] R. W. N. Sari, "Penerapan Long Short-Term Memory (LSTM) untuk Analisis Sentimen Teks Berbahasa

- Indonesia," *J. Sains dan Teknol. Inf.*, vol. 9, no. 1, pp. 20–28, 2023.
- [3] A. F. F. Hudaya, "Tren Penggunaan Deep Learning untuk Analisis Sentimen: Studi Literatur pada 400 Artikel Terindeks Scopus," *J. Teknol. dan Sist. Inf.*, vol. 9, no. 1, pp. 45–53, 2023.
- [4] S. F. Handayani, R. W. Pratiwi, and dan D. I. A. Dairoh, "Analisis Sentimen pada Ulasan Twitter Menggunakan Metode Long Short-Term Memory (LSTM)," *J. RESTI (Rekayasa Sist. dan Teknol. Inf.)*, vol. 6, no. 2, pp. 384–390, 2022.
- [5] L. Farsiah, A. Misbullah, and H. Husaini, "Analisis Sentimen Menggunakan Arsitektur Long Short-Term Memory (Lstm) Terhadap Fenomena Citayam Fashion Week," *Cybersp. J. Pendidik. Teknol. Inf.*, vol. 6, no. 2, p. 86, 2022, doi: 10.22373/cj.v6i2.14687, 2022.
- [6] M. R. Putra and T. Rahman, "Sentiment Analysis of Social Media Comments in Indonesian Informal Language Using LSTM with TF-IDF Feature Engineering," *J. Inf. Syst. Informatics*, vol. 5, no. 1, pp. 77–86, 2024.
- [7] E. Suryani and A. Kurniawan, "Analisis Sentimen Komentar YouTube terhadap Konten Edukasi Menggunakan LSTM dan Word2Vec," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 9, no. 4, pp. 533–540, 2023.
- [8] R. W. G. Kusumawardana, "Analisis Sentimen pada Ulasan Aplikasi Grab Menggunakan Metode Support Vector Machine," *J. Teknol. Inf. dan Ilmu Komput. (JTIK)*, vol. X, no. X, 2021.
- [9] N. H. A. Pratama, "Pemanfaatan Arsitektur LSTM dalam Pemrosesan Bahasa Alami: Studi Struktur dan Efektivitas," *J. Ilmu Komput. dan Teknol. Inf.*, vol. 7, no. 1, pp. 66–74, 2023.
- [10] A. A. M. A. A. Abd Karim, "Sentiment Analysis on Social Media Using Long Short-Term Memory (LSTM)," *J. Inf. Syst. Informatics*, vol. 4, no. 2, pp. 145–152, 2022.
- [11] D. R. Alghifari, M. Edi, and L. Firmansyah, "Analisis Sentimen Ulasan Layanan Grab Indonesia Menggunakan Metode Bidirectional LSTM," *J. Teknol. Inf. dan Komput.*, vol. 7, no. 2, pp. 134–141, 2021.
- [12] R. L. Tatulus and L. A. Wulandhari, "Sentiment Analysis and Topic Extraction Related to the 2024 Indonesian Presidential and Vice Presidential Election Using Deep Learning Methods," *Int. J. Artif. Intell. Res.*, 2025, [Online]. Available: <http://www.ijair.id/index.php/ijair/article/view/1378>
- [13] D. A. Pratama and H. A. Nugroho, "Comparative Study of LSTM and BiLSTM for Sentiment Analysis of Indonesian YouTube Comments," *J. Teknol. dan Sist. Inf.*, vol. 10, no. 2, pp. 112–124, 2024.
- [14] P. K. Jain, V. Saravanan, and R. Pamula, "A Hybrid CNN-LSTM Model for Sentiment Analysis Using User-Generated Content," *Int. J. Eng. Technol.*, vol. X, no. X, p., 2021.
- [15] A. P. Giovani, A. Ardiansyah, T. Haryanti, L. Kurniawati, and W. Gata, "Analisis Sentimen terhadap Aplikasi Ruangguru Menggunakan Data Twitter," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. X, no. X, 2023.
- [16] A. K. NiKomang, "Analisis Sentimen Publik terhadap Tokoh Publik di Twitter dengan Teknologi Big Data," in *Seminar Nasional Teknologi Informasi dan Komunikasi (SENTIKA)*, 2022.
- [17] H. P. P. Zuriel and A. Fahrurozi, "Analisis Sentimen Menggunakan Support Vector Machine terhadap Opini Pengguna Twitter tentang PSBB," *J. Teknol. dan Sist. Komput.*, vol. X, no. X, 2020.
- [18] B. S. Rintyarna, "Analisis Sentimen Konsumsi Makanan Organik di Indonesia Selama Pandemi COVID-19," *J. Teknol. Inf. dan Komput.*, vol. X, no. X, 2021.
- [19] A. Liawati, R. Narasati, D. Solihudin, C. L. Rohmat, and S. E. Permana, "Analisis Sentimen Komentar Politik di Media Sosial Menggunakan LSTM," *J. Teknol. dan Sains Inf.*, vol. 5, no. 2, pp. 101–108, 2023
- [20] S. J. P. H. Kurniawan, "Analisis Sentimen Terhadap Kebijakan Merdeka Belajar Kampus Merdeka (MBKM) Menggunakan LSTM pada Media Sosial Twitter," *J. Ilm. Kebijak. Publik dan Sos.*, vol. 6, no. 1, pp. 55–62, 2022.