

WEB SCRAPING DAN NATURAL LANGUAGE PROCESSING MENGGUNAKAN CNN UNTUK ANALISIS SENTIMEN LINTAS PLATFORM DIGITAL

Oleh:

Arjuna Adlina Martha (221080200115)

Dosen Pembimbing : Ade Eviyanti, S.Kom., M.Kom.

Dosen Penguji 1 : Arif Senja Fitriani, S.Kom., M.Kom.

Dosen Penguji 2 : Ika Ratna Indra Astutik, S.Kom., M.T.

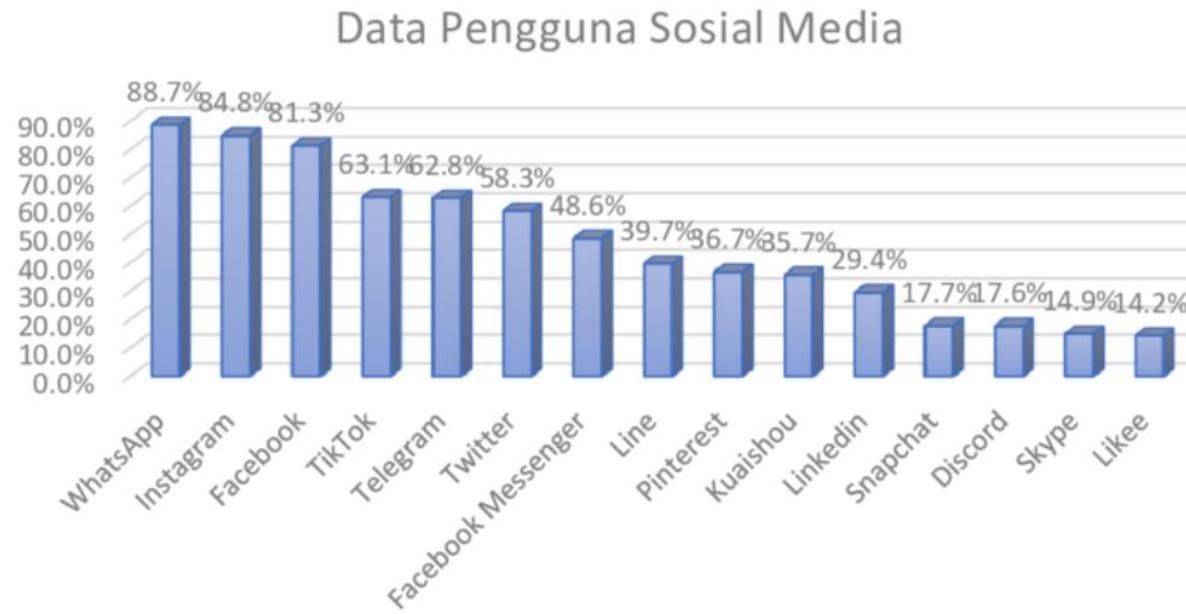
Program Studi Informatika

Universitas Muhammadiyah Sidoarjo

2026



Latar Belakang



Pertumbuhan pesat platform digital seperti media sosial, marketplace, dan platform video telah menghasilkan data opini dan ulasan dalam jumlah besar. Data ini mengandung informasi sentimen yang sangat penting untuk memahami persepsi publik terhadap produk, layanan, atau isu tertentu.

Analisis manual terhadap data dalam jumlah besar tidak efisien dan rentan bias, sehingga dibutuhkan metode otomatis untuk pengambilan (web scraping) dan analisis sentimen secara akurat dan cepat.

Salah satu metode NLP yang efektif untuk analisis sentimen adalah Convolutional Neural Network (CNN), yang terbukti mampu mengklasifikasikan sentimen positif, negatif, maupun netral pada data teks lintas platform digital.

Untuk mempermudah proses analisis sentimen, diperlukan sistem otomatis yang dapat mengekstrak data dari berbagai platform digital hanya dengan menginputkan keyword. Dengan demikian, user akan dimudahkan dalam menganalisa tanpa harus memikirkan coding.

Rumusan Masalah

- Bagaimana merancang sistem pengambilan data opini/ulasan lintas platform digital secara otomatis?
- Bagaimana menerapkan metode CNN untuk klasifikasi sentimen pada data yang diperoleh?
- Bagaimana membangun aplikasi web interaktif berbasis Next.js untuk menampilkan hasil analisis sentimen?
- Bagaimana mengambil data berita detik.com secara otomatis?



Tujuan dan Batasan Masalah

Tujuan

- Mengimplementasikan web scraping untuk pengambilan data opini/ulasan dari berbagai platform digital.
- Mengembangkan model analisis sentimen berbasis CNN untuk klasifikasi sentimen.
- Membangun aplikasi web berbasis Next.js untuk visualisasi hasil analisis sentimen secara real-time.

Batasan Masalah

- Data yang diambil berupa teks opini atau ulasan berbahasa Indonesia.
- Analisis sentimen dibatasi pada tiga kategori, yaitu sentimen positif, negatif, dan netral
- Data yang diproses hanya berupa teks, sehingga simbol, emoji, dan angka akan dihapus atau tidak diperhitungkan dalam analisis
- Pengambilan sentimen hanya dilakukan pada platform youtube.
- Pengambilan berita hanya dilakukan pada portal detik.com dan tidak diolah, hanya pengambilan data.
- Model yang digunakan untuk klasifikasi sentimen adalah Convolutional Neural Network (CNN).
- Visualisasi dan input keyword dilakukan melalui aplikasi web berbasis Next.js yang terintegrasi dengan backend Python.

Ringkasan Penelitian Terdahulu

Judul dan Peneliti	Pembahasan	Metode	Kelebihan atau Kekurangan
Analisis Sentimen Citayam Fashion Week pada Komentar YouTube dengan Metode Convolutional Neural Network (Suwito, G. A., Cholissodin, I., & Adikara, P. P., 2022)	Menganalisis sentimen masyarakat terhadap fenomena <i>Citayam Fashion Week</i> melalui komentar YouTube untuk memetakan persepsi publik (positif/negatif) dan menghindari kesalahpahaman sosial ¹ .	<i>Convolutional Neural Network (CNN)</i> , <i>Word2Vec</i> , <i>Stratified K-Fold Cross Validation</i>	Kelebihan: Tingkat akurasi sangat tinggi (97%), serta presisi dan <i>recall</i> yang seimbang, membuktikan efektivitas CNN dengan <i>Word2Vec</i> untuk fenomena sosial viral. Kekurangan: Fokus penelitian berat pada optimasi parameter model (filter size, learning rate), namun belum membahas implementasi sistem secara <i>real-time</i> atau antarmuka pengguna akhir.
Penggunaan Convolutional Neural Network untuk Analisis Sentimen Opini Lingkungan Hidup Kota Depok di Twitter (Parameswari, P. L., & Prihandoko, 2022)	Mengevaluasi opini masyarakat mengenai kondisi dan pengelolaan lingkungan hidup di Kota Depok (banjir, sampah, udara) menggunakan data dari Twitter sebagai alternatif evaluasi kebijakan ³ .	<i>Convolutional Neural Network (CNN) 1D</i> , <i>CRISP-DM</i> , Kerangka kerja <i>Flask</i> (Python)	Kelebihan: Telah berhasil mengembangkan prototipe aplikasi berbasis web (<i>deployment</i>) menggunakan Flask untuk memudahkan pengguna melakukan analisis ⁴ . Kekurangan: Akurasi model sebesar 86% ⁵ , lebih rendah dibanding penelitian sejenis. Model kesulitan mengenali sentimen kelas "Netral" (<i>recall</i> rendah 72%) dibandingkan kelas Negatif ⁶ .
Implementasi Algoritma Convolutional Neural Network Untuk Analisis Sentimen Bacapres 2024 Pada Kolom Komentar Youtube Mata Najwa (Saputra, D. E., & Isnain, A. R., 2024)	Mengklasifikasikan sentimen publik terhadap calon presiden 2024 dari tayangan "Mata Najwa" untuk memahami dinamika opini politik yang fluktuatif ⁷ .	<i>Convolutional Neural Network (CNN)</i> , <i>SMOTE (Oversampling)</i> , <i>Adam Optimizer</i>	Kelebihan: Menggunakan teknik SMOTE untuk menyeimbangkan data, sehingga prediksi tidak bias ke kelas mayoritas ⁸ . Akurasi tinggi (91%) mengungguli metode SVM pada studi terdahulu ⁹ . Kekurangan: Tantangan dalam menangani dinamika bahasa (slang/sarkasme) yang sangat beragam dalam konteks politik masih menjadi kendala yang diakui ¹⁰ .
Analisis Sentimen Terhadap Kebijakan Penggunaan Kendaraan Listrik di Media Youtube menggunakan Metode Convolutional Neural Network (CNN) (Putri, R. A., Muhsi, & Hamzah, M. A., 2025)	Menganalisis respons dan persepsi masyarakat terhadap kebijakan pemerintah mengenai transisi ke kendaraan listrik melalui komentar YouTube ¹¹ .	<i>Convolutional Neural Network (CNN)</i> , Pembobotan <i>TF-IDF</i>	Kelebihan: Mencapai akurasi yang sangat baik (97%) dengan menggunakan pembobotan TF-IDF untuk ekstraksi fitur sebelum masuk ke arsitektur CNN ¹²¹²¹²¹² . Kekurangan: Kualitas model masih sangat bergantung pada pelabelan manual dan tahap <i>preprocessing</i> (seperti normalisasi kata) yang perlu ditingkatkan untuk mengurangi noise data ¹³ .



Analisis Gap

- Keterbatasan pada Sistem yang Bersifat *End-to-End* dan Interaktif.

Mayoritas penelitian berorientasi pada hasil akademik (akurasi model) dan prosesnya berjalan di lingkungan pengembangan. Belum ada yang menawarkan sebuah platform interaktif di mana pengguna non-teknis dapat memanfaatkan sistem dari awal hingga melihat hasil analisisnya.

- Tidak Adanya Fitur Input Otomatis dari keyword oleh Pengguna

Dalam penelitian-penelitian terdahulu, pengguna harus menyediakan dataset terlebih dahulu dalam bentuk file atau hasil scraping manual. Belum ditemukan sistem yang memungkinkan pengguna hanya input keyword dan secara otomatis sistem akan mengambil data (scraping), mengolah, dan menampilkan hasil analisis sentimen.

- Minimnya Fokus pada Visualisasi Hasil yang Mudah Dipahami.

Hasil penelitian sering kali disajikan dalam format tabel metrik (akurasi, presisi) yang ditujukan untuk kalangan akademisi. Terdapat celah dalam penyajian hasil analisis sentimen ke dalam bentuk dasbor visual yang interaktif, yang lebih mudah dipahami oleh masyarakat umum atau pemangku kepentingan untuk pengambilan keputusan.

- Dengan mengintegrasikan filter kata kunci dinamis dan rekomendasi kebijakan otomatis, penelitian ini menjembatani kesenjangan antara analisis teknis dan pengambilan keputusan praktis.

Metodologi

- Pendekatan Penelitian : Kuantitatif dengan metode analisis.
- Pengambilan Data : Pengambilan data menggunakan library scraping (misal: BeautifulSoup).
- Preprocessing Data : Cleaning, Tokenisasi, Stopword Removal, Stemming, Labeling.
- Analisis data menggunakan CNN.
- LLM Llama untuk membantu pengambilan keputusan.
- Evaluasi model : akurasi, presisi, recall, f1-score.
- Front-end untuk menampilkan hasil analisis CNN.



www.umsida.ac.id



[umsida1912](https://www.instagram.com/umsida1912)



[umsida1912](https://twitter.com/umsida1912)

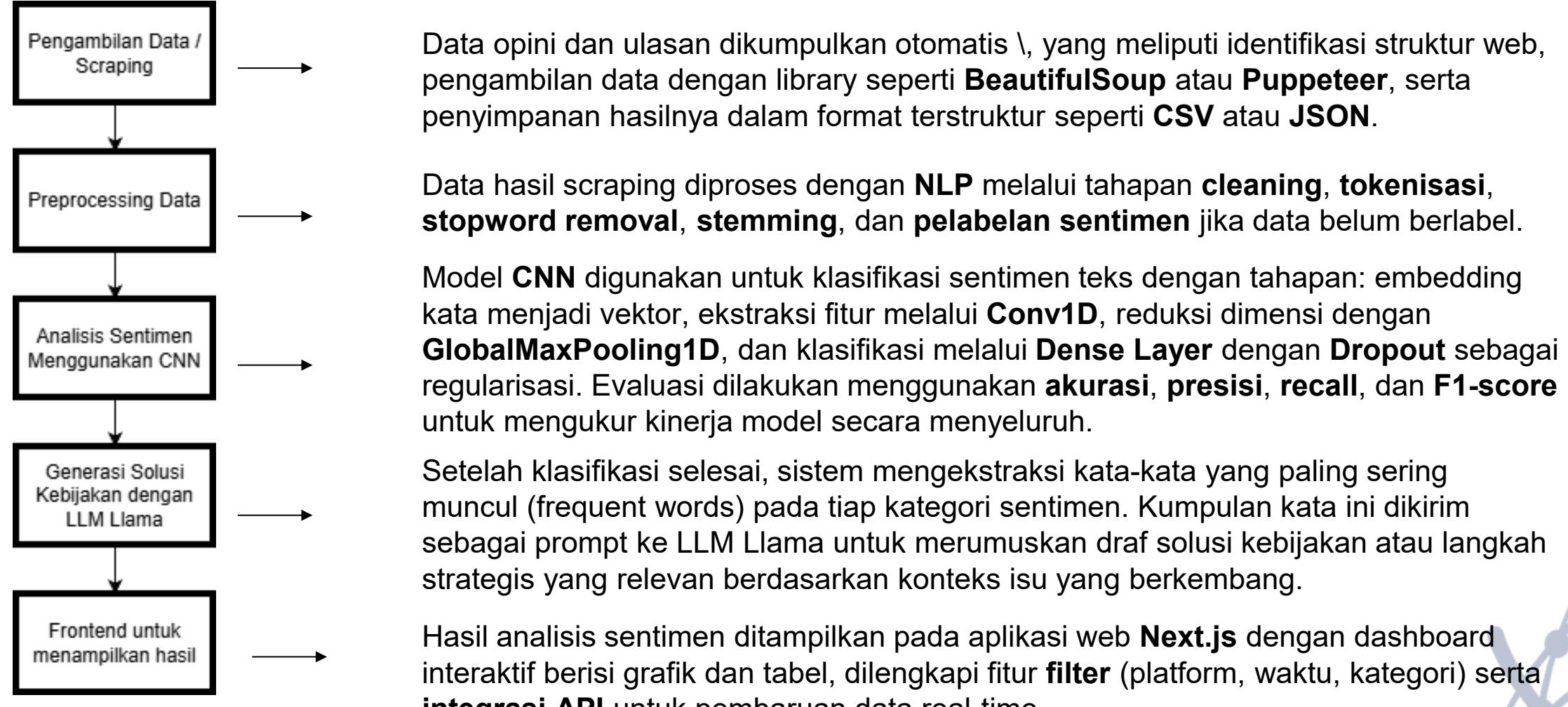


[universitas
muhammadiyah
sidoarjo](https://www.facebook.com/universitasmuhammadiyahsidoarjo)

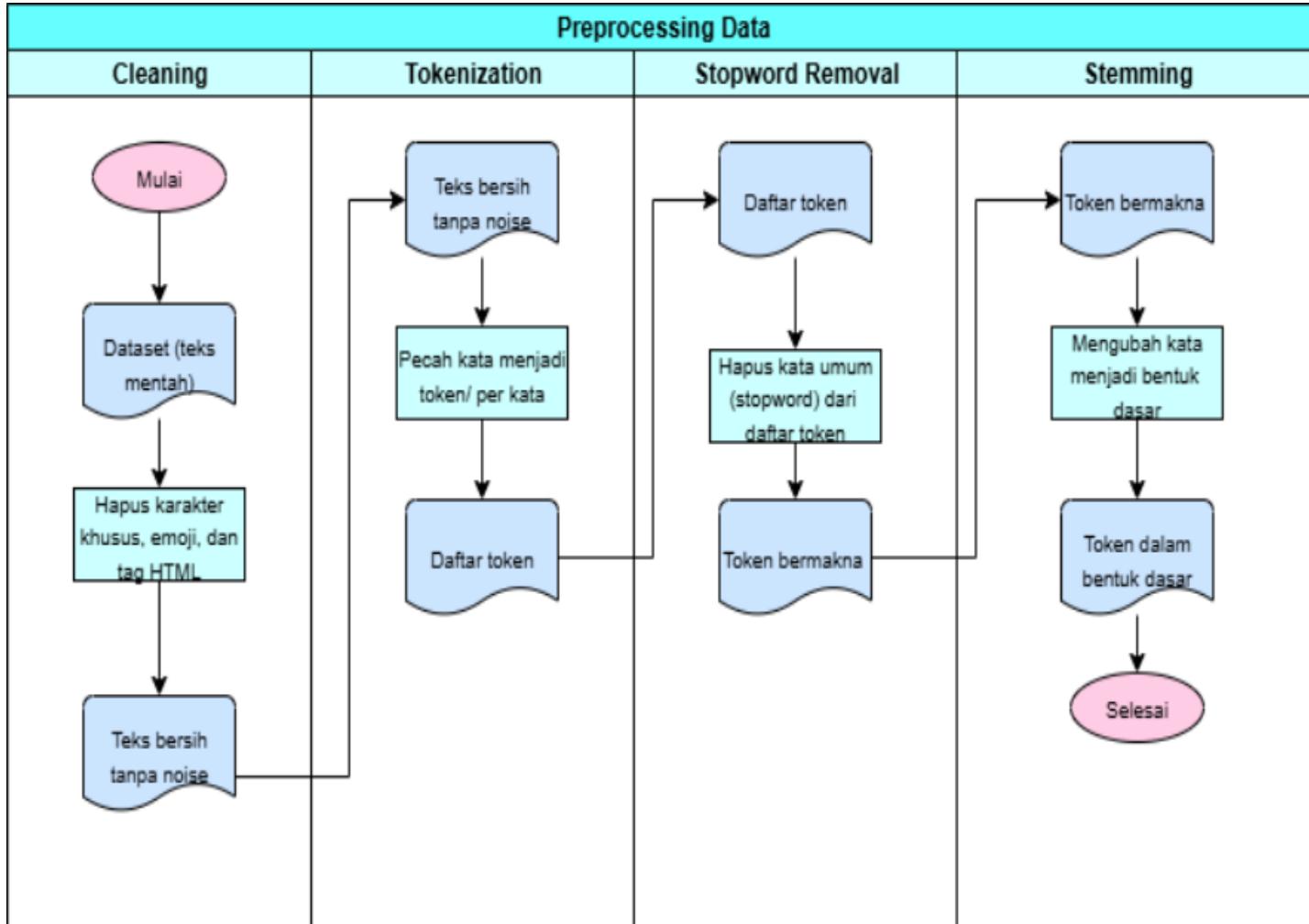


[umsida1912](https://www.youtube.com/umsida1912)

Tahap Penelitian



Tahap Penelitian



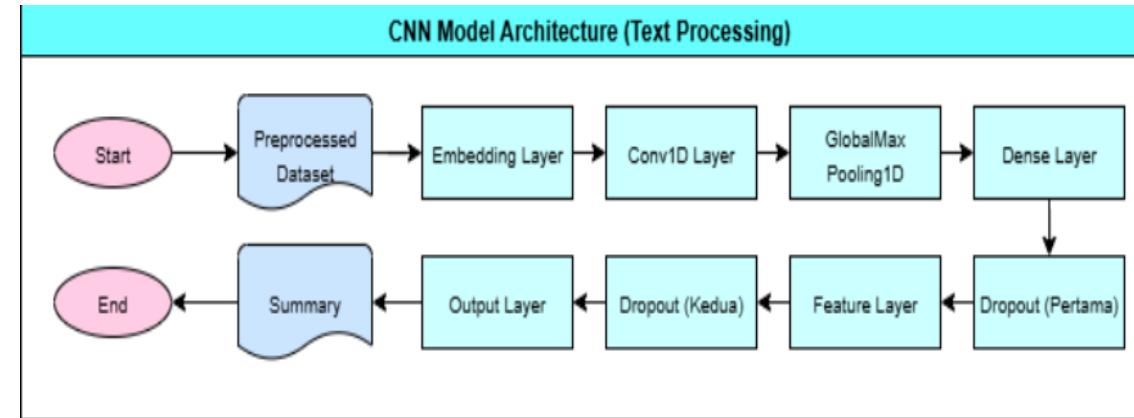
Preprocessing Data

Data hasil scraping diproses melalui beberapa tahapan NLP:

- Cleaning: menghapus karakter khusus, emoji, dan tag HTML.
- Tokenisasi: memecah kalimat menjadi token.
- Stopword removal: menghapus kata-kata umum yang tidak bermakna
- Stemming: mengubah kata ke bentuk dasar.
- Labeling: jika data belum berlabel, dilakukan pelabelan sentiment.

Tahap Penelitian

Arsitektur CNN



1. **Embedding Layer**, Mengubah token kata menjadi vektor numerik berdimensi tetap.
2. **Conv1D Layer (Convolution 1D)** Melakukan ekstraksi fitur lokal dari urutan kata, seperti mendeteksi frasa atau pola tertentu (misalnya kata "tidak bagus" → indikasi sentimen negatif)
3. **GlobalMaxPooling1D Layer** Mengambil nilai maksimum dari setiap filter hasil konvolusi, menyimpan fitur paling penting dari keseluruhan teks. Tujuannya untuk mereduksi dimensi dan mencegah overfitting.
4. **Dense Layer** yang berfungsi untuk menggabungkan fitur-fitur hasil pooling menjadi satu representasi padat.
5. **Dropout (Pertama dan Kedua)** Digunakan untuk mengurangi overfitting, dengan cara menonaktifkan sebagian neuron secara acak saat training (misalnya Dropout(0.5)).
6. **Feature Layer** ini memproses hasil dari Dense untuk membentuk fitur akhir yang akan digunakan klasifikasi.
7. **Output Layer** dengan fungsi aktivasi Softmax untuk klasifikasi multi-kelas (misalnya: positif, negatif, netral).
8. **Summary** Menampilkan total parameter dan struktur model



Tahap Penelitian

Evaluasi Model

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Metrik Evaluasi:

Performa sistem diukur menggunakan *Confusion Matrix* dengan empat parameter utama:

- 1. Accuracy:** Tingkat ketepatan keseluruhan sistem.
- 2. Precision (P):** Proporsi hasil ekstraksi yang benar.
- 3. Recall (R):** Proporsi informasi relevan yang berhasil diekstraksi.
- 4. F1-Score:** Rata-rata harmonis antara precision dan recall.

TP (True Positive): Informasi yang benar diekstraksi

TN (True Negative): Informasi yang benar tidak diekstraksi

FP (False Positive): Informasi yang salah diekstraksi

FN (False Negative): Informasi yang seharusnya diekstraksi tapi tidak



Tahap Penelitian

Generasi Solusi Kebijakan (LLM Llama)

Setelah klasifikasi selesai, sistem mengekstraksi kata-kata yang paling sering muncul (frequent words) pada tiap kategori sentimen. Kumpulan kata ini dikirim sebagai prompt ke LLM Llama untuk merumuskan draf solusi kebijakan atau langkah strategis yang relevan berdasarkan konteks isu yang berkembang.



www.umsida.ac.id



[umsida1912](https://www.instagram.com/umsida1912/)



[umsida1912](https://twitter.com/umsida1912)



universitas
muhammadiyah
sidoarjo

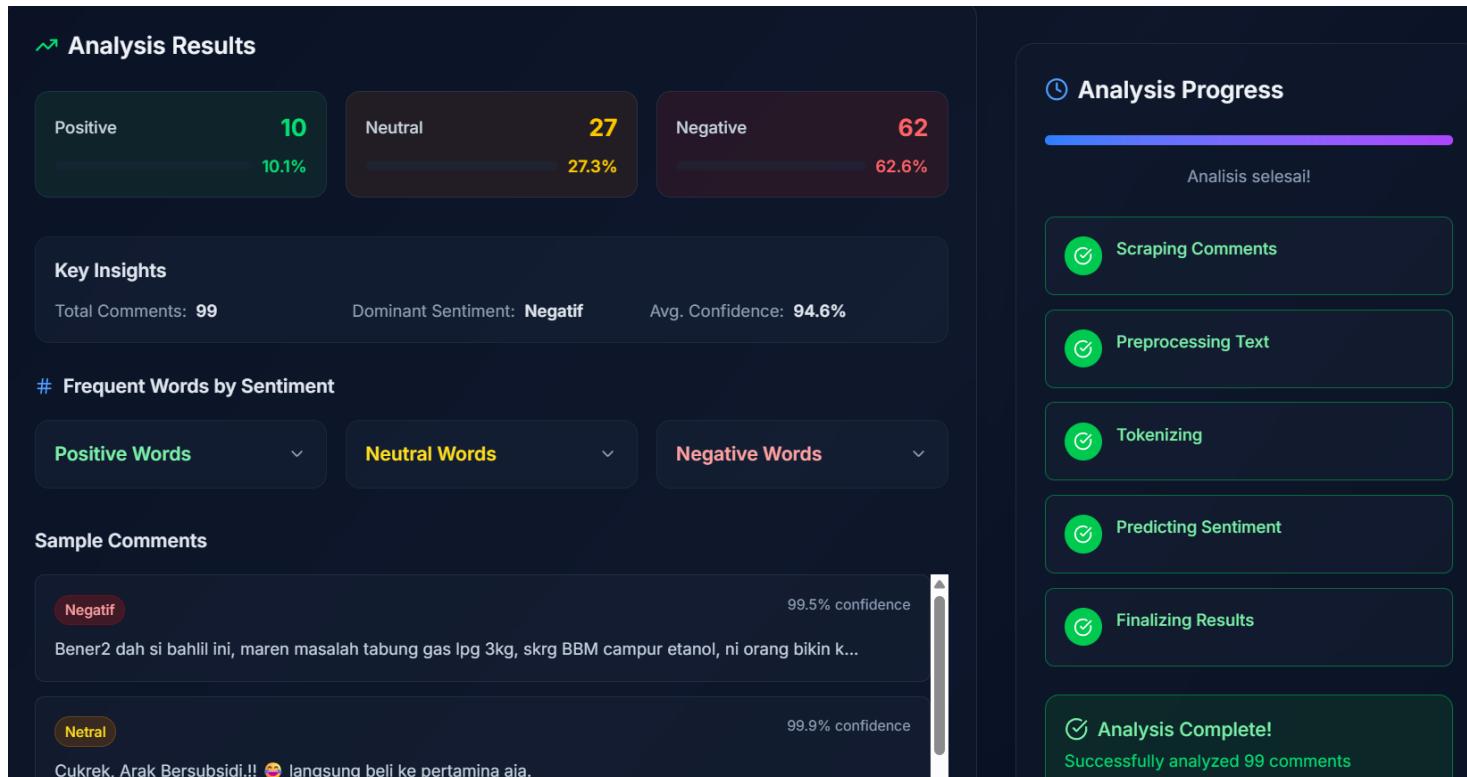


[umsida1912](https://www.youtube.com/umsida1912)



Tahap Penelitian

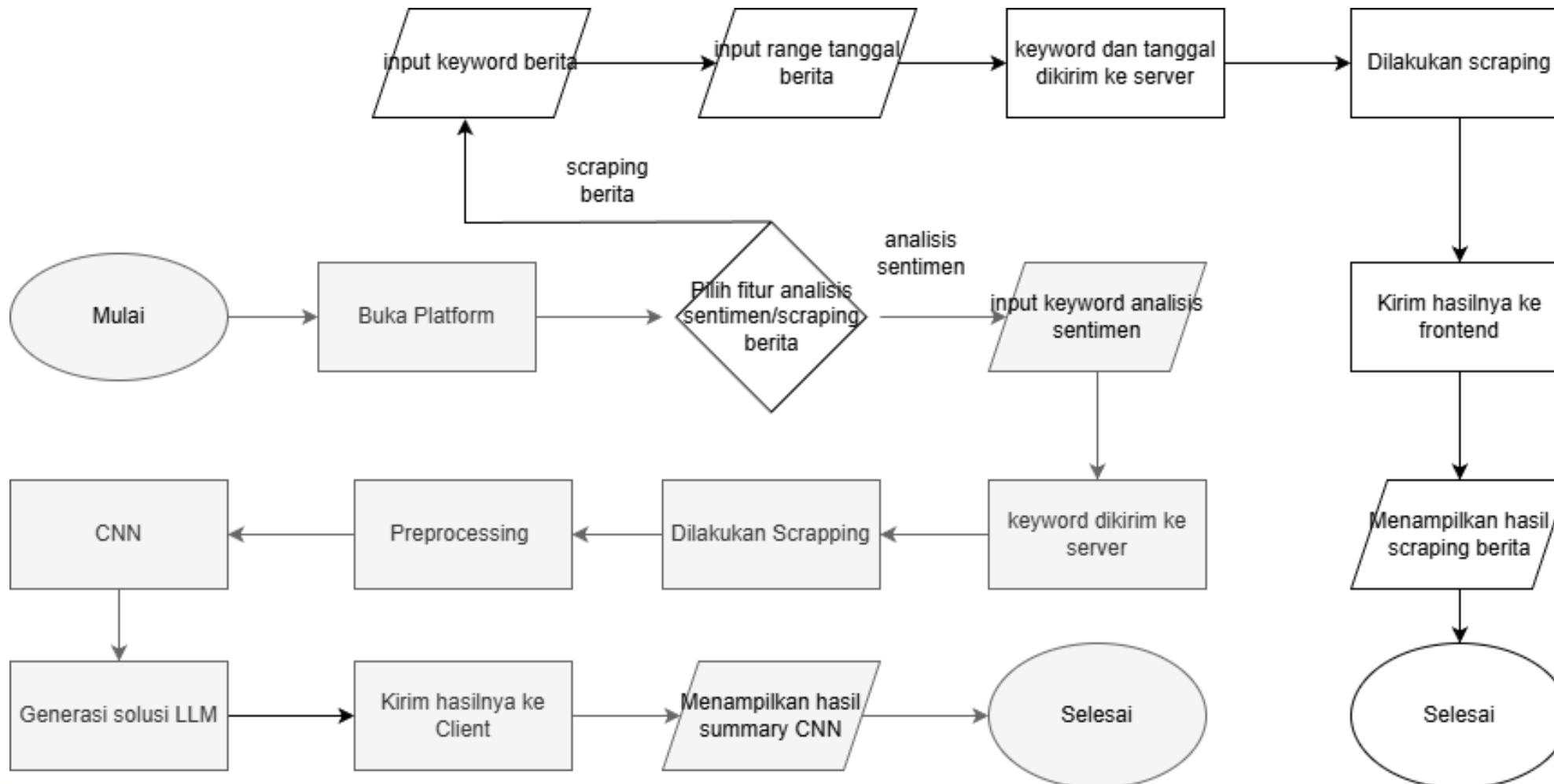
FrontEnd



Hasil analisis sentimen divisualisasikan pada aplikasi web berbasis Next.js, yang menawarkan keunggulan dalam hal rendering cepat dan kemampuan integrasi API. Fitur utama frontend meliputi:

- Dashboard visualisasi sentimen (grafik dan tabel).
- Fitur filter berdasarkan platform, waktu, atau kategori.
- Integrasi API untuk update data secara real-time.

Alur Penggunaan Sistem



Hasil dan Pembahasan

Analisis Kebutuhan

Analisis kebutuhan sistem difokuskan pada penyediaan platform end-to-end yang memungkinkan pengguna awam untuk melakukan analisis sentimen tanpa memerlukan keahlian teknis dalam pemrograman. Kebutuhan utama mencakup antarmuka web yang intuitif berbasis Next.js, di mana pengguna cukup memasukkan kata kunci untuk menginstruksikan sistem melakukan scraping secara otomatis pada 5 video teratas YouTube dan portal berita Detik.com berdasarkan parameter tanggal tertentu. Selain itu, sistem harus mampu mengolah hasil klasifikasi CNN yang bersifat teknis menjadi rekomendasi kebijakan praktis melalui integrasi LLM Llama, sehingga informasi yang dihasilkan dapat langsung dipahami dan digunakan oleh pemangku kepentingan dalam pengambilan keputusan tanpa harus berurusan dengan kerumitan kode sumber atau manajemen dataset manual.

Perancangan Sistem

Perancangan sistem menggunakan arsitektur client-server yang mengintegrasikan beberapa modul utama untuk memastikan aliran data berjalan secara otomatis dan efisien. Rancangan dimulai dengan modul antarmuka pada frontend Next.js yang berfungsi menerima input kata kunci dari pengguna, yang kemudian dikirimkan ke server untuk memicu proses scraping dinamis pada platform YouTube (20 komentar per video) dan Detik.com. Pada sisi backend, sistem dirancang untuk melakukan preprocessing dan klasifikasi menggunakan model CNN. Setelah label sentimen diperoleh, sistem mengekstraksi kata kunci dominan (frequent words) yang dikirimkan ke modul LLM Llama untuk menghasilkan rekomendasi kebijakan, sebelum akhirnya seluruh data divisualisasikan kembali ke dasbor pengguna secara real-time.



Hasil dan Pembahasan

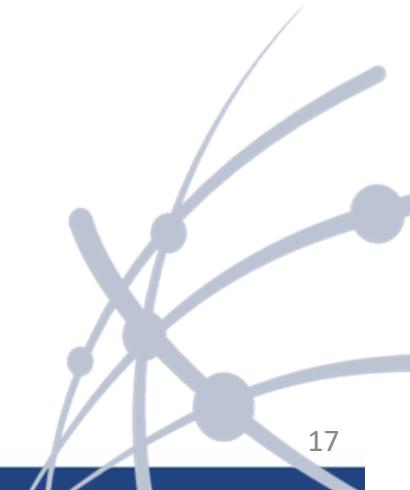
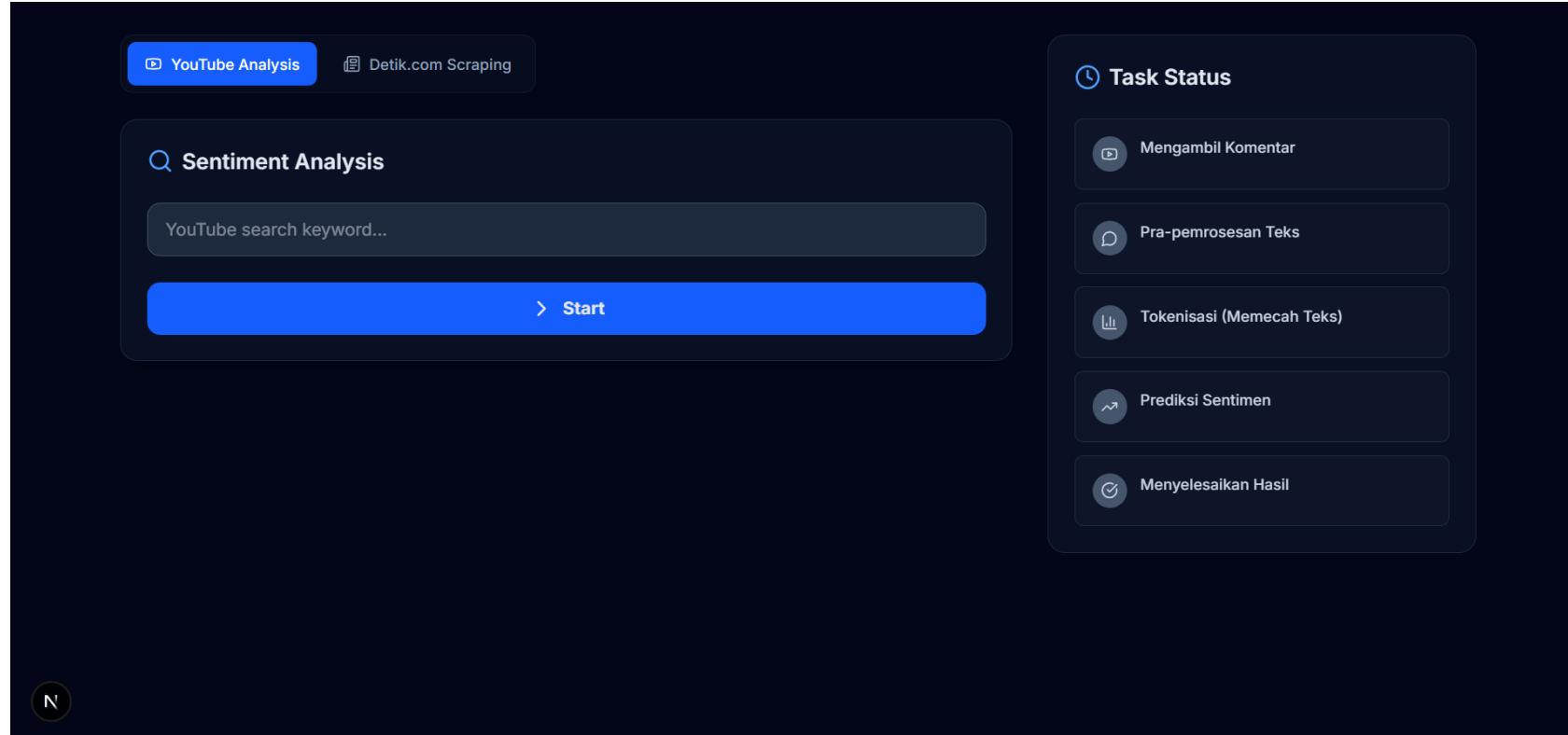
Performa Model CNN

Sentimen	Precision	Recall	F1-Score	Support
Negatif (-1)	0.78	0.76	0.77	323
Netral (0)	0.87	0.93	0.90	323
Positif (1)	0.76	0.73	0.74	316
Accuracy			0.81	962

Secara keseluruhan, distribusi data (support) yang hampir seimbang antara ketiga kelas (323, 323, dan 316) berkontribusi positif terhadap stabilitas model dalam melakukan klasifikasi

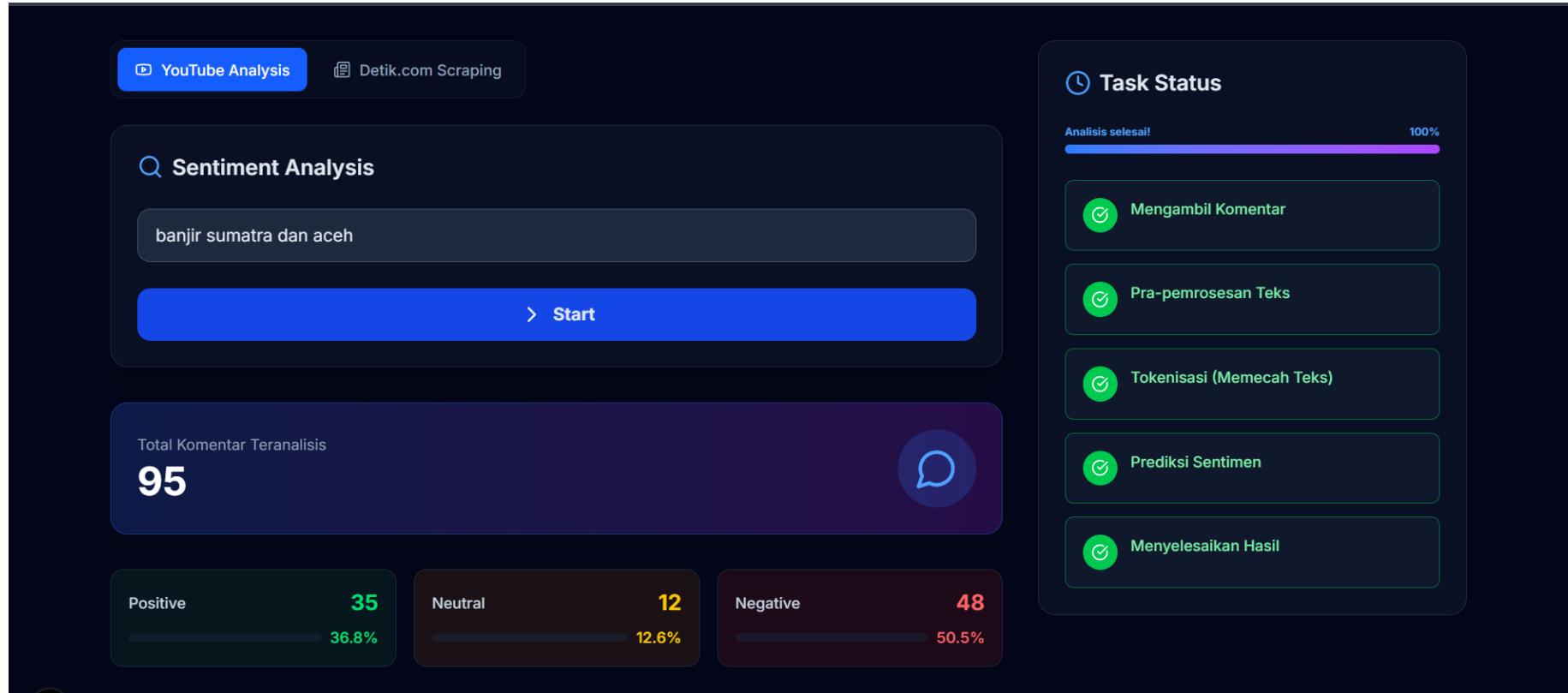
Hasil dan Pembahasan

Dashboard analisis sentimen



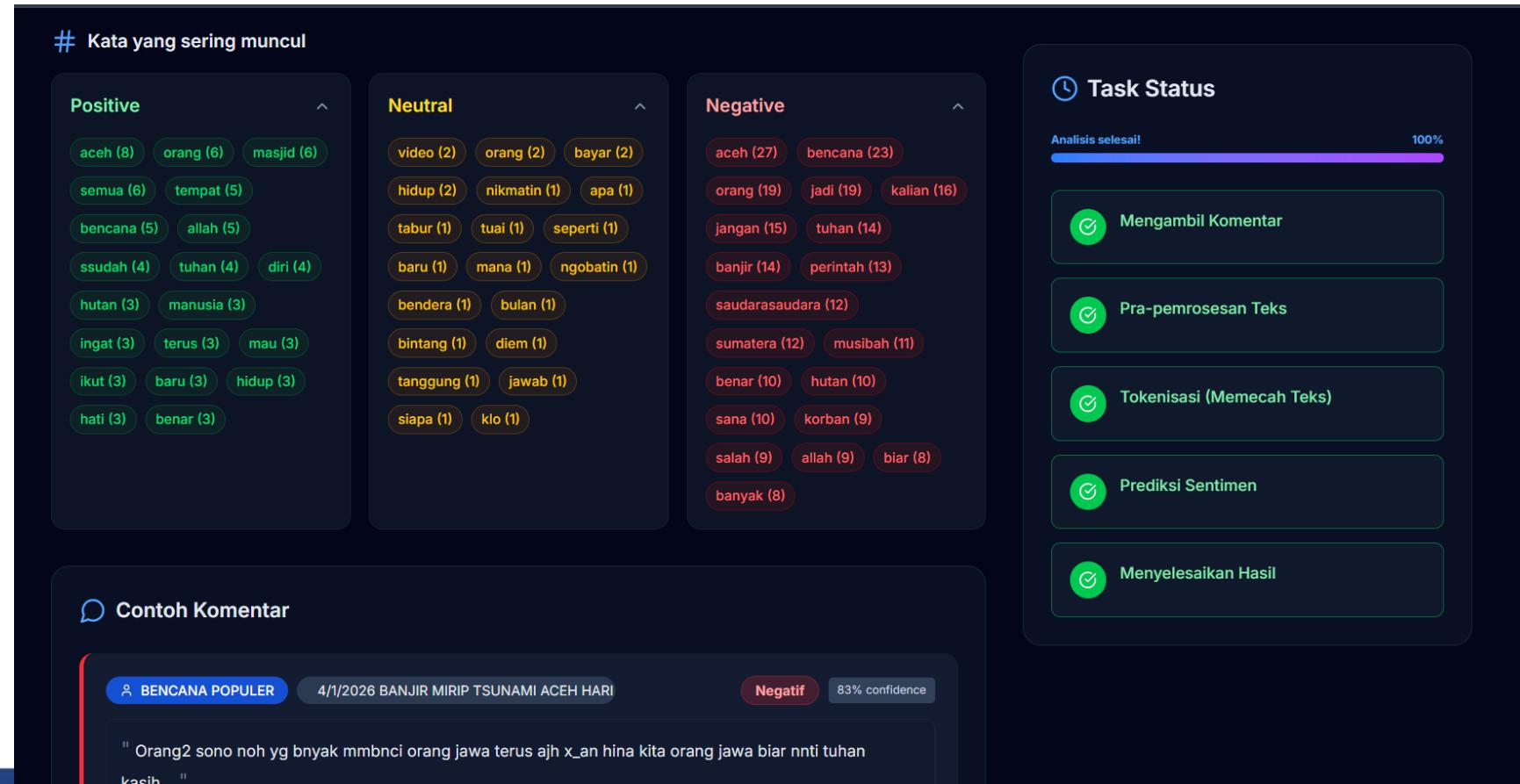
Hasil dan Pembahasan

Visualisasi analisis sentimen



Hasil dan Pembahasan

Frekuensi kata yang sering muncul



Hasil dan Pembahasan

Ringkasan LLM Llama

Ringkasan AI dan Rekomendasi Kebijakan

Ringkasan

Rekomendasi Kebijakan

Tindakan Konkret

Sentimen publik menunjukkan bahwa warga Aceh masih terkesan dengan bencana banjir dan musibah yang telah terjadi, dengan kata-kata seperti 'bencana', 'banjir', dan 'korban' yang sering muncul dalam komentar negatif. Namun, ada juga kata-kata positif seperti 'Allah', 'Tuhan', dan 'benar' yang menunjukkan bahwa warga Aceh masih memiliki harapan dan kepercayaan pada kekuatan alam.



Tokenisasi (Memecah Teks)



Prediksi Sentimen



Menyelesaikan Hasil



Hasil dan Pembahasan

Rekomendasi kebijakan LLM Llama

The diagram illustrates the AI processing pipeline for policy recommendations. It starts with a 'Ringkasan AI dan Rekomendasi Kebijakan' section containing three categories: 'MENGURANGI DAMPAK BENCANA', 'MENGEMBANGKAN INFRASTRUKTUR', and 'MENGGALANG DANA'. Below these are three boxes: 'Pra-pemrosesan Teks', 'Tokenisasi (Memecah Teks)', and 'Prediksi Sentimen', followed by a final box for 'Menyelesaikan Hasil'.

Ringkasan AI dan Rekomendasi Kebijakan

Ringkasan Rekomendasi Kebijakan Tindakan Konkret

MENGURANGI DAMPAK BENCANA
Meningkatkan pengawasan dan mitigasi bencana, serta meningkatkan kemampuan masyarakat dalam menghadapi bencana.

MENGEMBANGKAN INFRASTRUKTUR
Meningkatkan infrastruktur di daerah yang rentan bencana, seperti jembatan, jalan, dan bangunan.

MENGGALANG DANA
Menggalang dana untuk membantu korban bencana.

Pra-pemrosesan Teks

Tokenisasi (Memecah Teks)

Prediksi Sentimen

Menyelesaikan Hasil

Hasil dan Pembahasan

Rekomendasi tindakan konkret LLM Llama

Ringkasan AI dan Rekomendasi Kebijakan

Ringkasan Rekomendasi Kebijakan **Tindakan Konkret**

- Mengadakan survei dan analisis lebih lanjut untuk memahami sentimen publik lebih dalam.
- Mengembangkan program mitigasi bencana yang efektif dan efisien.
- Menggalang dana dan bantuan untuk korban bencana.

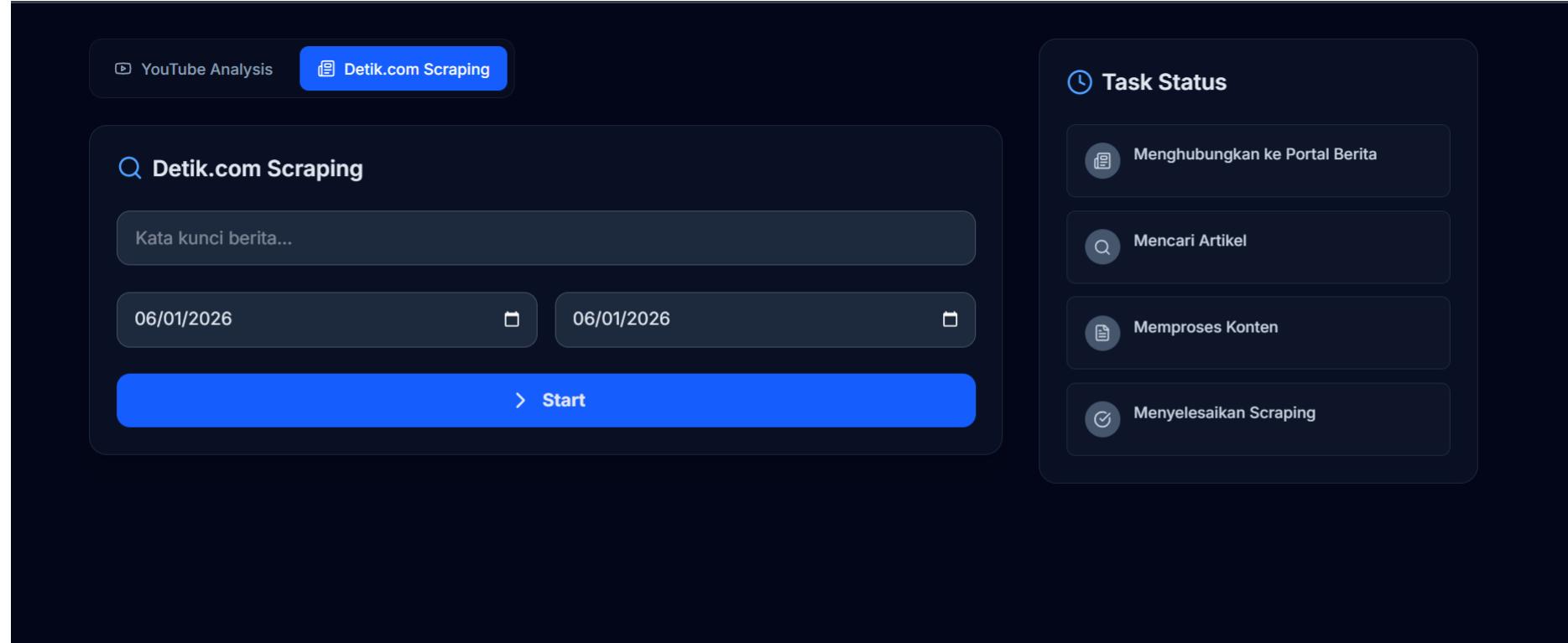
Tokenisasi (Memecah Teks)

Prediksi Sentimen

Menyelesaikan Hasil

Hasil dan Pembahasan

Antarmuka scraping berita



Hasil dan Pembahasan

Visualisasi scraping berita

The screenshot shows a user interface for news scraping. At the top, there are two tabs: "YouTube Analysis" and "Detik.com Scraping", with "Detik.com Scraping" being active. Below the tabs, a search bar contains the query "banjir sumatra". Underneath the search bar are two date inputs: "03/12/2025" and "07/12/2025". A large blue button labeled "Start" is positioned below these inputs. To the right, a "Task Status" section displays a progress bar at 100% completion with the message "Berhasil scraping undefined berita". Below the progress bar, four green status boxes are listed: "Menghubungkan ke Portal Berita", "Mencari Artikel", "Memproses Konten", and "Menyelesaikan Scraping". In the bottom left, a section titled "Berita Ditemukan (49)" shows a preview of an article titled "Warga Bantu Warga saat Banjir Sumatra, Netizen: Miris, Bangga, Haru Jadi Satu". A "Download Excel" button is located next to this section. The overall background is dark, and the interface uses a clean, modern design.



Hasil dan Pembahasan

Hasil Scraping Berita

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Judul	Tanggal Tayang	Link	Isi Berita													
2	Warga Bar -	https://ine	Dalam sepekan terakhir, limmasa media sosial dipenuhi unggahan berbagai inisiatif masyarakat yang bahu membahu mengoordinasi bermacam bantuan untuk korban														
3	Banjir Sum -	https://ine	Bencana banjir bandang di Sumatra turut meningkatkan pencarian berbagai informasi terupdate seputar peristiwa tersebut. Google Indonesia menyebutkan, tren penc														
4	Honda Kas Sabtu, 06 Des 20	https://otc	PT Honda Prospect Motor (HPM) memberikan program servis khusus konsumen yang kendaraannya terdampak bencana banjir di wilayah Aceh, Medan, dan Padang. Pe														
5	Wuling Kas Jumat, 05 Des 20	https://otc	Bencana Banjir di Sumatra jadi pusat perhatian semua elemen masyarakat termasuk pada stakeholder otomotif di Indonesia. Salah satunya Wuling Motors yang mema														
6	Ribuan Per Jumat, 05 Des 20	https://wv	Polda Jatim menggelar salat gaib dan doa bersama untuk para korban bencana alam. Khususnya, korban banjir bandang dan tanah longsor yang melanda Aceh, Sumate														
7	Aksi Solidar Minggu, 07 Des 2	https://otc	Jakarta- Komunitas motor di Jakarta menggalang dana untuk membantu korban banjir bandang dan longsor di Sumatra. Aksi ini jadi wujud kepedulian terhadap warga														
8	Tim dari Ti Sabtu, 06 Des 20	https://wv	Lima orang relawan dari Tiongkok tiba di Aceh untuk membantu mendeteksi korban bencana banjir dan longsor yang masih tertimbun. Mereka disebut memiliki alat ur														
9	Prasasti He Minggu, 07 Des 2	https://wv	Jauh sebelum teknologi beton modern mendominasi infrastruktur sipil, peradaban Jawa Kuno ternyata telah memiliki sistem mitigasi bencana yang canggih untuk men														
10	Menteri Li Sabtu, 06 Des 20	https://wv	Menteri Lingkungan Hidup/Badan Pengendalian Lingkungan Hidup (KLH/BPLH) Hanif Faisol Nurofiq mengatakan, Desa Garoga, yang terletak di Kecamatan Batang Toru														
11	Polres Bat Sabtu, 06 Des 20	https://wv	Polres Batu menggelar salat ghaib dan doa bersama untuk para korban bencana alam banjir bandang dan tanah longsor di Aceh, Sumatra Utara dan Sumatra Barat. Sa														
12	Peringatan -	https://ine	Banjir bandang yang menerjang sejumlah wilayah di Sumatra pekan lalu kembali mengingatkan bahwa pulau ini memang punya sejarah panjang terkait bencana alam./														
13	Kisah Perju Minggu, 07 Des 2	https://ne	Korban banjir dan longsor di Aceh, Sumatera Utara dan Sumatera Barat, melakukan berbagai cara untuk bertahan hidup. Ada yang harus kabur dan mengungsi ke teng														
14	Ini Profil B Sabtu, 06 Des 20	https://wv	Bupati Aceh Selatan Mirwan MS menjadi sorotan karena berangkat umrah di tengah bencana yang menerjang wilayahnya. Mirwan sudah dicopot dari Ketua DPC Gerir														
15	Viral Air Da -	https://ine	Fenomena air di Bendungan Kawasan Ombilin dari Danau Singkarak, Sumatra Barat, berwarna jernih bikin geger warga TikTok. Airnya biru kehijauan. Berbagai unggaha														
16	Kemensos Minggu, 07 Des 2	https://ne	Kementerian Sosial (Kemensos) melalui Sentra Darussaadah Aceh berkolaborasi dengan Dinas Sosial Aceh menyelenggarakan layanan dukungan psikososial (LDP) bagi														
17	Ria Ricis U -	https://wv	Ria Ricis mengungkapkan fakta memilukan yang ia temui di lapangan. Meski sebagian pemberitaan menyebut banjir mulai surut, ia menegaskan keadaan di lokasi masih														
18	Bupati Ace Sabtu, 06 Des 20	https://wv	Bupati Aceh Selatan Mirwan MS keluar daerah untuk melaksanakan umrah di tengah kondisi bencana alam yang menimpak daerahnya. Pemerintah Kabupaten menyebu														
19	Gubsu Bob Minggu, 07 Des 2	https://wv	Gubernur Sumatera Utara (Gubsu) Bobby Nasution menyebutkan kerugian akibat banjir bandang-longsor di Sumut mencapai Rp 9,98 triliun. Jumlah itu mencakup keru														
20	Video Deti 8,406 Views	https://20	Banjir lahar Gunung Semerum menghantam permukiman warga di Desa Jugosari, Kecamatan Candipuro, Lumajang, Minggu (7/12). Material banjir lahar lalu merendam ru														
21	Korban Ba Minggu, 07 Des 2	https://wv	Korban banjir bandang di Sumatra rencananya akan direlokasi ke hunian sementara (Huntara). Rumah tersebut rencananya akan dibangun dalam waktu enam bulan s														
22	Utang KUF Minggu, 07 Des 2	https://fin	Presiden Prabowo Subianto mengumumkan akan menghapus utang kredit usaha rakyat (KUR) bagi petani yang terdampak bencana banjir dan longsor di Sumatera. Pen														
23	Andien Sia Minggu, 07 Des 2	https://ho	Penyanyi Andien Aisyah mempersiapkan diri secara fisik dan mental sebelum berangkat ke daerah banjir dan longsor di Sumatera. Ia diladangkan berangkat pada Rahu /														



Kesimpulan dan Saran

Penelitian ini berhasil membangun sistem analisis sentimen lintas platform yang mengintegrasikan pengambilan data otomatis dari komentar YouTube dan portal berita Detik.com. Implementasi metode *Convolutional Neural Network* (CNN) menunjukkan performa yang sangat baik dengan tingkat akurasi mencapai 81% dalam mengklasifikasikan sentimen menjadi kategori positif, negatif, dan netral. Mekanisme pengambilan data berbasis kata kunci terbukti memberikan fleksibilitas tinggi bagi pengguna awam, sementara fitur scraping berita secara otomatis mampu mengekstraksi informasi dalam rentang waktu tertentu dan menyediakannya dalam format tabel yang terstruktur. Selain itu, integrasi LLM Llama memberikan nilai tambah signifikan dengan mengubah tren kata kunci menjadi rekomendasi kebijakan otomatis, sehingga menjembatani celah antara analisis data teknis dan pengambilan keputusan praktis.



Kesimpulan dan Saran

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran untuk pengembangan selanjutnya. Saran pertama adalah melakukan *fine-tuning* pada LLM Llama menggunakan dataset kebijakan publik spesifik guna meningkatkan ketajaman, relevansi, dan detail dari rekomendasi solusi yang dihasilkan. Selanjutnya, disarankan untuk memperluas jangkauan scraping berita ke berbagai portal berita nasional lainnya untuk mendapatkan cakupan data opini publik yang lebih luas dan komprehensif. Selain itu, perlu ditambahkan fitur deteksi ulasan duplikat atau *bot* pada proses scraping guna meningkatkan kualitas dataset yang akan diolah oleh model klasifikasi. Terakhir, pengembangan sistem sebaiknya mengintegrasikan rekomendasi kebijakan dan tindakan konkret dengan hukum yang berlaku di Indonesia.



Referensi

- [1] G. A. Suwito, I. Cholissodin, and P. P. Adikara, "Analisis Sentimen Citayam Fashion Week pada Komentar YouTube dengan Metode Convolutional Neural Network," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 6, no. 12, pp. 5948–5956, 2022, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [2] D. E. Saputra and A. R. Isnain, "Implementasi Algoritma Convolutional Neural Network Untuk Analisis Sentimen Bacapres 2024 Pada Kolom Komentar Youtube Mata Najwa," *JIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 9, no. 3, pp. 1431–1441, 2024, doi: 10.29100/jipi.v9i3.5420.
- [3] Y. F. Qitfirul Dwi Cahyono, Ade Eviyanti , Metatia Intan Mauliana, "Perancangan Sistem Informasi Layanan Surat Menyurat dan Pengaduan Masyarakat Berbasis WEB," *J. Technol. Syst. Inf.*, vol. 1, no. 1, pp. 86–102, 2024.
- [4] R. C. Rivaldi, T. D. Wismarini, J. T. Lomba, and J. Semarang, "Analisis Sentimen Pada Ulasan Produk Dengan Metode Natural Language Processing (NLP) (Studi Kasus Zalika Store 88 Shopee)," *Elkom*, vol. 17, no. 1, pp. 120–128, 2024.
- [5] M. Ujaran et al., "Implementasi Convolutional Neural Network (C NN) Untuk," vol. 14, no. 2, pp. 314–325, 2024.
- [6] P. L. Parameswari and Prihandoko, "Penggunaan Convolutional Neural Network Untuk Analisis Sentimen Opini Lingkungan Hidup Kota Depok Di Twitter," *J. Ilm. Teknol. dan Rekayasa*, vol. 27, no. 1, pp. 29–42, 2022, doi: 10.35760/tr.2022.v27i1.4671.
- [7] G. Tamara and Kemas Muslim L, "Sentiment Analysis on Acute Kidney Syrup Videos Using CNN and LSTM Algorithms," *Int. J. Inf. Commun. Technol.*, vol. 9, no. 2, pp. 53–65, 2023, doi: 10.21108/ijocit.v9i2.818.
- [8] A. Maulana, D. Dyantono, and R. E. Putra, "Perbandingan Sent2vec TF-IDF Logistic Regression dan Word2vec CNN pada hasil Sentiment Analysis Youtube Comment," *J. Informatics Comput. Sci.*, vol. 05, pp. 63–72, 2023, [Online]. Available: <https://ejournal.unesa.ac.id/index.php/jinacs/article/view/54621%0Ahttps://ejournal.unesa.ac.id/index.php/jinacs/article/download/54621/43435>
- [9] M. Fariz, S. Lazuardy, and D. Anggraini, "Modern Front End Web Architectures with React.Js and Next.Js," *Int. Res. J. Adv. Eng. Sci.*, vol. 7, no. 1, pp. 132–141, 2022.
- [10] H. L. Yuzeфа and A. Eviyanti, "Perancangan Perpustakaan Digital Berbasis Web dengan Pendekatan Development Life Cycle," *Indones. J. Appl. Technol.*, vol. 1, no. 2, p. 22, 2024, doi: 10.47134/ijat.v1i2.3042.
- [11] A. Z. Pratama, A. Marinta, B. Triyudanto, M. Saman, and T. N. Fatyanosa, "Retrieval-Augmented Generation for Indonesian Criminal Law Information Using the LLaMA Model," *Innov. Informatics Artif. Intell. Res.*, vol. 1, no. 1, pp. 35–41, 2025, [Online]. Available: <http://doi.org/10.35718/iiair.v1i1.1306https://journal.itk.ac.id/index.php/IIAIRAvailable:https://doi.org/10.35718/iiair.v1i1.1306>
- [12] R. A. Putri and M. A. Hamzah, "Analisis Sentimen Terhadap Kebijakan Penggunaan Kendaraan Listrik di Media Youtube menggunakan Metode Convolutional Neural Network (CNN)," vol. 11, no. 1, pp. 182–192, 2025.
- [13] A. Fatma and P. Indah, "Analisis Sentimen Komentar YouTube MV K-Pop Menggunakan Naïve Bayes dan SVM : Studi Kasus Jung Jaehyun ' Horizon ,'" vol. 15, no. 2, pp. 383–395, 2025.
- [14] S. Khairunnisa, A. Adiwijaya, and S. Al Faraby, "Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19)," *J. Media Inform. Budidarma*, vol. 5, no. 2, p. 406, 2021, doi: 10.30865/mib.v5i2.2835.
- [15] E. Y. Hidayat and D. Handayani, "Penerapan 1D-CNN untuk Analisis Sentimen Ulasan Produk Kosmetik Berdasarkan Female Daily Review," *J. Nas. Teknol. dan Sist. Inf.*, vol. 8, no. 3, pp. 153–163, 2023, doi: 10.25077/teknosi.v8i3.2022.153-163.
- [16] S. T. Andini, A. Eviyanti, H. Setiawan, and C. Taurusta, "Analisis Sentimen Pengguna Aplikasi Tantan: Perbandingan Kinerja Metode Naive Bayes dan SVM Sentiment Analysis of Tantan Application Users: A Performance Comparison Between Naive Bayes and SVM," vol. 15, no. 2, pp. 396–407, 2025.
- [17] M. A. Naufal Dzaki, H. Hindarto, A. Eviyanti, and Nuril Lutvi Azizan, "Analisis Sentimen Layanan Pelanggan Provider Internet dengan Algoritma Support Vector Machine dan Naive Bayes," *Semant. Tek. Inf.*, vol. 11, no. 1, pp. 84–93, 2025, doi: 10.55679/semantik.v1i1.127.



DARI SINI PENCERAHAN BERSEMI